

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS

Departamento de Bioquímica y Biología Molecular I



TESIS DOCTORAL

**NMR insights into molecular recognition: structure and interactions of
peptides and proteins**

**Investigaciones sobre reconocimiento molecular mediante RMN :
estructura e interacciones de péptidos y proteínas**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Héctor Zamora Carrera

Directoras

M^a Ángeles Jiménez López
Marta Bruix Bayés

Madrid, 2017



UNIVERSIDAD
COMPLUTENSE
MADRID

**NMR INSIGHTS INTO MOLECULAR RECOGNITION:
STRUCTURE AND INTERACTIONS OF PEPTIDES AND PROTEINS**

**INVESTIGACIONES SOBRE RECONOCIMIENTO MOLECULAR
MEDIANTE RMN: ESTRUCTURA E INTERACCIONES DE
PÉPTIDOS Y PROTEÍNAS**

Memoria con la que
Héctor Zamora Carreras
aspira al Grado de Doctor

Tesis dirigida por:
Dra. M^a Ángeles Jiménez López
Dra. Marta Bruix Bayés

Tutor:
Dr. Álvaro Martínez del Pozo

Departamento de Bioquímica y Biología Molecular I
Facultad de Ciencias Químicas (UCM)

Departamento de Química Física Biológica
Instituto de Química Física Rocasolano (CSIC)

MADRID 2016



Por y para todos los míos.

"The ancient teachers of this science," said he, "promised impossibilities, and performed nothing. The modern masters promise very little; they know that metals cannot be transmuted, and that the elixir of life is a chimera. But these philosophers, whose hands seem only made to dabble in dirt, and their eyes to pore over the microscope or crucible, have indeed performed miracles. They penetrate into the recesses of nature, and show how she works in her hiding places. They ascend into the heavens: they have discovered how the blood circulates, and the nature of the air we breathe. They have acquired new and almost unlimited powers; they can command the thunders of heaven, mimic the earthquake, and even mock the invisible world with its own shadows."

Frankenstein; or, The Modern Prometheus

Chapter 3

Mary Wollstonecraft Shelley

Plasmar en negro sobre blanco los sentimientos de gratitud asociados a este período dedicado a mi formación como investigador y a la construcción de una tesis doctoral, se me antoja un maravilloso ejercicio de reflexión y, al fin y al cabo, de justicia. Me temo, en cambio, que la justicia no será más que parcial, pues me sería imposible siquiera enumerar todas las personas que de alguna manera han influido indirectamente en todo esto.

No quiero dejar de mencionar aquí la influencia de nuestros tiempos, tan complicados para muchas personas. Es sobradamente conocido que crisis es igual a oportunidad y que, en momentos así, se suelen tomar decisiones que ni siquiera se considerarían en otras circunstancias, buscando siempre que la desesperación no se torne en desesperanza. Sin embargo, al margen de las crisis que nos son impuestas, deberíamos también declarar nuestras pequeñas revoluciones de vez en cuando, tomar decisiones y atravesar nuestra burbuja de confort, pues al otro lado todo puede suceder. En definitiva, debemos recorrer nuevos caminos, y decidir. Avanzar sin miedo a equivocarse, pues la única decisión errónea es la que no se toma. Esta tesis es fruto de una decisión.

Durante esta intensa etapa de formación ha habido dos pilares fundamentales cuyos nombres son M^a Ángeles y Marta. Ellas confiaron en mí y en mis capacidades desde el primer momento y sólo espero haberles devuelto al menos una pequeña fracción de todo lo que ellas me han dado. Durante estos años he sentido que formábamos un equipo muy bien engranado y complementado, y para mí son un modelo de trabajo, esfuerzo y organización. Es una suerte haber podido compartir estos años todo tipo de vivencias y haber tejido una cercana amistad que ha trascendido la, en ocasiones, fría relación entre maestro y pupilo.

Durante estos años he podido disfrutar de la compañía de grandes científicos, cuyas opiniones y comentarios han sido de gran utilidad para dar una vuelta de tuerca adicional a los resultados obtenidos durante mi tesis. Carlos, Doug, Jorge, Jose, Irene, Miguel Treviño, Flor, no sólo han sido compañeros de trabajo y colegas científicos, sino gente con la que compartir anécdotas y risas. Y más allá del “Rocasolano”, he tenido la inmensa suerte de colaborar con y recibir la ayuda de personas extraordinarias, gracias a quienes todo el trabajo recogido en esta tesis ha sido posible: Jesús Sanz y Beatriz Maestro de la Universidad Miguel Hernández de Elche; David Andreu y su gente de la Universitat Pompeu Fabra de Barcelona; María Torres, Mayte Villalba, Rosalía Rodríguez, Javier Merino, Nacho Rodríguez–Crespo, Álvaro Martínez del Pozo, Begoña García del departamento BBMI de la Universidad Complutense de Madrid; Charo González del IQM de Madrid...

It was a great pleasure and a privilege for me to work in the group of Prof. Anne S. Ulrich, in the Institute for Biological Interfaces (KIT). My time in Karlsruhe is a beautiful memory. I worked hard there and I learnt a lot in a very short time, thanks to all the people who nicely helped me, especially Erik and Anne, but also Jochen, Parvesh, Phillip, Bianca, Andrea, Kerstin, Marina, Stephan and Markus. I want to say thank you also to all my mates, who made me feel like home: Sezgin, Papia, Benjamin, Marie-Claude, Xu... I feel particularly grateful to Tamta and Marco, who were incredibly kind to me and facilitated my adaptation to the German way of life.

Ni puedo, ni debo olvidarme de algunas personas hacia las cuales sentí simpatía, respeto y admiración y que, desgraciadamente, ya no están con nosotros. Sus nombres aquí plasmados son un tributo a su legado: Manolo, Alfredo y Noé.

Guardo un especial recuerdo de todos los camaradas becarios y contratados *erremeneros* con los que he compartido trabajo y solaz: Santi, Miguel Garavís, Soraya, Diana, Vero, Angélica, Ismael, Belén y Sergio. Y por supuesto, otras muchas personas que habitan o han habitado el edificio Rockefeller, como Javi, Mon, María, M^a Angélica, Palma... Y muchos más que me dejaré en el tintero.

De la playa a la montaña y de una punta a otra de España, hay dos personas que durante estos años han sido algo más que unos excelentes colegas. Es un privilegio encontrar en dos compañeros de trabajo a dos amigos y compañeros de aventuras y, al mismo tiempo, dos científicos con mayúsculas. Cuántas veces escucharos hablar ha reavivado en mí la llama de la curiosidad científica, las ganas de trabajar, de tomar la iniciativa y la inspiración. David y Miguel, nunca dejaré de aprender de vosotros, me habéis ayudado a crecer como investigador.

Esta tesis, como todo cuando hago en la vida, sin mi gente no valdría nada, ni podría haberse realizado. Mis amigos no pueden dejar de estar presentes en esta etapa de mi vida. De toda la vida o de hace poco, de cerca o de lejos, de la guardería o de la universidad, todos ellos ponen de su parte para que todo sea mejor. Todos, parecidos y diferentes, tienen siempre sus manos, su comprensión, su empatía y su capacidad para escuchar dispuestas. Da igual el tiempo transcurrido, los kilómetros de distancia, me enseñáis una y otra vez que aún queda gente buena y leal, y que la risa siempre tiene su lugar. Vosotros ya sabéis quiénes sois. Gracias por estar ahí.

Y al final, la familia. Todo empezó ahí. Esta tesis refleja parte de lo que soy, y de lo que soy, sin duda, tiene una enorme responsabilidad mi familia. Puedo decir sin tapujos que en mis padres puedo encontrar probablemente dos de las mejores personas que jamás haya conocido, y que posiblemente conoceré. Ellos me dieron las herramientas, y he intentado usarlas lo mejor posible, pues cada acción de mi vida intenta homenajear todo el esfuerzo que han puesto siempre en darnos a mi hermana y a mí todo aquello que pudiéramos necesitar. De mi familia he cogido gran parte de lo que soy, y en mí siempre estáis presentes, tanto los que sois, como los que fueron, incluido esa persona medio amigo, medio hermano que sobre sus cuatro patas aún sigue viniendo a jugar conmigo en sueños.

Y tú, Sara. Si estos años he aprendido algo, ha sido sin duda gracias a ti. Contigo he aprendido un nuevo significado de la palabra “familia”. He aprendido que el hogar no es un lugar geográfico, que puede viajar contigo y estar en cualquier parte. He aprendido que hay que aprovechar cada instante, que hay que relativizar y, sobre todo, que hay que estar decidido a ser feliz. Amigos, familia, compañeros de aventuras, gente de la que aprender... Si hay un hallazgo a destacar en mi carrera es el de haber encontrado a una persona que aúne todo eso, y haber tenido la fortuna de formar un equipo perfecto con ella. Tu contribución en todo esto ha sido determinante, sin ti remando en el otro lado, esta barca sólo se movería en círculos.

TABLE OF CONTENTS

Table of contents	i
List of figures	vii
List of tables	xiii
List of acronyms	xv
Summary	xix
CHAPTER 1. INTRODUCTION	1
1. Biomolecular interactions and recognition	3
1.1 Mechanistic theories for molecular recognition in proteins	3
1.2 Non-covalent interactions	4
1.2.1 Electrostatic interactions	4
1.2.2 Van der Waals forces	5
1.2.3 π -effects	6
1.2.4 Hydrophobic effect	6
2. Interactions with biological membranes	7
2.1 Cell membranes: A model	7
2.2 Components of cell membranes	8
2.3 Importance of protein/peptide-membrane interactions	9
2.4 Model membranes	10
3. Protein-carbohydrate interactions	12
3.1 Classes of GBPs	13
3.2 The nature of the protein-carbohydrate recognition	14
3.3 Importance of protein-carbohydrate interactions	14
4. Protein-protein interactions (PPIs)	15
4.1 Classes of PPIs: Protein complexes	15
4.2 Protein-binding domains and interaction interface	16
4.3 PPIs in biological functions	17
5. Characterization of biomolecular interactions	17
5.1 Theoretical approaches	17
5.2 Experimental approaches	18
5.3 Spectroscopic methods to study protein interactions	19
5.4 Investigating protein interactions by NMR	21
5.4.1 Information obtained from NMR experiments	22
5.4.2 Solution NMR to study protein interactions	22
5.4.3 Protein interactions by solution NMR: Limitations	23
5.4.4 Solid-state NMR to study peptide-membrane interactions	24
6. Aims and structure of this thesis	24
Bibliography	27
CHAPTER 2. SWITCH PEPTIDES	33
INTRODUCTION	35
1. Choline-binding proteins (CBPs)	35
1.1 Pneumococcal autolysin LytA	36
1.2 Applicability of the study of LytA	40
1.3 Peptides derived from LytA	40
2. Chameleon sequences	41
3. Conformational switches	43
3.1 Features of the conformational switching	43
3.2 Classes of conformational switches	44

3.3 Relevance of the conformational switching	46
OBJECTIVES	49
MATERIALS AND METHODS	51
1. Peptide synthesis	51
2. C–LytA domain purification	52
3. Preparation of small unilamellar vesicles (SUVs)	52
4. Circular dichroism (CD)	53
5. NMR sample preparation	55
6. NMR spectra acquisition	55
7. NMR spectra assignment	56
8. Structure calculation	56
9. Fluorescence measurements	56
10. Effect of paramagnetic compounds on NMR spectra	57
RESULTS AND DISCUSSION	59
1. Structural characterization of LytA ₂₃₉₋₂₅₂ and LytA ₂₅₉₋₂₇₂	59
1.1 CD study of LytA ₂₃₉₋₂₅₂ and LytA ₂₅₉₋₂₇₂ in aqueous solution	59
1.2 NMR structure calculations in aqueous solution	62
1.3 Effect of DPC micelles	66
1.4 Characterization of the interaction with DPC micelles	72
1.5 Effect of TFE	76
1.6 Effect of negatively–charged SDS micelles	77
1.7 Behaviour in the presence of lipid vesicles	80
2. Structural characterization of LytA ₂₃₉₋₂₇₂	81
2.1 CD study in aqueous solution and interaction with choline	81
2.2 NMR structure calculation in aqueous solution	84
2.3 Effect of DPC and SDS micelles	86
2.4 Characterization of the interaction with DPC micelles	91
2.5 Effect of TFE	91
2.6 Behaviour in the presence of lipid vesicles	93
3. Effect of DPC micelles on the full–length C–LytA module	93
4. Are DPC micelles general helix inducers?	95
5. Study of LytA ₂₃₉₋₂₅₂ variants designed to investigate the β –to– α transition	96
5.1 Design of LytA ₂₃₉₋₂₅₂ variants	96
5.2 CD and NMR study in aqueous solution	98
5.3 CD and NMR study in DPC micelles	103
5.4 Fluorescence measurements on K3W5–LytA ₂₃₉₋₂₅₂ and W5K10–LytA ₂₃₉₋₂₅₂	106
6. Discussion	108
6.1 LytA ₂₃₉₋₂₅₂ , LytA ₂₅₉₋₂₇₂ , and LytA ₂₃₉₋₂₇₂ have native–like β –hairpin structure and maintain ability to bind choline with low affinity	108
6.2 Determinants of β –hairpin stability in aqueous solution	109
6.3 LytA ₂₃₉₋₂₅₂ , a novel switch peptide	112
6.4 Other LytA switch peptides	114
6.5 Determinants of α –helix formation in DPC micelles	116
CONCLUSIONS	119
BIBLIOGRAPHY	121
CHAPTER 3. MEMBRANE–ACTIVE PEPTIDES (MAPs)	129
GENERAL INTRODUCTION	131
1. Membrane–active peptides (MAPs)	131
2. Antimicrobial peptides (AMPs)	131
2.1 Classification and origin of the AMPs	132

2.2 AMPs cell–targeting and mechanisms of action	134
2.3 Structure–function relationships (SAR) of AMPs	137
2.4 Antibiotic resistance and clinical importance of AMPs	138
3. Anticancer peptides (ACPs)	140
3.1 Factors contributing to ACPs specificity	140
3.2 Structure and modes of action of ACPs	141
3.3 Development of therapeutic ACPs	142
4. Cell–penetrating peptides (CPPs)	143
4.1 Classification of the CPPs	143
4.2 Mechanisms of uptake	144
4.3 Conjugated cargoes	145
4.4 Clinical relevance of CPPs	145
5. Other membrane–active peptides	146
5.1 Antiviral peptides (AVPs)	147
5.2 Antifungal peptides	147
5.3 Antiparasitic peptides (APPs)	147
5.4 Immunomodulatory peptides	147
5.5 Fusogenic peptides (FPs)	148
5.6 Amyloid peptides	148
6. MAPs studied in this thesis	148
PART 1. BP100, a short peptide with good antimicrobial properties	149
INTRODUCTION	151
1. Origin of BP100	151
1.1 Melittin	151
1.2 Cecropin A	152
1.3 Cecropin A–melittin hybrids	152
2. Previous studies on BP100	153
OBJECTIVES	155
MATERIALS AND METHODS	157
1. Materials	157
2. Peptide synthesis	157
3. Minimum inhibitory concentration assays	158
4. Haemolysis assays	158
5. Circular dichroism (CD) spectropolarimetry	159
6. NMR sample preparation	160
7. Solid–state NMR	161
8. NMR data analysis	161
RESULTS AND DISCUSSION	163
1. Antimicrobial and haemolytic activity of BP100 analogues	163
2. CD characterization of BP100 analogues	165
3. Solid–state NMR study of BP100 analogues in lipid systems	166
4. Orientation of BP100 in lipid bilayers	172
5. Discussion	175
CONCLUSIONS	183
PART 2. Crotalicidin, a rattlesnake venom cathelicidin with antimicrobial and antitumour activity	185
INTRODUCTION	187
1. Cathelicidins	187
2. Viperidins	188
OBJECTIVES	189
MATERIALS AND METHODS	191

1. Bioinformatic analyses	191
2. Peptide synthesis	191
3. Circular dichroism	192
4. NMR spectroscopy	193
5. NMR structure calculation	193
6. Bacterial strains and antibacterial assays	194
7. Cell culture	194
8. Peptide cytotoxicity	194
9. Haemolytic activity	195
10. Serum stability	195
RESULTS AND DISCUSSION	197
1. Structural dissection of Ctn, Ctn ₁₋₁₄ and Ctn ₁₅₋₃₄ fragments	197
2. Circular dichroism	197
3. NMR structures of Ctn and derived peptides in DPC micelles	198
4. Antimicrobial activity	203
5. Antitumour activity	203
6. Toxicity to eukaryotic cells	205
7. Serum stability	206
8. Discussion	207
CONCLUSIONS	211
BIBLIOGRAPHY	213
 CHAPTER 4. CARBOHYDRATE-BINDING MODULES (CBMs)	 225
INTRODUCTION	227
1. β -1,3-glucanases	227
1.1 β -1,3-glucans	227
1.2 Functions of the β -1,3-glucanases	228
1.3 Classification of the β -1,3-glucanases	229
1.4 Plant endo- β -1,3-glucanases	230
2. Carbohydrate-binding modules (CBMs)	230
2.1 Classification and structure of the CBMs	231
2.2 Carbohydrate-CBM interaction	233
3. Ole e 9 and Fra e 9: Two homologue endo- β -1,3-glucanases	235
3.1 Ole e 9	235
3.2 Fra e 9	235
3.3 CtD-Ole e 9 and CtD-Fra e 9: Two homologue CBMs with different ligand affinity	236
3.4 A model β -1,3-glucan: Laminarin	237
OBJECTIVES	239
MATERIAL AND METHODS	241
1. rCtD-Ole e 9 and rCtD-Fra e 9 production and purification	241
2. N-glycosylation prediction analysis	241
3. Carbohydrate-binding assays	241
4. NMR spectroscopy and spectral assignment	241
5. Titration with laminarin monitored by NMR	242
6. NMR dynamics	242
7. Analytical ultracentrifugation	243
8. rCtD-Fra e 9 modelling and molecular docking	244
RESULTS AND DISCUSSION	245
1. N-glycosylation prediction analysis	245

2. NMR spectra assignment, secondary structure and global fold of rCtD–Fra e 9	246
3. Carbohydrate binding activity of rCtD–Fra e 9	248
4. Testing the interaction of rCtD–Fra e 9 and rCtD–Ole e 9 with laminarin by NMR	250
5. Hydrodynamic data of free rCtD–Ole e 9 and rCtD–Fra e 9 and their complexes with laminarin	251
6. Molecular docking for rCtD–Ole e 9 and rCtD–Fra e 9 with laminarin	254
7. Discussion	256
CONCLUSIONS	261
BIBLIOGRAPHY	263
 CHAPTER 5. CYTOSKELETAL MOTORS	 269
INTRODUCTION	271
1. Molecular motors	271
2. Cytoskeletal motors	271
2.1 Components of the cytoskeleton	271
2.2 Classes of cytoskeletal motors	274
2.3 Mechanisms of cytoskeletal motors	275
2.4 Functions of cytoskeletal motors	276
2.5 Membrane association of cytoskeletal motors	278
3. Components of the dynein complex	279
3.1 Dynein intermediate chains (DICs)	280
3.2 Dynein light chains (LICs)	281
3.2.1 DYNLL family	281
3.2.2 DYNLRB family	283
4. DYNLT family	283
4.1 Structure of DYNLT' proteins	284
4.2 Functions of DYNLT' proteins	285
4.3 Ligand binding to DYNLT'	286
OBJECTIVES	289
MATERIALS AND METHODS	291
1. Materials	291
2. Constructs	291
3. Yeast–two hybrid screen and β –galactosidase assay	292
4. Recombinant protein expression and purification	292
5. Isothermal titration calorimetry (ITC)	292
6. Cell transfection and immunofluorescence	293
7. Fluorescence polarization assays	293
8. NMR experiments	293
9. Structure calculation	294
10. NMR dynamics	294
11. Circular dichroism measurements	295
12. Luciferase reporter assay	295
13. Cell transfection and pulldown assays	295
14. Pepscan experiments	295
RESULTS AND DISCUSSION	297
1. The polybasic sequence motifs of mammalian dynein intermediate chains are not required for DYNLT1 binding	297
2. NMR solution structure of human DYNLT1 in complex with DIC	298

3. 3D structure by solution NMR: Analysis of the dimer interface and canonical binding groove	302
4. Identification of novel DYNLT1–interacting proteins capable of occupying the canonical hydrophobic groove	305
5. DYNLT1 interacts with the intracellular domain of activating receptor IIB	307
6. The binding site of Lfc within DYNLT1 extends well–beyond the canonical binding groove	309
7. Other DYNLT1–interacting partners might adopt a similar way of interaction	312
8. Discussion	313
CONCLUSIONS	317
BIBLIOGRAPHY	319
 CHAPTER 6. SUMMARY AND CONCLUDING REMARKS	 325
 APPENDICES	 331
TABLES	333
FIGURES	375

LIST OF FIGURES

Figure	Description	Page
Chapter 1		
1.1	Classical molecular recognition models for proteins.	3
1.2	Scheme of the “conformational selection model”.	4
1.3	Fluid mosaic membrane model (S–N model).	7
1.4	Main membrane models used for the study of protein/peptide–membrane interactions.	12
1.5	Scheme of the different approaches to characterise biomolecular interactions.	21
Chapter 2		
2.1	Representation of the peptidoglycan layer.	35
2.2	X–ray crystallographic structure of pneumococcal LytA.	38
2.3	Choline–binding sites in the CBD of pneumococcal autolysin LytA.	39
2.4	Schematic representation of peptide LytA _{197–210} .	41
2.5	Chameleon sequences.	42
2.6	Scheme of the different kinds of conformational switches.	46
2.7	CD spectra for LytA _{239–252} and LytA _{259–272} aqueous samples.	59
2.8	Thermal unfolding of LytA _{239–252} and LytA _{259–272} monitored by near–UV CD.	60
2.9	Thermal unfolding of peptides LytA _{239–252} and LytA _{259–272} monitored by CD, in the presence and the absence of choline.	61
2.10	CD–monitored titration of LytA _{239–252} with choline.	62
2.11	Affinity of LytA _{259–272} for choline by fluorescence spectroscopy.	62
2.12	NMR data for LytA _{239–252} and LytA _{259–272} .	64
2.13	Structures of LytA _{239–252} and LytA _{259–272} in aqueous solution.	65
2.14	Far–UV CD spectra of LytA _{239–252} and LytA _{259–272} in the presence of DPC micelles.	66
2.15	Titration of LytA _{239–252} with DPC monitored by CD.	66
2.16	Thermal unfolding of LytA _{239–252} in the presence of 30 mM DPC.	67
2.17	Determination of the critical micelle concentration (cmc) of DPC.	67
2.18	Effect of DPC below cmc on LytA _{239–252} and LytA _{259–272} peptides monitored by NMR.	68
2.19	Effect of DPC above cmc on LytA _{239–252} and LytA _{259–272} peptides monitored by NMR.	70
2.20	Structures of LytA _{239–252} and LytA _{259–272} in DPC micelles.	71
2.21	Effect of paramagnetic compounds on LytA _{239–252} in the presence of DPC micelles.	73
2.22	Effect of hydro–soluble paramagnetic MnCl ₂ on LytA _{259–272} in the presence of DPC micelles.	73
2.23	Characterization of the interaction between LytA _{239–252} and DPC micelles by NMR.	74
2.24	Characterization of the interaction between LytA _{259–272} and DPC micelles by NMR.	74
2.25	Intrinsic fluorescence of LytA _{239–252} .	75

2.26	Cartoon representation of a hypothetical model for LytA ₂₃₉₋₂₅₂ /DPC micelle complex.	76
2.27	Bar plots of $\Delta\delta H_\alpha$ as a function of sequence for the peptides LytA ₂₃₉₋₂₅₂ and LytA ₂₅₉₋₂₇₂ in aqueous solution and in the presence of TFE.	77
2.28	Effect of SDS on peptides LytA ₂₃₉₋₂₅₂ and LytA ₂₅₉₋₂₇₂ .	78
2.29	Structures of LytA ₂₃₉₋₂₅₂ and LytA ₂₅₉₋₂₇₂ in SDS micelles.	79
2.30	Characterization of the interaction between LytA ₂₃₉₋₂₅₂ and SDS micelles by NMR.	80
2.31	Far-UV CD spectra of LytA ₂₃₉₋₂₅₂ and LytA ₂₅₉₋₂₇₂ in the presence of lipid vesicles.	81
2.32	Far- and near-CD spectra of LytA ₂₃₉₋₂₇₂ .	82
2.33	Thermal unfolding of peptide LytA ₂₃₉₋₂₇₂ monitored by CD in the presence of choline.	83
2.34	Affinity of LytA ₂₃₉₋₂₇₂ for choline by fluorescence spectroscopy.	83
2.35	Affinity of LytA ₂₃₉₋₂₇₂ for choline by fluorescence spectroscopy (data fitting).	84
2.36	NMR data for LytA ₂₃₉₋₂₅₂ .	85
2.37	Structure of LytA ₂₃₉₋₂₇₂ in aqueous solution.	86
2.38	Far-UV CD spectra of LytA ₂₃₉₋₂₇₂ in the presence of DPC and SDS micelles.	87
2.39	Effect of DPC and SDS above cmc on LytA ₂₃₉₋₂₇₂ peptide monitored by NMR.	88
2.40	NOE summaries for LytA ₂₃₉₋₂₇₂ in DPC and SDS micelles.	89
2.41	Structure of LytA ₂₃₉₋₂₇₂ in SDS micelles.	90
2.42	Effect of hydro-soluble paramagnetic MnCl ₂ on LytA ₂₃₉₋₂₇₂ in the presence of DPC micelles.	91
2.43	Far-UV CD spectra of LytA ₂₃₉₋₂₇₂ in the presence of 30 % vol. TFE.	92
2.44	Bar plots of $\Delta\delta H_\alpha$ as a function of sequence for peptide LytA ₂₃₉₋₂₇₂ in aqueous solution and in the presence of TFE.	92
2.45	Far-UV CD spectra of LytA ₂₃₉₋₂₇₂ in the presence of SUVs.	93
2.46	Effect of detergent micelles on the full-length C-LytA module.	94
2.47	Effect of temperature on DPC-induced C-LytA conformational changes.	95
2.48	Bar plots of $\Delta\delta H_\alpha$ as a function of sequence for the peptides SESYV11 and SESYW11 in aqueous solution and in DPC micelles.	96
2.49	Structural schemes for peptide LytA ₂₃₉₋₂₅₂ and designed variants in aqueous solution and in the presence of detergent micelles.	97
2.50	Far-UV CD spectra of LytA ₂₃₉₋₂₅₂ variants in aqueous solution.	98
2.51	Bar plots of $\Delta\delta H_\alpha$ as a function of sequence for peptides K3W5-LytA ₂₃₉₋₂₅₂ and W5K10-LytA ₂₃₉₋₂₅₂ in aqueous solution.	99
2.52	NOESY spectra of K3W5-LytA ₂₃₉₋₂₅₂ and W5K10-LytA ₂₃₉₋₂₅₂ in aqueous solution.	100
2.53	Bar plots of $\Delta\delta H_\alpha$ as a function of sequence for peptides S3S10-LytA ₂₃₉₋₂₇₂ , I3V10-LytA ₂₃₉₋₂₅₂ , and I5Y6T11T13-LytA ₂₃₉₋₂₅₂ in aqueous solution.	101
2.54	NOESY spectra of S3S10-LytA ₂₃₉₋₂₅₂ and I3V10-LytA ₂₃₉₋₂₅₂ in aqueous solution.	102
2.55	NOESY spectrum of I5Y6T11T13-LytA ₂₃₉₋₂₅₂ in aqueous solution	103
2.56	Far-UV CD spectra of LytA ₂₃₉₋₂₅₂ variants in DPC micelles.	104
2.57	Bar plots of $\Delta\delta H_\alpha$ as a function of sequence for LytA ₂₃₉₋₂₅₂ variants in DPC micelles.	105
2.58	NOE summaries for LytA ₂₃₉₋₂₅₂ variants in DPC micelles.	105

2.59	Effect of DPC below cmc on LytA _{239–252} variants monitored by NMR.	106
2.60	Intrinsic fluorescence of LytA _{239–252} , K3W5–LytA _{239–252} , and W5K10–LytA _{239–252} .	107
2.61	Fluorescence quenching experiments for K3W5–LytA _{239–252} , and W5K10–LytA _{239–252} .	107
2.62	Normalized fluorescence spectra for LytA _{239–252} , K3W5–LytA _{239–252} , and W5K10–LytA _{239–252} , in the presence of 30 mM DPC.	108
2.63	β -hairpin schemes and α -helical wheel representations for LytA _{239–252} and LytA _{259–272} in aqueous solution and in detergent micelles.	112
2.64	Chameleon sequence identified in LytA _{239–252} .	114
2.65	Chameleon sequence identified in LytA _{259–272} .	115

Chapter 3

Cover	NMR structure of anoplin R5W in DPC micelles (PDB code: 2MJR)	129
3.1	Examples of the AMPs classes by sequence and structure.	133
3.2	Schematic representation of the cell membrane structure of Gram-positive and Gram-negative bacteria.	134
3.3	Carpet model of AMPs.	135
3.4	Barrel-stave model of AMPs.	136
3.5	Toroidal-pore model of AMPs.	136
3.6	Mechanisms of action of ACPs.	142
3.7	Structure of melittin.	151
3.8	Helical wheel projection of BP100.	157
3.9	Definition of τ and ϱ angles.	162
3.10	Haemolysis of BP100–PP and Ala-d3 substituted analogues.	164
3.11	CD spectra of BP100 and Ala-d3 substituted analogues.	165
3.12	³¹ P NMR spectra of oriented samples of lipids with BP100K5A.	167
3.13	² H NMR spectra of oriented samples of BP100K5A.	168
3.14	³¹ P and ¹⁵ N NMR spectra of oriented samples of BP100.	170
3.15	² H NMR spectra of oriented samples of BP100.	171
3.16	Analysis of ² H NMR data of BP100 in POPC:POPG (3:1) bilayers.	173
3.17	Analysis of ² H NMR data of BP100 in DMPC:DMPG (3:1) bilayers.	174
3.18	Analysis of ² H NMR data of BP100 in DMPC:DMPG:lyso-MPC (1:1:1) bilayers.	175
3.19	Orientation and dynamical behaviour of BP100 in lipid bilayers.	178
3.20	Definition of peptide orientation and meaning of the angular parameters.	179
3.21	Schematic representation of a cathelicidin holoprotein.	187
3.22	Example of the structural families of cathelicidins.	188
3.23	Helical-wheel projections for Ctn and its fragments.	193
3.24	CD spectra of Ctn and its Ctn _{1–14} and Ctn _{15–34} fragments in aqueous buffer.	199
3.25	Bar plots of $\Delta\delta H_\alpha$ and $\Delta\delta C_\alpha$ as a function of sequence for Ctn and its Ctn _{1–14} and Ctn _{15–34} in DPC micelles.	201
3.26	Summary of NOE signals observed for Ctn and Ctn _{1–14} .	202
3.27	Ctn structures in DPC micelles.	204
3.28	Local structure around Pro33 in Ctn and Ctn _{15–34} .	204
3.29	Viability of HeLa and leukaemia cells upon treatment with Ctn and fragments for 4 h.	206
3.30	Relative cell viabilities of tumour cells after treatment with Ctn and fragments for 24 h.	207

3.31	Toxicity of Ctn and fragments to eukaryotic cells.	208
3.32	Time-course of the treatment of Ctn, Ctn _{1–14} , and Ctn _{15–34} with human serum and.	208

Chapter 4

Cover	Crystal structure of xylan-binding module CBM15 (PDB code: 1GNY)	225
4.1	Chemical structure of the general repetitive unit of β -1,3-glucans.	227
4.2	Examples of CBMs structures.	232
4.3	Sequence alignment for rCtD-Ole e 9 and rCtD-Fra e 9.	236
4.4	Representation of the chemical structure of the repetitive unit of laminarin.	237
4.5	Assignment of the ^{15}N - ^1H -HSQC spectrum of rCtD-Fra e 9.	247
4.6	C_α and C_β conformational shift plots for rCtD-Ole e 9 and rCtD-Fra e 9.	247
4.7	rCtD-Fra e 9 and rCtD-Ole e 9 structures.	248
4.8	Carbohydrate binding assays with rCtD-Ole e 9 and rCtD-Fra e 9.	249
4.9	Calculation of K_d for laminarin complexes.	249
4.10	Titration of rCtD-Fra e 9 and rCtD-Ole e 9 with laminarin monitored by NMR.	250
4.11	Chemical shift perturbations of rCtD-Fra e 9 and rCtD-Ole e 9 upon the addition of laminarin.	251
4.12	Mapping of the chemical shift perturbations on rCtD-Fra e 9 and rCtD-Ole e 9 surfaces.	251
4.13	Velocity and equilibrium sedimentation experiments.	253
4.14	Driven docking models for rCtD-Ole e 9 and rCtD-Fra e 9 with laminarin.	255
4.15	Sequence alignment of rCtD-Ole e 9 and rCtD-Fra e 9 with homologous sequences found in the Uniprot database.	259

Chapter 5

Cover	Crystal structure of human cytoplasmic dynein 2 in complex with ADP.Vi (PDB code: 4RH7)	269
5.1	Actin assembly scheme.	272
5.2	Microtubule assembly scheme.	273
5.3	Classes of myosins.	274
5.4	Scheme of the intramolecular communication mechanism in cytoskeletal motors.	276
5.5	General scheme of the organelle transport mediated by kinesins and dyneins in the cell.	276
5.6	Sliding-filament model for contraction in striated muscle.	277
5.7	Cross-section scheme of a typical flagellum.	278
5.8	Types of cargo-linkage for kinesins, dyneins, and myosins.	279
5.9	Sequence alignment of human DYNLL1 and DYNLL2.	281
5.10	Sequence alignment of human DYNLL1 and homologues.	282
5.11	DYNLL1 structure and dimer formation.	282
5.12	Sequence alignment of human DYNLRB1 and DYNLRB2.	283
5.13	Sequence alignment of human DYNLT1 and DYNLT3.	284
5.14	Sequence alignment of human DYNLT1 and homologues.	284
5.15	DYNLT1 structure and dimer formation.	285

5.16	Interaction of DYNLT1 with DIC.	286
5.17	Sequence alignment of several mammalian DIC isoforms.	297
5.18	Yeast–two hybrid assay using DYNLT1 and various dynein intermediate chain constructs.	297
5.19	Far–UV CD spectra of purified recombinant DYNLT1 saturated with a DIC1 peptide and its self–saturated DYNLT1–DIC2 counterpart.	298
5.20	ITC analysis of the binding of a DIC peptide to purified DYNLT1.	298
5.21	Representative binding curve for a FITC–labelled DIC peptide to DYNLT1, measured by fluorescence polarization.	299
5.22	Molecular features of the DYNLT1–DIC2 chimera.	299
5.23	¹ H– ¹⁵ N heteronuclear NOE data of the purified DYNLT1–DIC chimera.	300
5.24	1D ¹ H spectra of DYNLT1 and DYNLT1–DIC2 construct.	301
5.25	Bar plots of $\Delta\delta H_\alpha$ and $\Delta\delta C_\alpha$ as a function of sequence for the purified DYNLT1–DIC chimera in aqueous solution.	301
5.26	¹ H– ¹⁵ N–HSQC spectrum of DYNLT1–DIC2.	302
5.27	Solution structure of human DYNLT1 self–saturated with a DIC polypeptide occupying the canonical binding groove.	303
5.28	Solution structure of human DYNLT1 self–saturated with a DIC2 polypeptide occupying the canonical binding groove.	304
5.29	Analysis of the binding plasticity of DYNLT1 canonical groove.	305
5.30	Yeast two–hybrid assay using DYNLT1 in the bait plasmid and various DIC constructs.	306
5.31	List of proposed cellular DYNLT1–interacting partners.	307
5.32	Sequence of ActRIIB at the binding site within the canonical groove of DYNLT1.	308
5.33	Confocal microscopy immunofluorescence of HEK293 cells transfected with mCherry–tagged DYNLT1 together with GFP–tagged ActRIIB constructs.	308
5.34	ActRIIB pull–down assays.	309
5.35	Effect of DYNLT1 on ActRIIB signalling using a luciferase reporter construct	309
5.36	Sequence alignment between DIC and Lfc.	310
5.37	Yeast–two hybrid assay in which DYNLT1 was confronted to various DIC and Lfc constructs.	310
5.38	ITC analysis of the binding of peptides Lfc(131–161), Lfc(131–154) and Lfc(137–154) to purified DYNLT1.	311
5.39	Yeast–two hybrid assay in which DYNLT1 was confronted to various Lfc constructs.	311
5.40	Binding curve for a FITC–labelled DIC1 peptide (118–178) to DYNLT1 measured by fluorescence polarization.	312
5.41	Sequence comparison between six DYNLT1–binding partners.	313
5.42	Structural comparison of the human, <i>C. reinhardtii</i> , and <i>D. melanogaster</i> DYNLT1 in complex with DIC peptides.	314
5.43	Proposed interaction model of DYNLT1 with various characterised targets.	316

Appendices

A1	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptide LytA _{239–252} in aqueous solution, DPC and SDS micelles, and in the presence of TFE.	377
A2	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptide LytA _{259–272} in aqueous solution, DPC and SDS micelles, and in the presence of TFE.	378
A3	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptide LytA _{239–272} in aqueous solution, DPC and SDS micelles, and in the presence of TFE.	379
A4	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptide SESYV11 in aqueous solution and in DPC micelles.	380
A5	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptide SESYW11 in aqueous solution and in DPC micelles.	381
A6	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptides K3W5–LytA _{239–252} and W5K10–LytA _{239–252} in aqueous solution.	382
A7	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptides K3W5–LytA _{239–252} and W5K10–LytA _{239–252} in DPC micelles.	383
A8	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptides S3S10–LytA _{239–252} , I3V10–LytA _{239–252} and I5Y6T11T13–LytA _{239–252} in aqueous solution.	384
A9	Bar plots of $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of sequence for peptides S3S10–LytA _{239–252} , I3V10–LytA _{239–252} and I5Y6T11T13–LytA _{239–252} in DPC micelles.	385

LIST OF TABLES

Table	Description	Page
Chapter 2		
2.1	List of molar extinction coefficients.	53
2.2	List of solvents used in NMR sample preparation.	55
2.3	Thermodynamic parameters calculated for LytA _{239–252} and LytA _{259–272} .	61
2.4	Summary of non-sequential NOEs involving side chain protons observed in aqueous solution for LytA _{239–252} and LytA _{259–272} .	63
2.5	Statistical parameters for the LytA _{239–252} and LytA _{259–272} calculated conformers in aqueous solution.	64
2.6	Statistical parameters for the LytA _{239–252} and LytA _{259–272} calculated conformers in 30 mM [D ₃₈]-DPC.	71
2.7	Statistical parameters for the LytA _{239–252} and LytA _{259–272} calculated conformers in 30 mM [D ₂₅]-SDS.	79
2.8	Thermodynamic parameters calculated for LytA _{239–272} .	82
2.9	Statistical parameters for the LytA _{239–272} calculated conformers in aqueous solution.	85
2.10	Statistical parameters for the LytA _{239–272} calculated conformers in 30 mM [D ₂₅]-SDS.	90
2.11	Structure populations for LytA _{239–272} and its variants in aqueous solution and in DPC micelles.	99
2.12	Summary of non-sequential NOEs involving side chain protons observed in aqueous solution for LytA _{239–252} and its variants.	101
Chapter 3		
3.1	Chronological summary of the published works related to BP100.	153
3.2	Sequences of BP100 analogues.	158
3.3	MICs of BP100-PP peptide and Ala-substituted analogues.	163
3.4	Haemolysis of BP100 and Ala-d ₃ substituted analogues.	164
3.5	Secondary structure fractions of BP100-PP peptide and Ala-d ₃ analogues in lipid vesicles.	166
3.6	² H quadrupolar splittings of BP100K5A in different oriented lipid bilayers.	168
3.7	² H quadrupolar couplings of Ala-d ₃ labelled BP100 analogues.	171
3.8	Best-fit orientations of BP100 in lipid bilayers from ² H NMR data.	172
3.9	Viperidins identified in the venom of South American pit vipers and their characteristics.	188
3.10	Primary structure and physicochemical properties of crotalicidin and fragments.	193
3.11	Estimated helix populations of Ctn and its Ctn _{1–14} and Ctn _{15–34} fragments in 30 mM DPC.	201
3.12	Non-sequential NOE signals observed in Ctn _{15–34} .	202
3.13	Structural quality data for the structures calculated for peptides Ctn, Ctn _{1–14} and Ctn _{15–34} .	203
3.14	MICs of Ctn and fragments Ctn _{1–14} and Ctn _{15–34} .	205

Chapter 4		
4.1	Examples of β -1,3-glucans and some of their properties.	228
4.2	Subfamilies of the endo- β -1,3-glucanases.	230
4.3	CBM folding families.	231
4.4	Prediction of N-glycosylation sites for Ole e 9 and Fra e 9.	245
4.5	Carbohydrates used in the binding assays.	248
4.6	NMR relaxation data for rCtD-Ole e 9 and rCtD-Fra e 9.	252
4.7	Characterization of the protein-laminarin complexes by sedimentation equilibrium.	254
4.8	Summary of the interactions involved in the laminarin-CBM complexes analysed.	256
4.9	Homologous sequences selected for comparing with rCtD-Ole e 9 and rCtD-Fra e 9.	258
Chapter 5		
5.1	NMR structural calculations statistics for DYNLT1-DIC2 construct.	303
Appendices		
A1	^1H and ^{13}C chemical shifts of LytA ₂₃₉₋₂₅₂ under different solvent conditions.	335
A2	^1H and ^{13}C chemical shifts of LytA ₂₅₉₋₂₇₂ under different solvent conditions.	338
A3	^1H and ^{13}C chemical shifts of LytA ₂₃₉₋₂₇₂ under different solvent conditions.	341
A4	^1H and ^{13}C chemical shifts of SESYW11 in DPC micelles.	347
A5	^1H and ^{13}C chemical shifts of SESYV11 in DPC micelles.	348
A6	^1H and ^{13}C chemical shifts of K3W5-LytA ₂₃₉₋₂₅₂ under different solvent conditions.	349
A7	^1H and ^{13}C chemical shifts of W5K10-LytA ₂₃₉₋₂₅₂ under different solvent conditions.	351
A8	^1H and ^{13}C chemical shifts of S3S10-LytA ₂₃₉₋₂₅₂ under different solvent conditions.	353
A9	^1H and ^{13}C chemical shifts of I3V10-LytA ₂₃₉₋₂₅₂ under different solvent conditions.	355
A10	^1H and ^{13}C chemical shifts of I5Y6T11T13-LytA ₂₃₉₋₂₅₂ under different solvent conditions.	357
A11	^1H and ^{13}C chemical shifts of crotalicidin in DPC micelles.	359
A12	^1H and ^{13}C chemical shifts of crotalicidin-Nt in DPC micelles.	363
A13	^1H and ^{13}C chemical shifts of crotalicidin-Ct in DPC micelles.	365
A14	Backbone ^1H and ^{13}C chemical shifts of rCtD-Fra e 9 in aqueous solution.	367
A15	Structural statistics parameters for the calculated conformers of LytA ₂₃₉₋₂₅₂ in aqueous solution and in micelles.	371
A16	Structural statistics parameters for the calculated conformers of LytA ₂₅₉₋₂₇₂ in aqueous solution and in micelles.	372
A17	Structural statistics parameters for the calculated conformers of LytA ₂₃₉₋₂₇₂ in aqueous solution and in micelles.	373

LIST OF ACRONYMS

ACP	Anticancer peptide
ADP	Adenosine diphosphate
AFM	Atomic force microscopy
AMP	Antimicrobial peptide
ATP	Adenosine triphosphate
BMG	Buffered minimal glycerol
BMM	Buffered minimal methanol
BSA	Bovine serum albumin
CA	Cecropin A
CBD	Choline-binding domain
CBM	Carbohydrate-binding module
CBP	Choline-binding protein
CBR	Choline-binding repeats
cDNA	Complementary DNA
CF₃-Bpg	3-(trifluoromethyl)-L-bicyclopent-[1.1.1]-1-ylglycine
CF₃-Phg	4-(trifluoromethyl)-phenylglycine
CLytA	C-terminal domain of LytA
cmc	Critical micelle concentration
ConA	Concanavalin A
COSY	Correlation spectroscopy
CP	Cross-polarization
CPP	Cell-penetrating peptide
CRAMP	Cathelicidin-related antimicrobial peptide
CSP	Chemical shift perturbation
Ctn	Crotalicidin
DEAE-cellulose	Diethylaminethanol-cellulose
DHC	Dynein heavy chains
DIC	Dynein intermediate chains
DIPSI	Decoupling in the presence of scalar interactions pulse sequence
DLC	Dynein light chains
DLIC	Dynein light-intermediate chains
DLS	Dynamic light scattering
DMEM	Dulbecco's modified Eagle's medium
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMPG	1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
DMSO	Dimethylsulfoxide
DPC	Dodecylphosphocholine
DPH	Diphenylhexatriene
DSC	Differential scanning calorimetry
DSS	Sodium 4,4-dimethyl-4-silapentane-1-sulfonate
DYNC1I2	Cytoplasmic dynein 1 intermediate chain 2
DYNLL1	Dynein light chain 1, cytoplasmic
DYNLT	Dynein light chain Tctex-type
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
ESI	Electrospray ionization
FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
Fmoc	Fluorenyl-9-methyloxycarbonyl

FP	Fluorescence polarization
GBP	Glycan-binding protein
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GH	Glycosylhydrolase
GHF	Glycosylhydrolase family
GPCR	G protein-coupled receptor
GTP	Guanosine triphosphate
H	Hydrophobicity
HEK	Human embryonic kidney
Hepes	4-(2-hydroxyethyl) -1-piperazineethanesulfonic acid
HFIP	Hexafluoroisopropanol
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation spectroscopy
IFN	Interferon
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IQFR	Instituto de química física Rocasolano
ITC	Isothermal titration calorimetry
LC-MS	Liquid chromatography-mass spectrometry
LDME	Laser Doppler micro-electrophoresis
Lfc	Lbl first cousin protein
Lfcin	Lactoferrin
LTA	Lipoteichoic acid
MAP	Membrane-active peptide
MAS	Magic angle spinning
MD	Molecular dynamics
MIC	Minimum inhibitory concentration
Mlt	Melittin
MLV	Multilamellar vesicle
MRE	Mean residue ellipticity
mRNA	Messenger RNA
MRW	Mean residue weight
MS	Mass spectrometry
Mw	Molecular weight
Ni-NTA	Nickel-nitrilotriacetic acid
NK	Natural killer cell
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
NRMSD	Normalized root mean square deviation
OB	Oligonucleotide/oligosaccharide binding
OCD	Oriented circular dichroism
ORF	Open reading frame
PB	Phosphate buffer
PBS	Phosphate buffered-saline
PC	Phosphocholine
PE	Phosphoethanolamine
PG	Phospho(1'-rac-glycerol)

PISEMA	Polarization inversion spin exchange at the magic angle
PMO	Phosphorodiamidate morphodiamidate morpholino-oligomer
PMSF	Phenylmethylsulfonyl fluoride
PNA	Peptide nucleic acid
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
PPI	Protein-protein interaction
PR	Pathogenesis-related protein
PRIDE	Proton inverse-detected deuteron
PS	Phosphatidylserine
QM	Quantum mechanics
Rab	Ras-like in brain protein
RC	Random coil
REDOR	Rotational-echo double resonance
Rho	Ras homologous
RIPA	Radioimmunoprecipitation assay
RMSD	Root mean square deviation
RP3	Retinitis pigmentosa 3
rCtD	Recombinant carboxyl terminal domain
RDC	Residual dipolar coupling
RP-HPLC	Reverse phase high performance liquid chromatography
rpm	Revolutions per minute
SATB1	Special AT-rich sequence-binding protein-1
SD	Standard deviation
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
SELCON	Self-consistent method
siRNA	Short interfering RNA
SPPS	Solid-phase peptide synthesis
ssNMR	Solid-state nuclear magnetic resonance
SSRS-MASE	Slow-spinning rotor-synchronized MAS exchange
SUV	Small unilamellar vesicle
SVD	Singular value decomposition
TA	Teichoic acid
Tctex1	T-complex testis-expressed dynein 1
TEV	Tobacco etch virus
TGFβ	Transforming growth factor β
TIM	Triosephosphate isomerase
TFA	Trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TOCSY	Total correlation spectroscopy
TOF	Time of flight
Tris	Tris(hydroxymethyl)aminomethane
TRNOE	Transferred nuclear Overhauser enhancement
X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranose

SUMMARY

The investigation of the molecular bases governing biological processes is essential to understand them, as well as their implications and impact in the organism of living creatures. In these processes, biomolecular recognition events are pivotal and the search for a comprehensive description at atomic level is necessary to know the details of mechanisms, regulation, selectivity, specificity, etc. In this context, NMR spectroscopy is a technique that provides a vast diversity of reliable information allowing the study of the characteristics of biomolecular recognition events at atomic resolution, which is the key to infer and clarify the fundamentals of relevant biological processes.

In this work, NMR spectroscopy combined with diverse biochemical, physicochemical, and computational techniques have been utilised to study several interesting biological systems involving some representative biomolecular recognition events, such as peptide–membrane interactions, carbohydrate–protein interactions, and protein–protein interactions. The aim of this thesis is to go deeper in the knowledge of some biomolecular recognition events, and reinforce that NMR spectroscopy as a powerful tool to address the characterization of biomolecular interactions.

This thesis has been structured in six chapters. Chapter 2 describes the study of a series of peptides derived from the choline-binding domain of pneumococcal autolysin LytA, aimed to find out whether they maintain their native structure and the ability to bind choline when isolated from the full-length protein. Using solution NMR, CD and fluorescence techniques, three peptides were found to show native-like, β -hairpin conformations in aqueous solution, as intended, and undergo an unexpected, reversible β -to- α transition in the presence of detergent micelles. An explanation to the interaction between these peptides and detergent micelle has been proposed, and the physicochemical bases of the observed structural transition were characterised by studying variants of one of these peptides.

Chapter 3 comprises the study of two different membrane-active peptides (MAPs) with antimicrobial properties: BP100 and crotalicidin. First, the characterisation of the interaction of BP100 with different lipid bilayers used as complex membrane mimetics was addressed by solid state NMR techniques, together with CD and microbiological assays. Accurate measurements of the peptide orientation were achieved and a mechanism of action was proposed. On the other hand, a reductionist approach was applied to study crotalicidin. Interesting information about the structure–function relationship was obtained using NMR and CD spectroscopies and microbiological evaluations. These results will facilitate the design of new drugs with improved therapeutic indexes.

In chapter 4, the origin of the different affinity of two homologous carbohydrate-binding modules (CBMs) for a β -1,3-glucan substrate was investigated. The interaction of CtD-Ole e 9 (from *Olea europaea*) and CtD-Fra e 9 (from *Fraxinus excelsior*), two CBMs from pollen allergens, with laminarin was examined by biochemical and physicochemical techniques, combined with a complete solution NMR analysis, including structural and

relaxation studies. A detailed description of the interaction was achieved, which led to propose a binding mechanism and an explanation for the observed differences in the affinity. These differences arise from small variations in the residues present in the binding site of the two CBMs.

Finally, chapter 5 addresses the description of the interaction between human DYNLT1, a dynein light chain, and DIC, a dynein intermediate chain. A chimera designed for this purpose was used to structurally characterise the canonical binding site by solution NMR. Results evidence many similarities with other members of the protein family. Potential binding partners for DYNLT1 were identified, including ActRIIB, which was able to bind DYNLT1 *in vitro*, indicating that DYNLT1 may be implicated in the signalling pathway of TGF- β . Analysis of the interaction with another binding partner, Lfc, showed that the binding process in DYNLT1 may involve contacts with residues located outside the canonical binding groove. All the obtained results lead to propose an interaction model for DYNLT1 with different partners.

In conclusion, the utilization of NMR techniques for the study of biomolecular recognition processes and interactions constitutes the central axis of this thesis. Many biochemical, physicochemical or computational techniques have been employed to complement NMR data in order to attain a complete description of the investigated systems and to unveil the questions considered, leading to new interesting findings that open the door to further investigations.

RESUMEN

La investigación de las bases moleculares que gobiernan los procesos biológicos es esencial para comprenderlos, así como para entender sus implicaciones y el impacto que tienen en el organismo de los seres vivos. En dichos procesos, los eventos de reconocimiento biomolecular son cruciales y es necesario lograr una descripción detallada a nivel atómico para conocer las particularidades de los mecanismos, su regulación, selectividad, especificidad... En este contexto, la espectroscopía de RMN es una técnica que proporciona gran cantidad de información fiable que permite estudiar las características de los eventos de reconocimiento biomolecular con resolución atómica, lo que es clave para deducir y clarificar los fundamentos de los relevantes procesos biológicos.

En este trabajo se ha empleado la espectroscopía de RMN, en combinación con diversas técnicas bioquímicas, fisicoquímicas y computacionales, para estudiar varios sistemas biológicos de interés en los que están presentes algunos eventos de reconocimiento biomolecular representativos, tales como las interacciones péptido–membrana, las interacciones carbohidrato–proteína, o las interacciones proteína–proteína. El objeto de esta tesis es profundizar en el conocimiento sobre varios casos de los citados eventos de reconocimiento biomolecular y reforzar el papel de la espectroscopía de RMN como una potente herramienta para abordar la caracterización de las interacciones biomoleculares.

Esta tesis está estructurada en seis capítulos. El capítulo 2 describe el estudio de una serie de péptidos derivados del dominio de unión a colina de la autolisina de pneumococo, LytA. Este estudio está dirigido a averiguar si dichos péptidos mantienen su estructura nativa y su capacidad de unir colina cuando se encuentran aislados de la proteína completa. Utilizando espectroscopía de RMN en disolución, dicroísmo circular (CD) y técnicas de fluorescencia, se han descubierto tres péptidos que adoptan conformaciones de horquilla β nativas en disolución acuosa, tal y como se pretendía, y que sufren una inesperada transición reversible de estructura β a estructura α en presencia de micelas de detergente. Se ha propuesto una explicación para la interacción que tiene lugar entre estos péptidos y las micelas de detergente, y se han caracterizado las bases fisicoquímicas de la transición estructural observada mediante el estudio de variantes de uno de estos péptidos.

El capítulo 3 comprende el estudio de dos péptidos activos de membrana (MAPs) diferentes con propiedades antimicrobianas: BP100 y crotalidina. En primer lugar, se ha abordado la caracterización de la interacción de BP100 con diferentes bicapas lipídicas usadas como miméticos de membrana complejos. Para ello, se han utilizado técnicas de espectroscopía de RMN en estado sólido, junto con CD y ensayos microbiológicos. Se ha logrado obtener medidas precisas de la orientación del péptido y se ha propuesto un mecanismo de acción. Por otra parte, para el estudio de la crotalidina se aplicó un enfoque reduccionista. Empleando espectroscopía de RMN y CD, así como evaluaciones microbiológicas, se ha obtenido una interesante información acerca de la relación estructura–función del péptido. Estos resultados facilitarán el diseño de nuevos fármacos con índices terapéuticos mejorados.

En el capítulo 4, se ha investigado el origen de la diferente afinidad de dos módulos homólogos de unión a carbohidratos (CBMs) por un sustrato de tipo β -1,3-glucano. Se ha examinado la interacción con laminarina de dos CBMs presentes en alérgenos del polen: CtD-Ole e 9 (procedente de *Olea europaea*) y CtD-Fra e 9 (procedente de *Fraxinus excelsior*). Para ello, se han usado técnicas bioquímicas y fisicoquímicas, combinadas con un completo análisis por espectroscopía de RMN en disolución que incluyó análisis estructural y medidas de relajación. Se ha logrado una descripción detallada de la interacción que ha permitido proponer un mecanismo de unión y una explicación para las diferencias observadas en la afinidad. Estas diferencias surgen de pequeñas variaciones en los residuos presentes en el sitio de unión de ambas CBMs.

Finalmente, en el capítulo 5 se ha procedido a la descripción de la interacción entre la DYNLT1 humana, una cadena ligera de la dineína, y DIC, una cadena intermedia de la dineína. Con este propósito, se diseñó una quimera que fue utilizada para realizar la caracterización estructural del sitio canónico de unión mediante espectroscopía de RMN en disolución. Los resultados evidencian muchas similitudes con otras proteínas de la misma familia. Se han identificado varias proteínas que potencialmente pueden unirse a DYNLT1, incluyendo ActRIIB, que se ha demostrado que es capaz de unirse a DYNLT1 *in vitro*, indicando que DYNLT1 podría estar implicada en la ruta de señalización de TGF- β . El análisis de la interacción con otra proteína, Lfc, ha mostrado que en el proceso de unión pueden intervenir residuos situados fuera del surco canónico de unión. El conjunto de resultados obtenidos ha permitido establecer un modelo de interacción de DYNLT1 con varias proteínas.

En conclusión, la utilización de técnicas de espectroscopía de RMN para el estudio de procesos de reconocimiento biomolecular y de las interacciones constituye el eje central de esta tesis. Varias técnicas bioquímicas, fisicoquímicas y computacionales se han empleado para complementar los datos de RMN, y así conseguir una completa descripción de los sistemas investigados y dar respuesta a las cuestiones planteadas, lo que ha llevado a nuevos hallazgos de gran interés que abren la puerta a futuras investigaciones.

CHAPTER 1

INTRODUCTION

1. BIOMOLECULAR INTERACTIONS AND RECOGNITION

Biochemical systems are constituted by a highly complex set of molecules and structures that are in close contact, sometimes transiently. Interactions governing these contacts are crucial to maintain and define the function of each and every element of the system. In this regard, molecular recognition is defined as the specific interaction between molecules through non-covalent forces (such as electrostatic interactions, hydrogen bonding, van der Waals forces, hydrophobic forces, π effects, metal coordination, halogen bonding...), in which solvent molecules can play a significant role. Highly specific regulated interactions between biomolecules are basic for every process taking place in living organisms, and thus, molecular recognition is a fundamental phenomenon in biological processes such as cellular signalling, enzyme catalysis, cell transport, ligand–receptor binding, protein complex formation, regulation functions, etc., just to name a few. (Baron & McCammon, 2013; Boehr, Nussinov, & Wright, 2009; Breiten et al., 2013; Cleaves, 2011; Cosic, 1994)

1.1 MECHANISTIC THEORIES FOR MOLECULAR RECOGNITION IN PROTEINS

In the last decade of the 19th century, German chemist H. E. Fischer proposed a mechanistic explanation for the enzyme–substrate binding in enzymatic reactions. The model, called “lock and key model”, (Fischer, 1894) stated that the substrate (key) accommodates specifically in the enzyme active site (lock), and therefore, only the right substrate is able to fit into the active site (Figure 1.1A). Substrates with different size, structure or spatial organization will not fit into the enzyme, as it exhibits an essentially fixed conformation. In the 1950s, D. E. Koshland suggested a new hypothesis to explain the enzyme–substrate recognition, known as “induced fit model”, which stated that interaction with ligand induces a conformational change in the enzyme to a more complementary form, in order to accommodate and bind the substrate more tightly (Figure 1.1B). (Koshland, 1958) Since these simple models, theoretical description of molecular recognition has been revised numerous times as the knowledge in this field became more extensive. (Boehr et al., 2009)

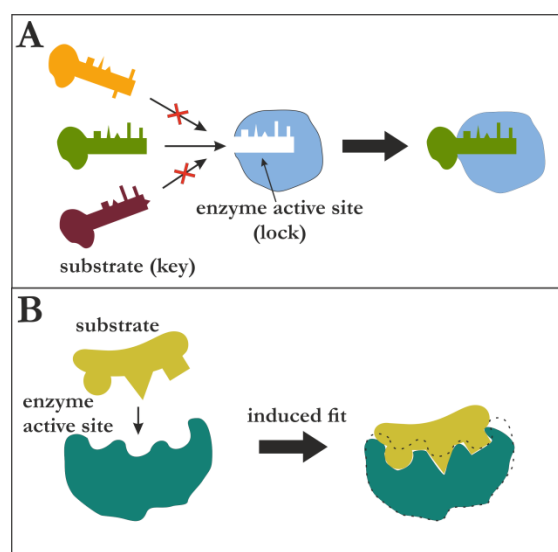


Figure 1.1. Classical molecular recognition models for proteins. **A.** Scheme of the “lock and key” recognition model proposed by H.E. Fischer. **B.** Scheme of the “induced fit” recognition model proposed by D. E. Koshland.

Nowadays, it is known that proteins are dynamic entities and they can adopt a huge amount of different conformations. Besides the native state (i.e. lowest energy), proteins usually are capable of accessing other conformational substates that may play an important role in molecular recognition. A more modern model named “conformational selection model” takes into account that conformational diversity and suggests that weakly populated, high energy non-native states are responsible for recognizing and binding molecules, and subsequently, population shifts toward these conformers. This model is derived from the energy landscape theory of protein dynamics (the well-known protein folding funnel). According to this theory, several substates exist in a dynamic equilibrium; they are populated following statistical thermodynamic distributions, and the conformational exchange timescale is defined by the height of the energy barriers. In this regard, in contrast to older models, this one indicates that binding interaction does not induce a conformational change; it just leads to a redistribution of the relative populations of the pre-existent substates (Figure 1.2). (Boehr et al., 2009; Csermely et al., 2010; Frauenfelder, Sligar, & Wolynes, 1991; Kumar, Ma, Tsai, Sinha, & Nussinov, 2000; Miller & Dill, 1997; Tsai, Kumar, Ma, & Nussinov, 1999)

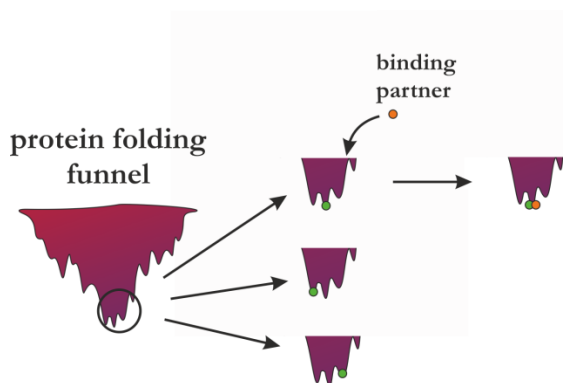


Figure 1.2. Scheme of the “conformational selection model”. The binding partner (orange) binds one of the several fluctuating conformations of the protein (green) and no further conformational rearrangement occurs. (Adapted from (Csermely, Palotai, & Nussinov, 2010)).

1.2 NON-COVALENT INTERACTIONS

As mentioned before, molecular recognition events take place by means of non-covalent interactions. Non-covalent interactions have some particular characteristics, being one of the most significant that they can exist at great distances, sometimes more than 10 Å. Electric and magnetic properties of the systems are the bases for this kind of interactions. Electrostatic interactions are dominant and arise from the proximity of charges, dipoles, quadrupoles, or higher multipoles. (Černý & Hobza, 2007) Non-covalent interactions between molecules can be of diverse origin and nature, and can be classified into four classes: electrostatic interactions, van der Waals forces, π -effects, and hydrophobic effects. (Lodish et al., 2008)

1.2.1 ELECTROSTATIC INTERACTIONS

Typically, electrostatic interactions involve the attraction of molecules with permanent net charges of opposite signs (ionic interactions). In proteins, these interactions

can occur when charged amino acids of different signs are present, such as Lys, His, Arg (positively-charged), and Asp, Glu (negatively-charged). Terminal charged peptide groups (positive amino and negative carboxyl ends) can also participate in ionic interactions.

Although they are not strictly electrostatic interactions, as they do not implicate net charges, hydrogen and halogen bonds are comparable to electrostatic interactions because of their strength. Actually, these two modalities of bonding are permanent dipole–permanent dipole interactions. Hydrogen bonds result from the interaction of a hydrogen atom with a partial positive charge (generated by the highly electronegative atom to which it is attached), and a very electronegative atom (that is partially negative), typically O or N. Hydrogen bonds are frequently found in biomolecules, including proteins. Many chemical groups (amines, carboxylates, hydroxyl...) from amino acid side chains or the backbone, are able to form hydrogen bonds between them, and also with solvent molecules. (Arunan et al., 2011; Lodish et al., 2008)

Halogen bonds are characterised by an electrophile halogen atom which interacts with an electron-rich species (nucleophile). The nucleophile is usually a strongly electronegative atom (such as O, N, or S), or even an anionic molecule, with a net negative charge. In spite of the fact that halogens are not natural constitutive blocks of biological macromolecules, the relevance of halogen bonds in the interaction of those biomolecules with small synthetic ligands, such as inhibitors, has been evidenced in some works. (Andrea, 2007; Auffinger, Hays, Westhof, & Ho, 2004; Desiraju et al., 2013; Legon, 2010)

In many biological systems a type of non-covalent interaction called “salt bridge” or “salt bond” can be found. This interaction involves both electrostatic interactions and hydrogen bonding. (Anslyn & Dougherty, 2006; Kumar & Nussinov, 1999)

1.2.2 VAN DER WAALS FORCES

Van der Waals forces occur between permanent or transient dipoles and they can be both attractive and repulsive. These forces are non-directional and short-ranged, as they decrease rapidly with increasing distance. Van der Waals interactions are weaker than hydrogen bonds. Three classes of van der Waals forces can be distinguished: (Lodish et al., 2008; McNaught & Wilkinson, 1997)

- Dipole–dipole interactions (Keesom interactions)

These interactions arise between permanent dipoles, and tend to align the molecules to reduce potential energy and increase attraction. They are usually found in molecules with electronegative atoms in which a separation of charges exists, giving rise to the above-mentioned permanent dipoles. (Anslyn & Dougherty, 2006)

- Dipole–induced dipole interactions (Debye force)

A permanent dipole is able to provoke distortions in the electronic cloud (polarize) of another (polarizable) molecule. As a result of the loss of symmetry in the electronic distribution, a dipole is generated and thus, an interaction between the permanent dipole and the induced dipole takes place. This kind of interaction is expected to occur between a polar molecule and a non–polar/symmetrical molecule. Dipole–induced dipole interactions are weaker than dipole–dipole interactions. (Anslyn & Dougherty, 2006)

- London dispersion force

Due to the presence of a polar molecule or to the repulsion between electron clouds in non–polar molecules, a non–zero instantaneous dipole momentum can arise in any non–polar molecule. Therefore, these dipole momenta are caused by random fluctuations of electron density in the molecular electronic clouds. As any molecule is polarizable, London force is present in every system. However, it is the weakest van der Waals interaction. (Anslyn & Dougherty, 2006)

1.2.3 π –EFFECTS

Systems containing π –bonds (π –systems) are capable of establishing a characteristic kind of interactions named π –effects. π –systems are electron–rich and, as a consequence, they can interact with many molecules, for instance, those with electron–deficient regions. There are different π –effects described so far: (Anslyn & Dougherty, 2006)

- Metal– π interactions: cationic or neutral metal atoms can interact with the face of a π –system.
- π –stacking: the π –systems of two aromatic rings interact with a specific geometry (edge–to–face, slipped).
- Cation– π interactions: an electronic–deficient system (cation) interacts with an electronic–rich system (π –system).
- Polar– π interactions: a polar molecule interacts with a π –system.
- Other: anion– π , π donor–acceptor, CH– π interactions...

1.2.4 HYDROPHOBIC EFFECT

Non–polar molecules show a high tendency to aggregate in polar environments in order to minimize the contact with the solvent. This effect is driven by a force called “hydrophobic bond”. In aqueous solution, the aggregation of hydrophobic molecules causes

the restriction in the motion of the polar solvent molecules, and thus, they become more organized, decreasing entropy (which is thermodynamically unfavourable). (Lodish et al., 2008)

2. INTERACTIONS WITH BIOLOGICAL MEMBRANES

Biological membranes are complex, two-dimensional fluids constituted by an amphipathic bilayer and a high density of proteins and carbohydrates. Despite their similar basic phospholipid bilayer structure and some common functions, each type of biological membrane shows different and distinctive features that are mainly defined by the particular set of associated proteins present in the membrane. (Lodish et al., 2008; A. W. Smith, 2012)

2.1 CELL MEMBRANES: A MODEL

In the early 1970s, S. J. Singer and G. L. Nicolson devised a model to explain the behaviour of cell membranes. This model, called “fluid mosaic membrane model” (S–N model), describes the membrane as a two-dimensional fluid lipid bilayer with a mosaic of proteins embedded in it. Membrane lipids and many associated proteins are able to move parallel to the surface and are, in fact, in constant motion (dynamic membrane). Some proteins occupy fixed positions, as they are anchored to cell structural elements, such as the cytoskeleton (Figure 1.3). (Singer & Nicolson, 1972)

The S–N model has been dominant for the last decades, and nowadays it is essentially valid. However, the advance in the knowledge of cell membrane properties has evidenced the necessity of addressing some issues. One of these issues is the fact that, despite the lateral and rotational freedom predicted by the model, the truth is that protein and lipid mobility in cell membrane is restricted. The presence of membrane domains (rafts) and the protein diffusion hindrance caused by the formation of supramolecular protein complexes are two observations that are in contradiction with the S–N model (Figure 1.3). (Vereb et al., 2003)

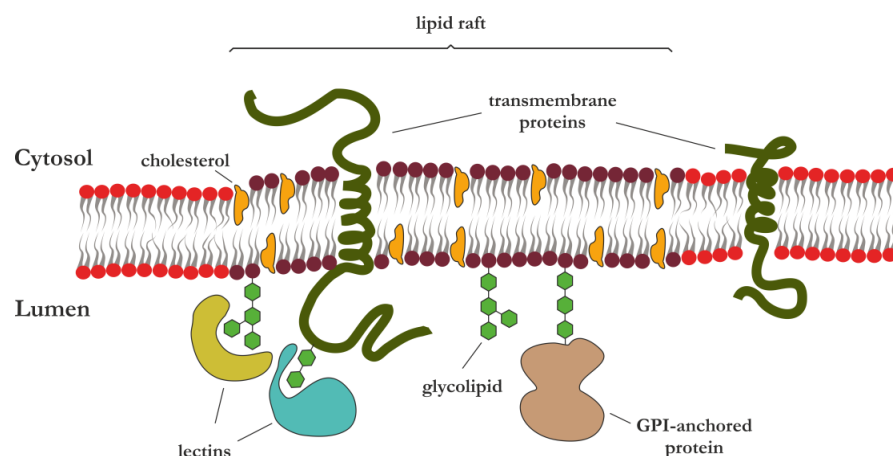


Figure 1.3. Fluid mosaic membrane model (S–N model). A fragment of the lipid bilayer is shown, as well as many of the typical elements present in the membrane. A lipid raft is also indicated.

Even though some experimental observations have revealed some limitations of the S–N model, it is a reasonably valid model, as stated before. However, it is necessary to be conscious of the fact that a simple model is not enough to describe such a complex system, as cell membrane is. Hence, S–N model can be accepted having in mind some considerations: free diffusion can occur within domain borders, where molecular interactions (lipid domain interactions, cytosolic interactions, or association with integral proteins) do not interfere. Baring this in mind, it has been suggested that membrane structure is highly compartmentalized and is more mosaic-like than fluid, and an updated model called “dynamically structured mosaic model” has been proposed. (Vereb et al., 2003)

2.2 COMPONENTS OF CELL MEMBRANES

Cell membranes are composed of lipids, proteins, and carbohydrates. Lipids are the main constituent of membranes, and can be classified into three types: phospholipids, glycolipids, and sterols. The concentration of each type of lipid varies depending on the class of cell, being typically phospholipids the most abundant. (Lodish et al., 2008)

- Phospholipids: they are amphipathic molecules that are capable of forming lipid bilayers. They usually possess two long fatty acid tails (hydrophobic region) bound to a glycerol moiety, and a polar head consisting of a phosphate group (also bound to the glycerol) to which many different classes of simple organic molecules can be covalently linked.
- Glycolipids: they are lipids with a carbohydrate moiety (monosaccharide or oligosaccharide) attached by a glycosidic bond. The lipid is often composed of either a glycerol or sphingosine backbone (glyceroglycolipids or sphingolipids, respectively).
- Sterols: they are a class of steroids, and thus, they show the classical sterane backbone, bearing a hydroxyl group in position 3.

After lipids, proteins are the next major component of cell membranes. If lipids play a fundamentally structural role in membranes, proteins are responsible for numerous and diverse biological functions. There are three classes of proteins associated to cell membrane: (Lodish et al., 2008)

- Integral or transmembrane proteins: they span the membrane and have hydrophilic domains able to interact with molecules present inside and outside the cell. The hydrophobic domain is responsible for anchoring the protein to the bilayer core. Examples of integral proteins are ion channels, proton pumps, etc.

- Membrane–anchored proteins: they are covalently attached to lipid molecules inserted in the cell membrane, so that the protein is not directly in contact with the membrane. For example, G proteins are membrane–anchored proteins.
- Peripheral proteins: they are associated with peripheral regions of the lipid bilayer or attached to integral membrane proteins. Interaction with membranes is only temporary. They are usually enzymes.

Finally, carbohydrates are also found in cell membranes, predominantly as glycoproteins, but also as glycolipids. Some cells possess a glycoprotein–polysaccharide covering called “glycocalix” enveloping the external surface of the cell. This structure and membrane–associated carbohydrates in general, play a crucial role in adhesion and recognition processes. (Lodish et al., 2008)

2.3 IMPORTANCE OF PROTEIN/PEPTIDE–MEMBRANE INTERACTIONS

Interaction of proteins with cell membranes is a pivotal subject in biology, as it is an essential step in many processes. Probably, two of the most relevant processes that involve protein–membrane interactions are cell signalling and membrane trafficking, in which many cytoplasmic proteins (peripheral proteins) are recruited to different cellular membranes. The strategies followed by those proteins to interact with membranes are diverse. Some of them have modular domains specialized in lipid binding (membrane–targeting domains). Other proteins utilize some regions of their molecular surface or a specific secondary structure element to interact with the membrane. In other cases, the proteins have covalently attached lipid anchors able to embed in the lipid bilayer. Membrane–protein interactions start with the formation of non–specific collisional complexes, driven by diffusion and electrostatic forces. Then, tightly bound complexes are formed, stabilized by specific interactions. (Cho & Stahelin, 2005)

As proteins, peptides are polymers of amino acids linked by amide bonds, but their size is normally smaller. Peptides are expected to use strategies to interact with micelles simpler than those used by proteins, but probably rather similar. In fact, peptides typically commence to interact non–specifically with membranes through diffusion and electrostatic forces, and subsequently, they establish more specific interactions thanks to their definite secondary structure. Membrane–interacting peptides are also very significant, as they are part of important processes related to the host defence, and it has been demonstrated that many peptides perform antimicrobial, anticancer, and other therapeutic effects, through membrane–related mechanisms.

2.4 MODEL MEMBRANES

Currently, the study of membrane-related proteins and membrane-active peptides are two of the greatest scientific themes of interest. Unfortunately, these biomolecules are often problematic: membrane proteins are difficult to purify and crystalize for X-ray crystallography and their strongly environment-dependent structure and functions make necessary to find an appropriate membrane milieu. In the case of membrane-active peptides, they are frequently poorly soluble in aqueous solutions. In addition, the structural complexity and physicochemical properties of biological membranes constitute a challenge for many experimental techniques, such as CD or NMR spectroscopy. Consequently, to study these molecules it is mandatory the utilization of suitable membrane models to reproduce as close as possible their biological environment. (Bechinger, 1999; Marcotte & Auger, 2005; Warschawski et al., 2011)

Many membrane models with different complexities have been investigated and employed in the last decades. These models vary in function of the experimental technique used, since each technique possesses its particular limitations. The final objective is to mimic the biological membrane in order to assure the presence of the same kind of interactions which lead to the adoption of the native folded state of the protein/peptide when it is in close contact with the membrane. (Warschawski et al., 2011) Some examples of the more frequently employed membrane models, from the simplest to the most complex, are:

- Simple organic molecule co-solvents (TFE/HFIP)

Mixtures of some halogenated alcohols, such as TFE or HFIP (Figure 1.4A), with water are used as the simplest model membranes as they are known to enhance secondary structure propensities of peptides and polypeptides. The mode of action of these molecules implies a decrease of the hydrogen bonding to the solvent, which enhances intramolecular hydrogen bonds stabilising the secondary structure. (Andersen et al., 1999; Santiveri, Pantoja-Uceda, Rico, & Jimenez, 2005; Sonnichsen, Van Eyk, Hodges, & Sykes, 1992; Warschawski et al., 2011)

- Detergent micelles

The use of micelle-forming detergents in water solution is a good choice to simulate membrane-like environments in CD or solution NMR experiments, because of their small size. Micelles are aggregates formed by self-assembly of amphipathic molecules in solution to minimize the exposition of their hydrophobic tails to the polar solvent (Figure 1.4B). The most popular detergents used to form micelles are DPC (zwitterionic), which is more similar to phosphatidylcholines abundant in eukaryotes; and SDS (anionic), which resembles more closely to bacterial membranes. Micelles show a small spherical (radius ~ 3 nm) monolayer with a rough surface, and they can adopt elliptical or rod-like shapes at high detergent concentrations or with weakly polar surfactants. Micelles are formed above a specific detergent concentration, called critical micelle concentration (cmc). Another characteristic parameter of micelles is the number

of aggregation (N), which is the number of molecules composing the micelle, and it is usually of several tens. (Bechinger, 1999; Erik Strandberg & Ulrich, 2004; Warschawski et al., 2011)

- Bicelles

Bicelles are an interesting membrane model because of their planar surface and lipid composition resembling biological membranes (Figure 1.4C). The morphologic and orientation properties, as well as easy preparation of isotropically tumbling bicelles make them appropriate for their use in solution NMR. In change, for solid-state NMR, macroscopically aligned bicelles are more suitable. Bicelles are typically made up of a mixture of lipids and detergents. The lipids are arranged in a disc-like shape with detergent molecules lining the curved circumference. A limitation of bicelles is that lipid composition cannot be modified much without disturbing the bicelle stability. (Marcotte & Auger, 2005; Erik Strandberg & Ulrich, 2004)

- Small unilamellar vesicles (SUVs)

Lipid vesicles are structures consisting of a portion of fluid enclosed by a lipid bilayer (Figure 1.4D). The smallest lipid vesicles are known as SUVs, and they have a diameter of approximately 30 nm (bigger than detergent micelles). Small-sized vesicles are preferred to bigger ones in order to avoid dispersion phenomena in some experimental techniques. A vast variety of lipids can be used to form SUVs, but adding charged lipids helps to slow down fusion and aggregation processes which lead to the formation of bigger vesicles. SUVs possess a very high curvature, and they provide moderately fast-tumbling species for solution NMR, but they are rather unstable. SUVs are also used in CD samples, but their large size often results in higher dispersion levels that lead to a decrease in the quality of spectra. (Warschawski et al., 2011)

- Nanolipoproteins (NLPs)

NLPs are phospholipid bilayers encircled by stabilizing amphipathic helical membrane scaffold proteins (MSPs) leading to nanoscale disc-shaped objects also known as “nanodiscs” (Figure 1.4E). These nanodiscs have a diameter of ~10 nm and a thickness of ~4 nm, equivalent to that of biomembranes. These entities can be solubilized in water, they can accommodate only one protein, and they are stable and monodisperse. Nanodiscs constitute a better membrane model than micelles or bicelles, because of their stability and flatness. NLPs are starting to be used in solution NMR. (Timothy H Bayburt, Grinkova, & Sligar, 2002; Timothy H. Bayburt & Sligar, 2010; Nath, Atkins, & Sligar, 2007; Warschawski et al., 2011)

▪ Oriented lipid bilayers

The use of samples containing oriented lipid bilayers (Figure 1.4F) is possible in some experimental techniques, such as solid-state NMR. Usually, they are prepared by spreading a lipid mixture diluted in organic solvents onto glass plates, evaporating the solvents and rehydrating the bilayer with water. This way, lipid bilayers are obtained oriented respect to the glass surface. Oriented lipid bilayers constitute a good membrane model due to their stability, flatness and a higher resemblance to natural biological membranes. In change, their use is limited to a few experimental techniques, such as solid state NMR. (Erik Strandberg & Ulrich, 2004; Warschawski et al., 2011)

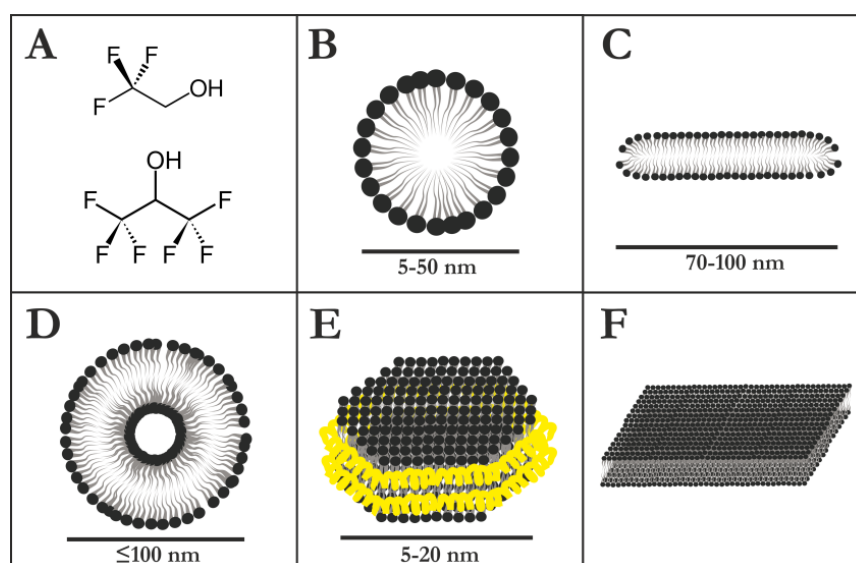


Figure 1.4. Main membrane models used for the study of protein/peptide-membrane interactions. **A.** Halogenated alcohols, such as TFE (top) or HFIP (bottom). **B.** Detergent micelles. **C.** Bicelles. **D.** Small unilamellar vesicles (SUVs). **E.** Nanolipoproteins (NLPs) Scaffold proteins are shown in yellow. **F.** Oriented lipid bilayers. Illustrative sizes are shown.

3. PROTEIN-CARBOHYDRATE INTERACTIONS

Proteins interacting non-covalently with carbohydrates are widespread in nature. This kind of proteins, known generally as carbohydrate-binding proteins or glycan-binding proteins (GBPs), recognizes and binds specifically glycans and mediate their biological function. Carbohydrates can be highly branched molecules and their monomers may be connected by many different linkages, generating a vast structural diversity. Taking advantage of this diversity, numerous proteins have evolved in order to adapt some of their modules to recognize specific glycans that mediate physiological or pathological processes. Carbohydrate characteristic features, such as mass, shape, charge, or other physicochemical properties, are responsible for their role in a wide variety of biological events. (Dwek, 1996; Lis & Sharon, 1998; Varki, Etzler, Cummings, & al., 2009)

3.1 CLASSES OF GBPs

There are many groups of proteins capable to bind glycan ligands, being the most significant:

- Lectins

Lectins are proteins able to bind mono- and oligosaccharides reversibly and with high specificity. In general, they are devoid of catalytic activity and contain several binding sites, i.e., they are multivalent. Lectins are found in most organisms, from viruses to humans, constituting a heterogeneous group of oligomeric proteins, varying widely in size, structure, molecular organization, etc. (Lis & Sharon, 1998)

- Carbohydrate-active enzymes

There are proteins with catalytic activity that are able to modify and/or degrade glycans, and these proteins need to bind the saccharide substrates in order to facilitate their access to the active site. As carbohydrates can show diverse degrees of complexity (monosaccharides, disaccharides, oligosaccharides, polysaccharides...), carbohydrate-active enzymes have developed different types of binding modules with distinct mechanisms. (Boraston, Bolam, Gilbert, & Davies, 2004; Davies & Williams, 2016)

- Glycan-specific antibodies

Glycans are usually not or just weakly immunogenic, as they cannot be processed by antigen-presenting cells to create a T-helper response. Naturally, only low-affinity IgMs arise as glycan-specific antibodies. However, high affinity IgG type antibodies against glycans have been developed biotechnologically. (Sterner, Flanagan, & Gildersleeve, 2016)

- Glycosaminoglycan-binding proteins

These proteins are capable of interacting with glycosaminoglycans (GAGs), such as heparin and heparan sulphate. GAGs are long unbranched polysaccharides with a disaccharide repeating unit consisting of an amino sugar along with a uronic sugar or galactose. These molecules exert their biological activities through the localization, stabilization, activation or inactivation of the interacting proteins. These interactions are involved in many physiological and pathological processes. (Hileman, Fromm, Weiler, & Linhardt, 1998; Varki et al., 2009)

3.2 THE NATURE OF THE PROTEIN–CARBOHYDRATE INTERACTION

Carbohydrates normally bind at regions containing shallow pockets located on the hydrophilic surface of the protein. The affinity of this binding is usually rather weak, and thus, multivalence is a pivotal feature in this process, as it permits an augmentation of the affinity and gives rise to other phenomena, such as agglutination. (Fernández-Alonso et al., 2012)

Many forces may be involved in carbohydrate–protein interactions, including hydrogen bonding, hydrophobic interactions, electrostatic interactions, and metal coordination: (Siebert et al., 2008)

- Hydrogen bonding: the large number of hydroxyl groups present in carbohydrates, and the occasional occurrence of other polar groups, such as amino or carboxyl, in some kind of saccharides, increases the probability of establishing hydrogen bonds with polar moieties of amino acids located in the binding site of the protein (for instance, side chains of Glu, Asp, Asn, Gln, Arg, or Ser; or backbone amino and carbonyl groups). Hydrogen bonding contributes both to affinity and selectivity of a protein towards a given sugar type. (Fernández-Alonso et al., 2012)
- Hydrophobic interactions: sugars are highly hydrophilic compounds, as they possess numerous hydroxyl groups. In spite of this, hydrophobic interactions of sugars with aromatic amino acids of GBPs play a significant role in the recognition process. It is well-known that pyranose rings can stack against aromatic rings of Trp, Tyr and Phe residues. The contacts take place between two or three CH groups from the pyranose ring and the π electron density of the aromatic rings, giving rise to the so-called “CH– π bonds”. (Asensio, Ardá, Cañada, & Jiménez-Barbero, 2012; Fernández-Alonso et al., 2012)
- Electrostatic interactions: sometimes, salt bridges can be formed between some types of saccharides and GBPs. For example, salt bridges may be established between the charged residues of sialic acid and protein amino acids of the opposite charge. (Fernández-Alonso et al., 2012)
- Metal coordination: sometimes, divalent cations act as bridges between sugar hydroxyl groups and negatively charged Asp or Glu. This is the case of some C-type lectins, which require Ca^{2+} to recognize their saccharide ligands. (Fernández-Alonso et al., 2012)

3.3 IMPORTANCE OF PROTEIN–CARBOHYDRATE INTERACTIONS

Many examples evidence the importance of this type of interactions. For instance, host–pathogen interactions are mediated by complex carbohydrates coating cell surfaces, which are recognized by proteins. Therefore, protein–carbohydrate interactions constitute

the first step in the contact between the host and the pathogen, and consequently they are involved in disease processes, such as inflammation, cancer or infections. A deeper understanding of the recognition of sugars by their corresponding receptors will help to the development of new drugs with improved properties respect to the natural saccharides (higher affinity, stability, and bioavailability). These drugs are compounds of low molecular mass based on the structure of functional carbohydrates, and they are called “glycomimetics”. (Fernández-Alonso et al., 2012; Magnani & Ernst, 2009)

Considering the sugars coating the cell surface of pathogenic bacteria, further insights into protein–carbohydrate interactions can potentially lead to the design of carbohydrate–based vaccines that may help our immune system to generate antibodies against those pathogens. (Avci & Kasper, 2009; Fernández-Alonso et al., 2012)

4. PROTEIN–PROTEIN INTERACTIONS (PPIS)

PPIs are of paramount importance in countless biological functions, and they are defined as non–covalent reversible contacts with molecular docking between proteins that occur *in vivo*, usually mediated by specific macromolecular recognition sites. Those contacts should be specific and oriented to the performance of a particular biological function by the formation of protein complexes. (De Las Rivas & Fontanillo, 2010; Uetz & Vollert, 2005; Yan, Wu, Jernigan, Dobbs, & Honavar, 2008)

4.1 CLASSES OF PPIS: PROTEIN COMPLEXES

PPIs can be classified attending to three main features: complex composition (homo– or hetero–oligomeric complexes), structural subsistence (obligate or non–obligate complexes), and protein interaction lifetime (transient or permanent complexes): (Alameer, 2012; Ozbabacan, Engin, Gursoy, & Keskin, 2011; Vinogradova & Qin, 2011)

- Homo–oligomeric and hetero–oligomeric complexes

Homo–oligomeric protein complexes are composed of two or more identical polypeptide units; they are symmetric and constitute good scaffolds for stable macromolecules. On the other hand, hetero–oligomeric protein complexes are made up of non–identical polypeptide chains, and their stability can be variable.

- Obligat and non–obligat complexes

Sometimes the constituents of a protein complex are unstable on their own *in vivo*, and thus they are forced to form a complex to enhance their stability. In these cases, the complexes are called “obligat”, and one example is the Ku proteins involved in DNA repair, which are obligat homodimers. In non–obligat complexes, the individual components can exist independently, as they are stable enough. (Ozbabacan et al., 2011)

- Transient and permanent complexes

It is said that a protein complex is permanent when PPIs involved in its formation are strong and irreversible, and therefore, the complex is very stable and long-lasting. Transient complexes are characterised by the transience of the PPIs implicated in the process, and they are associating and dissociating temporarily *in vivo*. Non-obligate interactions are principally transient. (Ozbabacan et al., 2011)

4.2 PROTEIN–BINDING DOMAINS AND INTERACTION INTERFACE

The formation of protein complexes involves the establishment of PPIs affecting specific regions of the constituent proteins. These regions can be located in protein-binding domains, and a large number of them have been described so far, such as WD40, SH2, SH3, PDZ, PTB... (Bock & Gough, 2001)

PPIs are greatly influenced by the properties of the protein–protein interaction interface. The size of the interaction interface is defined by the buried surface area (BSA), a magnitude determining the change in accessible surface area (ASA) in residues for proteins in their bound and unbound complexed state. Generally, larger interface surfaces ($> 2,000 \text{ \AA}^2$) involve significant conformational changes during complex formation, whereas smaller surfaces are related to less conformational flexibility during complex formation. (Alameer, 2012; Chothia & Janin, 1975; Susan Jones & Janet M Thornton, 1996)

Geometry and complementarity of the interaction interface are also essential elements for complex formation. Interface regions are usually more planar than the rest of the protein surface. It has been observed that planarity is higher in non-obligate complexes than in obligate complexes. Geometric complementarity is frequently present in protein complex interfaces, where a high atom packing density is found, comparable to that from the protein core. (Alameer, 2012; Susan Jones & Janet M Thornton, 1996; Murakami & Jones, 2006)

Secondary structural preferences are also a relevant feature in protein complex interfaces. Obligate homodimer interfaces show a higher proportion of α -helices than β -strands, whereas the proportion is similar in non-obligate hetero-complex interfaces. Non-obligate hetero-complex interfaces also show a high amount of loops, turns, and coils, compared to obligate homodimer interfaces. (Alameer, 2012; Guharoy & Chakrabarti, 2007; S. Jones & J. M. Thornton, 1996)

In homodimer interfaces, a slightly higher proportion of hydrophobic residues have been observed, compared to non-obligate heterodimer interfaces. Consequently, polar and charged residues are more abundant in non-obligate heterodimer interfaces than in homodimers. (Alameer, 2012; Susan Jones & Janet M Thornton, 1996)

4.3 PPIs IN BIOLOGICAL FUNCTIONS

PPIs and protein complex formation play a decisive role in a myriad of biological processes. It is very usual that protein complex formation leads to the activation or inhibition of any of the complex constituents. In this regard, PPIs and protein complex formation act as one of the numerous ways of regulating biological pathways.

PPIs are ubiquitous and absolutely indispensable for almost each and every biological process in which proteins are involved. Several examples of those essential processes can be cited: cell signalling, DNA repair, cytoskeletal trafficking and organization, protein synthesis, gene expression and regulation, membrane transport... (Feng & Walsh, 2001; Pawson & Nash, 2000, 2003)

In addition, aberrant PPIs are also of big interest, as they constitute the underlying cause of many serious diseases, such as cancer, Alzheimer's, Parkinson's or Creutzfeldt–Jakob diseases. (Cohen & Kelly, 2003; Dobson, 2002; Soto, 2003)

5. CHARACTERIZATION OF BIOMOLECULAR INTERACTIONS

Interactions between biomolecules can be predicted, detected and/or characterised by means of numerous techniques. Two main groups can be differentiated: theoretical approaches and experimental approaches.

5.1 THEORETICAL APPROACHES

Theoretical approaches to predict and study biomolecular interactions are based on the utilization of algorithms and computer models, and they usually require some experimental input data. The principal theoretical methods used to study biomolecular interactions are:

- Molecular docking

This method predicts the preferred orientation and describes the energetics of a biomolecule bound to other one forming a stable complex. Input data required for this prediction are the three-dimensional structures of receptor and ligand, usually obtained by X-ray crystallography or NMR. Over these structures, an appropriate search algorithm explores a space of all possible orientations and conformations of the molecular complex. This is a huge space, so docking programs normally explore the whole conformational space of the ligand, leaving the receptor as practically a rigid entity. Different docking possibilities are scored to evaluate the likelihood of a favourable binding interaction. (Lengauer & Rarey, 1996) There are many different docking approaches, such as the driven docking in which ambiguous experimental interaction restraints can be used to drive the docking process. This is the approach

utilized by the well-known protein-protein docking program HADDOCK. (Dominguez, Boelens, & Bonvin, 2003; van Zundert et al., 2016)

- Molecular dynamics (MD) simulation

MD simulation is a method able to calculate the time dependent behaviour of a molecular system, and it gives comprehensive information on the fluctuations and conformational changes occurring in biomolecules. At present, MD simulations are used routinely to study the dynamics and thermodynamics of biological ensembles. To perform MD, it is necessary to define appropriate force fields adapted to the system of interest. Some of the limitations of this method arise from the possible difficulties defining a force field that fits for the studied system. (Karplus & McCammon, 2002)

- Interaction network visualization

Combining computer science with mathematics, many interaction network visualization systems have been developed. These systems collect the information from biological datasets of physical, genetic and functional interactions, and then, they organize and display all the data in a flexible and expandable graphical format. (Alfarano et al., 2005; Breitkreutz, Stark, & Tyers, 2003)

5.2 EXPERIMENTAL APPROACHES

The high complexity is usually the main obstacle to address the experimental study of many biological and biochemical systems. For this reason, in this kind of studies the first step is often directed to the simplification of the selected system. The reductionist method is used in these cases, dissecting biological systems into their constituent parts. Reductionist studies are very usual when working with large proteins, and smaller fragments of them are analysed instead. The properties and characteristics observed in the fragments are very helpful to understand the features of the full-length protein. (Regenmortel, 2004)

Once the working system is defined, there are a lot of experimental methods capable of providing information about the potential interaction between biomolecules. Some of them are performed *in vivo*, and thus may procure valuable data about the actual interactions taking place in real biological systems.

The kind of information obtained varies depending on the experimental method considered. There are methods that are suitable for screening a big amount of potential binding partners, whereas other ones must be conducted on reduced sets of suspected binding molecules. Some experimental techniques just detect whether an interaction occurs or not, for instance by giving a visible response when a biomolecule binds another one.

Many techniques are oriented to investigate specific features of the binding process and the complex formation, such as kinetics, thermodynamics, stability, stoichiometry, structural effects, or atomic-level description of the process. Among those experimental techniques, biochemical and physicochemical methods can be found.

Some of the most usual biochemical methods employed to detect interactions in proteins are the yeast two-hybrid (Y2H) screen, phage display, pull-down immunoprecipitation, tandem affinity purification (TAP), affinity electrophoresis, microarrays, and proximity ligation assay (PLA). These techniques can provide information about the presence of protein interactions; reveal interaction–function relationships; determine crucial sequences for interaction; evaluate affinity constants; localize the subcellular positions where interactions occur, etc. (Bratkovič, 2010; Caufield, Sakhawalkar, & Uetz, 2012; Kuhn, Frei, & Christen, 1994; Raymond Sutandy, Qian, Chen, & Zhu, 2013; Rohila et al., 2006; Söderberg et al., 2008; Steinbrenner, Eldridge, Tomé, & Beynon, 2014)

Within the physicochemical methods utilized to study biomolecular interactions, two main groups can be distinguished: spectroscopic and non-spectroscopic methods. Spectroscopic methods are based on spectroscopic techniques, this is, techniques based on the interaction of electromagnetic radiation with the system, involving energy exchange between them. Non-spectroscopic methods can be used to detect conformational changes caused by interactions, calculate thermodynamical interaction parameters, and define complex stoichiometry and molecular mass. Some examples of non-spectroscopic methods are differential scanning calorimetry (DSC), isothermal titration calorimetry (ITC), analytical ultracentrifugation, and microscale thermophoresis (MST). (Chiu & Prenner, 2011; Jerabek-Willemsen, Wienken, Braun, Baaske, & Duhr, 2011; Pierce, Raman, & Nall, 1999; Schuck, 2003)

5.3 SPECTROSCOPIC METHODS TO STUDY PROTEIN INTERACTIONS

As stated before, spectroscopic methods to study biomolecular interactions are able to extract information about the system directly from the spectroscopic phenomenon, this is, the interaction between an electromagnetic radiation and the system involving an energy exchange between them.

Numerous spectroscopic techniques are regularly employed to obtain information about protein interactions, for instance:

- Dynamic light scattering (DLS): it can be used for the calculation of K_d , complex stoichiometry, hydrodynamic parameters, and thermodynamical characterization of interaction processes in proteins. (Hanlon, Larkin, & Reddick, 2010)
- Surface plasmon resonance (SPR): it is useful to measure K_d values, and to estimate kinetic parameters. (E. A. Smith, Thomas, Kiessling, & Corn, 2003)

- Fluorescence measurements: protein interactions can be characterised using fluorescent probes or quenchers for the intrinsic fluorescence of the protein. These measurements can be used to estimate binding affinities, and also to extract information about the interaction regions or molecular orientation. (Heyduk, Ma, Tang, & Ebright, 1996; Jähnig, 1979; Kakehi, Oda, & Kinoshita, 2001; Lundblad, Laurance, & Goodman, 1996)
- Förster resonance energy transfer (FRET): this technique permits the quantification of molecular dynamics of protein interactions, the detection of conformational changes, and even the monitoring of protein complex formation. (Truong & Ikura, 2001)
- Small-angle X-ray scattering (SAXS): it can be used to determine protein interaction potentials. (Kim, Dumont, & Gruebele, 2008)
- Electron spin resonance (EPR): using site-directed labelling, EPR experiments are able to reveal local dynamics, structure, and conformational changes produced by protein interactions. (Cafiso, 2012)

Besides the aforementioned ones, other spectroscopic techniques more suitable for obtaining structural information from the systems exist:

- IR spectroscopy: the use of novel techniques of IR spectroscopy has been reported as an interesting alternative to other usual techniques which eventually show time resolution-related limitations in the study of protein interactions and structure. Structural changes can be monitored by means of these novel IR methods, such as 2D IR spectroscopy. (Haris, 2010)
- Raman spectroscopy: Raman spectra of proteins serve as sensitive and selective fingerprints of their 3D structure, interactions, and dynamics. Protein assembly pathways, structural parameters, and recognition mechanisms have been studied and characterised using Raman spectroscopy. (Taraschi & Mendelsohn, 1980; Thomas Jr, 1999)
- Mass spectrometry (MS): MS is commonly used as a final stage in the purification of sub-proteomes, and it serves as a tool for protein identification. Data analysis leads to the deduction of protein-protein interactions. (Aebersold & Mann, 2003)
- Circular dichroism (CD): this technique provides information about the structural features and association processes of biomolecules. CD is very useful to perform secondary structure analysis, to characterise folding processes (thermodynamics, kinetics, intermediate states...) and conformational changes. It is also applicable to the study of molecular interactions. (Greenfield, 2000)

There are three techniques which stand out due to their atomic-level resolution:

- Atomic force microscopy (AFM): AFM is a high-resolution technique which has demonstrated to be useful for the direct quantification of the range and magnitude of interaction forces between proteins and other molecules. (Lee, Wang, Huang, & Lin, 2007)
- X-ray crystallography: thanks to its atomic-level resolution, X-ray crystallography offers a very accurate depiction of the spatial organization of the molecules interacting with proteins. However, this technique is limited by the possibility of getting a good-quality crystal form of the complex, which is not always easy to achieve. (Kobe et al., 2008; Palmer & Niwa, 2003)
- Nuclear magnetic resonance spectroscopy (NMR): this technique also provides an atomic-level resolution, with the advantage of being able to use samples in solution. NMR is a powerful tool to study protein interactions, since it is capable of detecting conformational changes, variations in the protein dynamic regime, interacting sites, etc. In addition, solid state NMR offers more possibilities when working with poorly soluble systems, or when investigating interactions with complex structures, such as biomembranes. (Baldus, 2006; Bonvin, Boelens, & Kaptein, 2005)

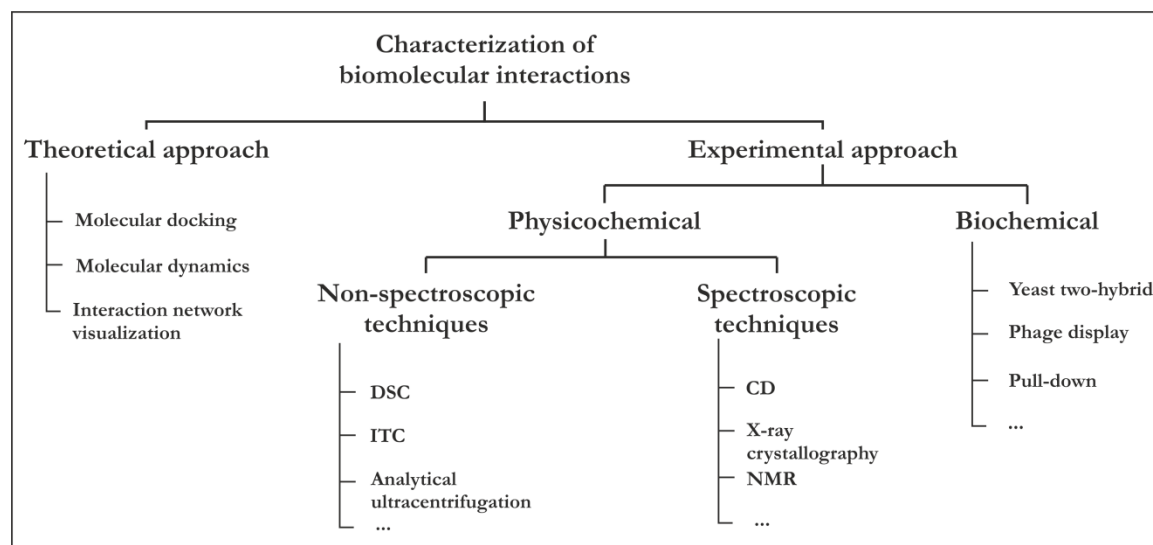


Figure 1.5. Scheme of the different approaches to characterise biomolecular interactions.

5.4 INVESTIGATING PROTEIN INTERACTIONS BY NMR

NMR is known to be one of the most sensitive tools for investigating biomolecular interactions at physiological conditions, especially those considered weak or ultra-weak ($K_d > 10^{-6}$ M). NMR can tackle samples that are solid or liquid, viscous or fluid, oriented or isotropic, static or spinning, cold or warm, etc.; and it can provide very accurate information

about local distances, orientation, and dynamics. (Vaynberg & Qin, 2006; Warschawski et al., 2011)

However, it is necessary to consider some general limitations of the NMR techniques that affect the study of PPIs. First, since NMR spectroscopy usually detects atomic nuclei with very low natural occurrences, to reach an acceptable sensitivity is mandatory to employ relative high amounts of sample (mg) and resort to isotopic labelling. Protein size is other limiting parameter, as the increase in molecular mass decreases solubility, homogeneity, stability, and favours undesired aggregation processes. Finally, the investigation of proteins by NMR is usually quite time-consuming because a big number of long experiments is required, and spectral analysis is complex.

5.4.1 INFORMATION OBTAINED FROM NMR EXPERIMENTS

There are several NMR parameters which can render structural information and evidence molecular interactions: chemical shifts, scalar coupling constants, residual dipolar coupling (RDCs), relaxation times, and nuclear Overhauser effect (NOE) measurements.

Chemical shifts, which are very sensitive to the chemical environment, can be used to estimate structural tendencies in the protein sequence. (Wishart, Sykes, & Richards, 1991) When the protein interacts with other molecule, changes in chemical shifts may be observed (chemical shift perturbations, CSP), and the position and magnitude of those changes allow for mapping the interaction region. (Vinogradova & Qin, 2011)

Scalar coupling constants vary when structural changes take place. (Ramachandran, Chandrasekaran, & Kopple, 1971) On the other hand, measurement of RDCs can provide information about the relative orientation of different regions of the protein. (Prestegard, Bougault, & Kishore, 2004)

Furthermore, NMR relaxation experiments can yield information about the dynamics of the system, showing where rigid and flexible regions are located, and which changes occur upon complex formation. Complex formation can also be confirmed based on alterations in the relaxation parameters (i.e., T_1 , T_2 , heteronuclear NOE...).

Finally, measurement of NOE effect is crucial to get fine structural data. NOE effect consists in a polarization transfer by means of an intermolecular cross relaxation process. Intensity of NOE effect is inversely proportional to the distance between the two nuclei originating the effect, and distance restraints can be deduced from it. From these distance restraints, structural calculation may be attempted. (Vinogradova & Qin, 2011)

5.4.2 SOLUTION NMR TO STUDY PROTEIN INTERACTIONS

Solution NMR has played a crucial role in the study of proteins, since it is able to provide structural information at atomic level, as well as information about the conformational dynamics. Determination of the protein structure is crucial when studying PPIs by solution NMR because it allows identifying and mapping interacting regions and the

residues involved in binding processes. Although most protein structures reported are elucidated by X-ray crystallography, solution NMR is a good alternative in many cases. For instance, when no good-quality crystals can be obtained for a protein sample, or when the protein is highly flexible, or when differences between the solution structure and the crystal are expected. At present, around 9 % of the structures deposited in the Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>) are determined by solution NMR. Structural determination by solution NMR usually implies the use of labelled samples and to obtain a set of distance and angle restraints from the assignment of NOESY type spectra. This set of restraints is used to find conformations that are consistent with them by theoretical calculations. These calculations are based on molecular dynamics/simulated annealing methods (MD/SA), which explore the conformational space of the molecule until an ensemble of conformers that fulfil the restrictions is obtained. (Hammes, 2005)

The study of PPIs by solution NMR can be focused on the observation of changes in the signals from the interacting molecules, but also on the study of how many other NMR parameters are affected upon interaction. Solution NMR not only allows confirming interactions, it also can be used to define and characterise binding regions, residues involved, dynamics, and even to calculate the structure of supramolecular complexes. (Barbero, 2003)

5.4.3 PROTEIN INTERACTIONS BY SOLUTION NMR: LIMITATIONS

Apart from the intrinsic limitations of NMR (mainly low sensitivity), when studying biomolecular interactions involving proteins or peptides some specific restrictions arise. First, proteins are usually large molecules involving the presence of a huge number of resonances in the spectra. This makes the assignment difficult, as many signals are often overlapped. To get the maximum of information, several different NMR experiments involving carbon and nitrogen nuclei must be performed, and thus, samples must be labelled with ^{13}C and ^{15}N to enhance resolution and sensitivity. Furthermore, in the case of the study of complexes the asymmetric labelling approach is often employed. This approach consists in labelling only one of the entities of the complex, facilitating the differentiation of intramolecular and intermolecular NOEs. Sometimes, interaction and complex formation lead to the precipitation of a significant amount of the sample and, consequently, spectra become poor as concentration falls below the sensitivity limit of the technique.

A challenging situation is faced when studying protein or peptide interaction with biomembranes. As explained before in this chapter, biomembranes are rather complex systems and the experimental utilization of models is mandatory in many techniques. Solution NMR can only deal with the most simple membrane models, such as micelles, bicelles, or nanodiscs at the most. To study tightly membrane-associated peptides or proteins, solution NMR is no longer directly applicable and solid-state NMR represents a good choice, allowing the study of the interaction with even more complex systems, like lipid bilayers. (Erik Strandberg & Ulrich, 2004; Warschawski et al., 2011)

5.4.4 SOLID-STATE NMR TO STUDY PEPTIDE-MEMBRANE INTERACTIONS

In the context of peptide-membrane interactions, the correlation times of small peptides associated with large macromolecular assemblies, such as lipid bilayers, are strongly determined by the size of the complex. Consequently, tumbling is considerably slower when compared to the free peptide in solution. In this situation, application of solid-state NMR (ssNMR) methods is very suitable to obtain information about structure and dynamics of the associated peptide. (Bechinger, 1999)

The principal objective of most of the studies performed on peptide-membrane interactions by ssNMR, is to characterise the orientation and dynamics of the peptide in the membrane-bound state. To reach this goal, a wide range of ssNMR experimental approaches has been employed. Sometimes, only the backbone is analysed, since it defines the overall fold and alignment of the peptide. Labelling strategies (^{15}N , ^2H , ^{13}C , or even other nuclei, such as ^{19}F) have been traditionally applied. NMR experiments used include cross-polarization (CP), CP-MAS, PISEMA, quad echo, Hahn echo, REDOR, TRNOE, PRIDE, SSRS-MASE... (Erik Strandberg & Ulrich, 2004)

Any simple α -helix or β -strand conformation can be identified from the isotropic chemical shifts of the ^{13}C backbone resonances by MAS. In the case of regular helices, they present a repetitive pattern of anisotropic spin interaction tensors along the primary sequence. The chemical shift of a ^{15}N or ^{13}C nucleus, the dipolar coupling within a bond, and the quadrupole splitting of ^2H , all exhibit characteristic orientation dependence. These anisotropic interactions can be measured for successive peptide planes in oriented membranes. The parameter depends on the given orientation of the tensor with the molecular frame of the peptide, on the time-averaged orientation of the peptide with respect to the bilayer normal, and on the alignment of the bilayer in the magnetic field. The periodic pattern of the anisotropic parameters with a periodicity of 3.6 residues is characteristic of helices and it is called “dipolar wave”, “chemical shift wave”, or “quadrupolar wave”, respectively. It is possible to estimate the orientation of the peptide from these parameters. (E. Strandberg et al., 2004; Erik Strandberg & Ulrich, 2004)

6. AIMS AND STRUCTURE OF THIS THESIS

As a first step to understand the complex biological processes, which have an enormous impact in our lives, it is crucial to explore and unravel their molecular bases. Molecular recognition events are essential in those processes, and an idea of their significance and biological implications has been depicted in the previous pages.

NMR spectroscopy is a technique capable of providing diverse and reliable data about biomolecular systems at an atomic level. This fact makes the technique very suitable to investigate the singularities behind the molecular recognition events, and to walk those first steps towards the interpretation and elucidation of a variety of biological processes of interest.

In the next chapters, the work carried out during my thesis will be explained in detail. The main objectives of this thesis have been to shed light on many different molecular recognition events with distinct biomolecular systems involved, and to confirm that NMR spectroscopy is a powerful tool to address the characterization of biomolecular interactions.

For those purposes, I have tried to make use of a wide repertoire of NMR experiments, complemented with many other different techniques. On the other hand, various interesting biological systems have been studied, involving some representative biomolecular recognition events, such as peptide–membrane interactions, carbohydrate–protein interactions, and protein–protein interactions.

This thesis has been structured in six chapters. In the present one, I tried to summarise the importance and variety of interactions between the principal types of biomolecules, as well as the main techniques utilised for the study of the interactions and binding processes between them. Chapters 2 to 5 are devoted to the experimental study of an assortment of different interactions in diverse selected biological systems. The relevance of all these system is highlighted and described in detail in the corresponding sections.

Chapter 2 describes the study of a set of switch peptides and the role of the interactions between them and model membranes in the switching process. Solution NMR has been used to characterise these peptides structurally, as well as the conformational changes promoted by different membrane mimetic environments. The study has been complemented with CD and fluorescence experiments.

Chapter 3 includes the study of two different membrane–active peptides (MAPs), and thus it is divided into two parts. In the first part, I have addressed a detailed characterization of the interaction of an antimicrobial peptide with different lipid bilayers used as complex membrane mimetics. For that purpose, solid –state NMR techniques has been employed, together with many other techniques including CD and microbiological assays. In the second part of chapter 3, a reductionist approach was followed to study the properties of an antimicrobial peptide with good antitumour activity. The peptide was dissected into two fragments and they were structurally analysed by solution NMR, to help the understanding of the results observed in the biological assays and the potential rational design of new drugs.

An interesting and relevant protein–carbohydrate interacting system has been investigated and described in chapter 4. We have structurally characterised by solution NMR, theoretical methods and hydrodynamic approaches, two homologous and related carbohydrate–binding modules (CBMs) that show different affinity for laminarin. On the bases of the obtained results, a model for the interaction has been proposed and the molecular bases of the observed different affinity are described.

Protein–protein interactions are the topic of chapter 5, focused in the description of specific interactions in proteins corresponding to cytoskeletal motors. Biochemical and biophysical methods were applied to analyse DYNLT1–DIC binding, as well as interactions between DYNLT1 and other binding partners.

Finally, the last chapter (chapter 6) includes a summary of the results obtained in this thesis with the main conclusions.

BIBLIOGRAPHY

- Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*, 422(6928), 198-207.
- Alameer, A. (2012). *Theoretical and experimental restraints to drive the docking of protein-protein complexes*. University of Leicester.
- Alfarano, C., Andrade, C. E., Anthony, K., Bahroos, N., Bajec, M., Bantoft, K., . . . Hogue, C. W. V. (2005). The Biomolecular Interaction Network Database and related tools 2005 update. *Nucleic Acids Res*, 33(suppl 1), D418-D424. doi:10.1093/nar/gki051
- Andersen, N. H., Dyer, R. B., Fesinmeyer, R. M., Gai, F., Liu, Z., Neidigh, J. W., & Tong, H. (1999). Effect of hexafluoroisopropanol on the thermodynamics of peptide secondary structure formation. *Journal of the American Chemical Society*, 121(42), 9879-9880.
- Andrea, V. R. (2007). The role of halogen bonding in inhibitor recognition and binding by protein kinases. *Current topics in medicinal chemistry*, 7(14), 1336-1348.
- Anslyn, E. V., & Dougherty, D. A. (2006). *Modern physical organic chemistry*: University Science Books.
- Arunan, E., Desiraju, G. R., Klein, R. A., Sadlej, J., Scheiner, S., Alkorta, I., . . . Hobza, P. (2011). Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure and Applied Chemistry*, 83(8), 1637-1641.
- Asensio, J. L., Ardá, A., Canada, F. J., & Jiménez-Barbero, J. s. (2012). Carbohydrate–aromatic interactions. *Accounts of chemical research*, 46(4), 946-954.
- Auffinger, P., Hays, F. A., Westhof, E., & Ho, P. S. (2004). Halogen bonds in biological molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(48), 16789-16794.
- Avci, F. Y., & Kasper, D. L. (2009). How bacterial carbohydrates influence the adaptive immune system. *Annual review of immunology*, 28, 107-130.
- Baldus, M. (2006). Molecular interactions investigated by multi-dimensional solid-state NMR. *Current Opinion in Structural Biology*, 16(5), 618-623.
- Barbero, J. J. (2003). *Aplicaciones de la RMN al estudio de interacciones entre moléculas*. Paper presented at the Anales de la Real Sociedad Española de Química.
- Baron, R., & McCammon, J. A. (2013). Molecular Recognition and Ligand Association. *Annual Review of Physical Chemistry*, 64(1), 151-175. doi:doi:10.1146/annurev-physchem-040412-110047
- Bayburt, T. H., Grinkova, Y. V., & Sligar, S. G. (2002). Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Letters*, 2(8), 853-856.
- Bayburt, T. H., & Sligar, S. G. (2010). Membrane protein assembly into Nanodiscs. *FEBS Letters*, 584(9), 1721-1727. doi:<http://dx.doi.org/10.1016/j.febslet.2009.10.024>
- Bechinger, B. (1999). The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1462(1–2), 157-183. doi:[http://dx.doi.org/10.1016/S0005-2736\(99\)00205-9](http://dx.doi.org/10.1016/S0005-2736(99)00205-9)
- Bock, J. R., & Gough, D. A. (2001). Predicting protein–protein interactions from primary structure. *Bioinformatics*, 17(5), 455-460.
- Boehr, D. D., Nussinov, R., & Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol*, 5(11), 789-796. doi:10.1038/nchembio.232
- Bonvin, A. M., Boelens, R., & Kaptein, R. (2005). NMR analysis of protein interactions. *Current Opinion in Chemical Biology*, 9(5), 501-508.

- Boraston, A., Bolam, D., Gilbert, H., & Davies, G. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.*, 382, 769-781.
- Bratkovič, T. (2010). Progress in phage display: evolution of the technique and its applications. *Cellular and Molecular Life Sciences*, 67(5), 749-767.
- Breiten, B., Lockett, M. R., Sherman, W., Fujita, S., Al-Sayah, M., Lange, H., . . . Whitesides, G. M. (2013). Water Networks Contribute to Enthalpy/Entropy Compensation in Protein–Ligand Binding. *Journal of the American Chemical Society*, 135(41), 15579-15584. doi:10.1021/ja4075776
- Breitkreutz, B.-J., Stark, C., & Tyers, M. (2003). Osprey: a network visualization system. *Genome Biol.*, 4(3), R22.
- Cafiso, David S. (2012). Taking the Pulse of Protein Interactions by EPR Spectroscopy. *Biophysical Journal*, 103(10), 2047-2048. doi:10.1016/j.bpj.2012.10.005
- Caufield, J., Sakhawalkar, N., & Uetz, P. (2012). A comparison and optimization of yeast two-hybrid systems. *Methods*, 58(4), 317-324.
- Černý, J., & Hobza, P. (2007). Non-covalent interactions in biomacromolecules. *Physical Chemistry Chemical Physics*, 9(39), 5291-5303.
- Cleaves, H. J. (2011). Molecular Recognition. In M. Gargaud, R. Amils, J. C. Quintanilla, H. J. Cleaves, W. M. Irvine, D. L. Pinti, & M. Viso (Eds.), *Encyclopedia of Astrobiology* (pp. 1079-1080). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Cohen, F. E., & Kelly, J. W. (2003). Therapeutic approaches to protein-misfolding diseases. *Nature*, 426(6968), 905-909.
- Cosic, I. (1994). Macromolecular bioactivity: is it resonant interaction between macromolecules?—Theory and applications. *Biomedical Engineering, IEEE Transactions on*, 41(12), 1101-1114.
- Csermely, P., Palotai, R., & Nussinov, R. (2010). Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends in Biochemical Sciences*, 35(10), 539-546. doi:10.1016/j.tibs.2010.04.009
- Chiu, M. H., & Prenner, E. J. (2011). Differential scanning calorimetry: An invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions. *Journal of Pharmacy and Bioallied Sciences*, 3(1), 39-59. doi:10.4103/0975-7406.76463
- Cho, W., & Stahelin, R. V. (2005). Membrane-Protein Interactions in Cell Signaling and Membrane Trafficking. *Annual Review of Biophysics and Biomolecular Structure*, 34(1), 119-151. doi:doi:10.1146/annurev.biophys.33.110502.133337
- Chothia, C., & Janin, J. (1975). Principles of protein-protein recognition. *Nature*, 256(5520), 705-708.
- Davies, G. J., & Williams, S. J. (2016). Carbohydrate-active enzymes: sequences, shapes, contortions and cells. *Biochemical Society Transactions*, 44(1), 79-87.
- De Las Rivas, J., & Fontanillo, C. (2010). Protein–Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks. *PLoS Computational Biology*, 6(6), e1000807. doi:10.1371/journal.pcbi.1000807
- Desiraju, G. R., Ho, P. S., Kloo, L., Legon, A. C., Marquardt, R., Metrangolo, P., . . . Rissanen, K. (2013). Definition of the halogen bond (IUPAC Recommendations 2013). *Pure and Applied Chemistry*, 85(8), 1711-1713.
- Dobson, C. M. (2002). Protein-misfolding diseases: Getting out of shape. *Nature*, 418(6899), 729-730.
- Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A Protein–Protein Docking Approach Based on Biochemical or Biophysical Information. *Journal of the American Chemical Society*, 125(7), 1731-1737. doi:10.1021/ja026939x
- Dwek, R. A. (1996). Glycobiology: toward understanding the function of sugars. *Chemical reviews*, 96(2), 683-720.

- Feng, Y., & Walsh, C. A. (2001). Protein–protein interactions, cytoskeletal regulation and neuronal migration. *Nature Reviews Neuroscience*, 2(6), 408-416.
- Fernández-Alonso, M. d. C., Díaz, D., Berbis, M. Á., Marcelo, F., Cañada, J., & Jiménez-Barbero, J. (2012). Protein–Carbohydrate Interactions Studied by NMR: From Molecular Recognition to Drug Design. *Current protein & peptide science*, 13(8), 816-830. doi:10.2174/138920312804871175
- Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft*, 27(3), 2985-2993. doi:10.1002/cber.18940270364
- Frauenfelder, H., Sligar, S. G., & Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science*, 254(5038), 1598-1603.
- Greenfield, N. J. (2000). Biomacromolecular applications of circular dichroism and ORD (pp. 117-130): Academic Press, London.
- Guharoy, M., & Chakrabarti, P. (2007). Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions. *Bioinformatics*, 23(15), 1909-1918.
- Hammes, G. G. (2005). *Spectroscopy for the biological sciences*. John Wiley & Sons.
- Hanlon, A. D., Larkin, M. I., & Reddick, R. M. (2010). Free-Solution, Label-Free Protein–Protein Interactions Characterized by Dynamic Light Scattering. *Biophysical journal*, 98(2), 297-304. doi:<http://dx.doi.org/10.1016/j.bpj.2009.09.061>
- Haris, P. I. (2010). Can infrared spectroscopy provide information on protein–protein interactions? *Biochemical Society Transactions*, 38(4), 940-946.
- Heyduk, T., Ma, Y., Tang, H., & Ebright, R. H. (1996). Fluorescence anisotropy: rapid, quantitative assay for protein–DNA and protein–protein interaction. *Methods in enzymology*, 274, 492.
- Hileman, R. E., Fromm, J. R., Weiler, J. M., & Linhardt, R. J. (1998). Glycosaminoglycan–protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *BioEssays*, 20(2), 156-167. doi:10.1002/(SICI)1521-1878(199802)20:2<156::AID-BIES8>3.0.CO;2-R
- Jähnig, F. (1979). Structural order of lipids and proteins in membranes: evaluation of fluorescence anisotropy data. *Proceedings of the National Academy of Sciences*, 76(12), 6361-6365.
- Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P., & Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. *Assay and drug development technologies*, 9(4), 342-353.
- Jones, S., & Thornton, J. M. (1996). Principles of protein–protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 93(1), 13-20.
- Jones, S., & Thornton, J. M. (1996). Principles of protein–protein interactions. *Proceedings of the National Academy of Sciences*, 93(1), 13-20.
- Takehi, K., Oda, Y., & Kinoshita, M. (2001). Fluorescence polarization: analysis of carbohydrate–protein interaction. *Analytical biochemistry*, 297(2), 111-116.
- Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology*, 9(9), 646-652.
- Kim, S. J., Dumont, C., & Gruebele, M. (2008). Simulation-Based Fitting of Protein–Protein Interaction Potentials to SAXS Experiments. *Biophysical journal*, 94(12), 4924-4931. doi:<http://dx.doi.org/10.1529/biophysj.107.123240>
- Kobe, B., Guncar, G., Buchholz, R., Huber, T., Maco, B., Cowieson, N., . . . Forwood, J. K. (2008). Crystallography and protein–protein interactions: biological interfaces and crystal contacts. *Biochemical Society Transactions*, 36(6), 1438-1441.
- Koshland, D. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences*, 44(2), 98-104.

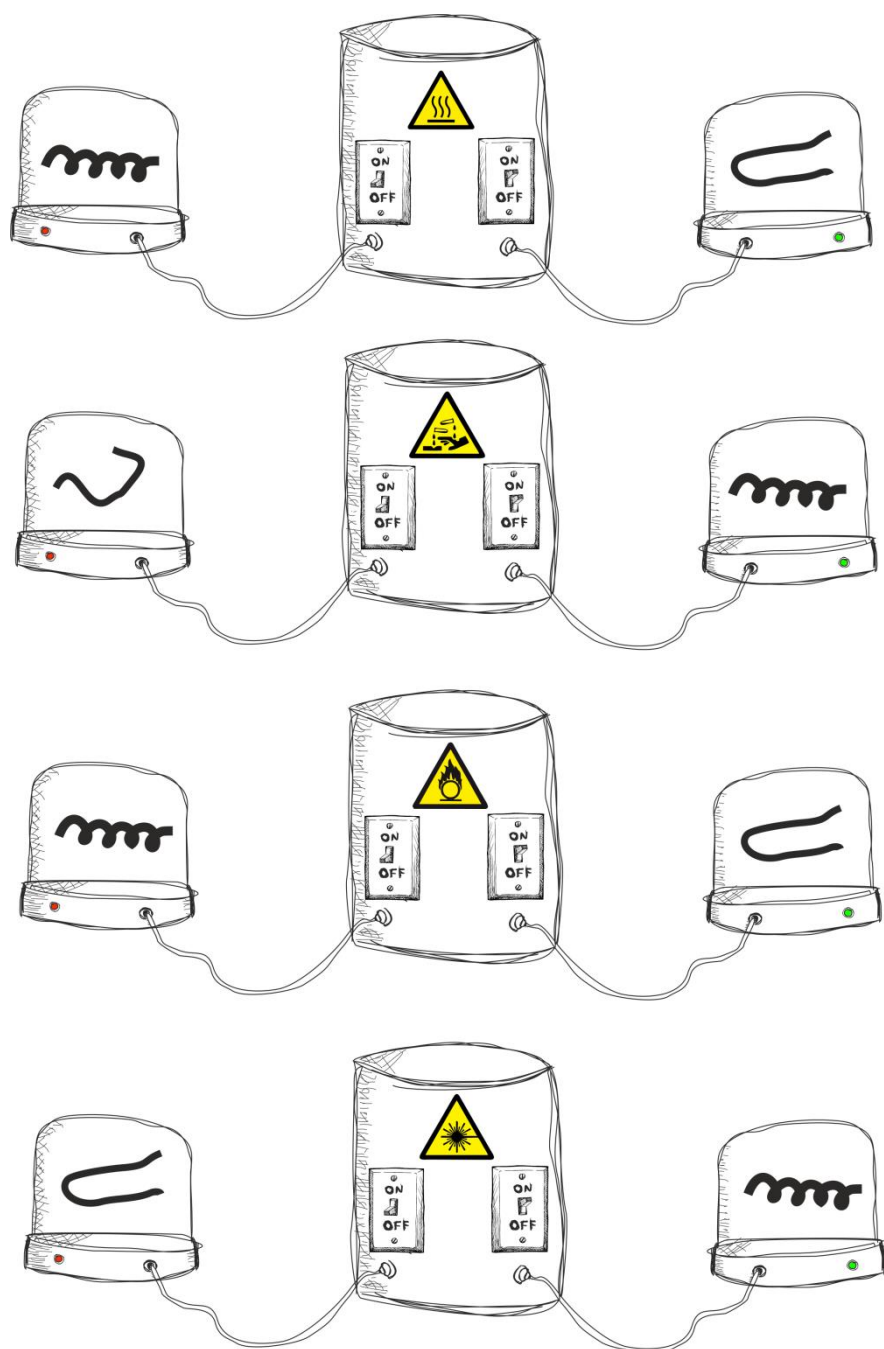
- Kuhn, R., Frei, R., & Christen, M. (1994). Use of capillary affinity electrophoresis for the determination of lectin-sugar interactions. *Analytical biochemistry*, 218(1), 131-135.
- Kumar, S., Ma, B., Tsai, C. J., Sinha, N., & Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Science*, 9(1), 10-19.
- Kumar, S., & Nussinov, R. (1999). Salt bridge stability in monomeric proteins. *J Mol Biol*, 293(5), 1241-1255.
- Lee, C.-K., Wang, Y.-M., Huang, L.-S., & Lin, S. (2007). Atomic force microscopy: Determination of unbinding force, off rate and energy barrier for protein–ligand interaction. *Micron*, 38(5), 446-461. doi:<http://dx.doi.org/10.1016/j.micron.2006.06.014>
- Legon, A. C. (2010). The halogen bond: an interim perspective. *Physical Chemistry Chemical Physics*, 12(28), 7736-7747.
- Lengauer, T., & Rarey, M. (1996). Computational methods for biomolecular docking. *Current Opinion in Structural Biology*, 6(3), 402-406. doi:[http://dx.doi.org/10.1016/S0959-440X\(96\)80061-3](http://dx.doi.org/10.1016/S0959-440X(96)80061-3)
- Lis, H., & Sharon, N. (1998). Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chemical reviews*, 98(2), 637-674.
- Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Scott, M. P., Bretscher, A., . . . Matsudaira, P. (2008). *Molecular Cell Biology* (W. H. Freeman Ed. 4th ed.).
- Lundblad, J. R., Laurance, M., & Goodman, R. H. (1996). Fluorescence polarization analysis of protein-DNA and protein-protein interactions. *Molecular Endocrinology*, 10(6), 607-612.
- Magnani, J. L., & Ernst, B. (2009). Glycomimetic drugs-a new source of therapeutic opportunities. *Discovery medicine*, 8(43), 247-252.
- Marcotte, I., & Auger, M. (2005). Bicelles as model membranes for solid-and solution-state NMR studies of membrane peptides and proteins. *Concepts in Magnetic Resonance Part A*, 24(1), 17-37.
- McNaught, A. D., & Wilkinson, A. (1997). *IUPAC. Compendium of Chemical Terminology* (2nd ed.): Blackwell Scientific Publications.
- Miller, D. W., & Dill, K. A. (1997). Ligand binding to proteins: the binding landscape model. *Protein Science*, 6(10), 2166-2179.
- Murakami, Y., & Jones, S. (2006). SHARP2: protein–protein interaction predictions using patch analysis. *Bioinformatics*, 22(14), 1794-1795.
- Nath, A., Atkins, W. M., & Sligar, S. G. (2007). Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry*, 46(8), 2059-2069.
- Ozbabacan, S. E. A., Engin, H. B., Gursoy, A., & Keskin, O. (2011). Transient protein–protein interactions. *Protein Engineering Design and Selection*, 24(9), 635-648.
- Palmer, R., & Niwa, H. (2003). X-ray crystallographic studies of protein–ligand interactions. *Biochemical Society Transactions*, 31(5), 973-979.
- Pawson, T., & Nash, P. (2000). Protein–protein interactions define specificity in signal transduction. *Genes & development*, 14(9), 1027-1047.
- Pawson, T., & Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. *Science*, 300(5618), 445-452.
- Pierce, M. M., Raman, C., & Nall, B. T. (1999). Isothermal titration calorimetry of protein–protein interactions. *Methods*, 19(2), 213-221.
- Prestegard, J., Bougault, C., & Kishore, A. (2004). Residual dipolar couplings in structure determination of biomolecules. *Chemical reviews*, 104(8), 3519-3540.
- Ramachandran, G., Chandrasekaran, R., & Kopple, K. D. (1971). Variation of the NH–C α H coupling constant with dihedral angle in the NMR spectra of peptides. *Biopolymers*, 10(11), 2113-2131.

- Regenmortel, M. H. V. V. (2004). Reductionism and complexity in molecular biology. *EMBO Reports*, 5(11), 1016-1020. doi:10.1038/sj.embor.7400284
- Reymond Sutandy, F. X., Qian, J., Chen, C.-S., & Zhu, H. (2013). Overview of Protein Microarrays. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, 0 27, Unit-27.21. doi:10.1002/0471140864.ps2701s72
- Rohila, J. S., Chen, M., Chen, S., Chen, J., Cerny, R., Dardick, C., . . . Kanrar, S. (2006). Protein-protein interactions of tandem affinity purification-tagged protein kinases in rice. *The Plant Journal*, 46(1), 1-13.
- Santiveri, C. M., Pantoja-Uceda, D., Rico, M., & Jimenez, M. (2005). β -Hairpin formation in aqueous solution and in the presence of trifluoroethanol: A ^1H and ^{13}C nuclear magnetic resonance conformational study of designed peptides. *Biopolymers*, 79(3), 150-162.
- Schuck, P. (2003). On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Analytical biochemistry*, 320(1), 104-124.
- Siebert, H.-C., Von Der Lieth, C.-W., Gilleron, M., Reuter, G., Wittmann, J., Vliegenthart, J. F. G., & Gabius, H.-J. (2008). Carbohydrate-Protein Interaction *Glycosciences* (pp. 291-310): Wiley-VCH Verlag GmbH.
- Singer, S. J., & Nicolson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. *Science*, 175(4023), 720-731.
- Smith, A. W. (2012). Lipid-protein interactions in biological membranes: A dynamic perspective. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1818(2), 172-177. doi:<http://dx.doi.org/10.1016/j.bbamem.2011.06.015>
- Smith, E. A., Thomas, W. D., Kiessling, L. L., & Corn, R. M. (2003). Surface plasmon resonance imaging studies of protein-carbohydrate interactions. *Journal of the American Chemical Society*, 125(20), 6140-6148.
- Söderberg, O., Leuchowius, K.-J., Gullberg, M., Jarvius, M., Weibrecht, I., Larsson, L.-G., & Landegren, U. (2008). Characterizing proteins and their interactions in cells and tissues using the in situ proximity ligation assay. *Methods*, 45(3), 227-232.
- Sonnichsen, F., Van Eyk, J., Hodges, R., & Sykes, B. (1992). Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry*, 31(37), 8790-8798.
- Soto, C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. *Nature Reviews Neuroscience*, 4(1), 49-60.
- Steinbrenner, J., Eldridge, M., Tomé, D. F. A., & Beynon, J. L. (2014). A Simple and Fast Protocol for the Protein Complex Immunoprecipitation (Co-IP) of Effector: Host Protein Complexes. In P. Birch, T. J. Jones, & I. B. J. Bos (Eds.), *Plant-Pathogen Interactions: Methods and Protocols* (pp. 195-211). Totowa, NJ: Humana Press.
- Sterner, E., Flanagan, N., & Gildersleeve, J. C. (2016). Perspectives on Anti-Glycan Antibodies Gleaned from Development of a Community Resource Database. *ACS chemical biology*. doi:10.1021/acscchembio.6b00244
- Strandberg, E., Ozdirekcan, S., Rijkers, D. T., van der Wel, P. C., Koeppe, R. E., 2nd, Liskamp, R. M., & Killian, J. A. (2004). Tilt angles of transmembrane model peptides in oriented and non-oriented lipid bilayers as determined by ^2H solid-state NMR. *Biophys J*, 86(6), 3709-3721. doi:10.1529/biophysj.103.035402
- Strandberg, E., & Ulrich, A. S. (2004). NMR methods for studying membrane-active antimicrobial peptides. *Concepts in Magnetic Resonance Part A*, 23(2), 89-120.
- Taraschi, T., & Mendelsohn, R. (1980). Lipid-protein interaction in the glyceophorin-dipalmitoylphosphatidylcholine system: Raman spectroscopic investigation. *Proceedings of the National Academy of Sciences*, 77(5), 2362-2366.
- Thomas Jr, G. J. (1999). Raman spectroscopy of protein and nucleic acid assemblies. *Annual Review of Biophysics and Biomolecular Structure*, 28(1), 1-27.

- Truong, K., & Ikura, M. (2001). The use of FRET imaging microscopy to detect protein–protein interactions and protein conformational changes in vivo. *Current Opinion in Structural Biology*, 11(5), 573-578.
- Tsai, C.-J., Kumar, S., Ma, B., & Nussinov, R. (1999). Folding funnels, binding funnels, and protein function. *Protein Science*, 8(06), 1181-1190.
- Uetz, P., & Vollert, C. S. (2005). Protein-Protein Interactions *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine* (pp. 1548-1552): Springer.
- van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastitis, P. L., Karaca, E., . . . Bonvin, A. M. J. J. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol*, 428(4), 720-725. doi:<http://dx.doi.org/10.1016/j.jmb.2015.09.014>
- Varki, A., Etzler, M. E., Cummings, R. D., & al., e. (2009). *Essentials of glycobiology* (C. S. H. L. Press Ed. 2 ed.).
- Vaynberg, J., & Qin, J. (2006). Weak protein–protein interactions as probed by NMR spectroscopy. *Trends in Biotechnology*, 24(1), 22-27.
- Vereb, G., Szöllősi, J., Matko, J., Nagy, P., Farkas, T., Vigh, L., . . . Damjanovich, S. (2003). Dynamic, yet structured: the cell membrane three decades after the Singer–Nicolson model. *Proceedings of the National Academy of Sciences*, 100(14), 8053-8058.
- Vinogradova, O., & Qin, J. (2011). NMR as a unique tool in assessment and complex determination of weak protein–protein interactions *NMR of Proteins and Small Biomolecules* (pp. 35-45): Springer.
- Warschawski, D. E., Arnold, A. A., Beaugrand, M., Gravel, A., Chartrand, É., & Marcotte, I. (2011). Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1808(8), 1957-1974.
- Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol*, 222(2), 311-333.
- Yan, C., Wu, F., Jernigan, R. L., Dobbs, D., & Honavar, V. (2008). Characterization of Protein–Protein Interfaces. *The Protein Journal*, 27(1), 59-70. doi:10.1007/s10930-007-9108-x

CHAPTER 2

SWITCH PEPTIDES



The work detailed in this chapter was carried out in collaboration with the group of Dr. J. M. Sanz, belonging to the Instituto de Biología Molecular y Celular from Universidad Miguel Hernández (Elche, Spain). Dr. B. Maestro, from the mentioned group, performed the fluorescence experiments and part of the CD experiments. I did the NMR study and the rest of CD experiments, which were performed during one of my short stays in the group of Prof. A. S. Ulrich, at the Institute for Biological Interfaces, from Karlsruhe Institute of Technology (Karlsruhe, Germany).

INTRODUCTION

1. CHOLINE-BINDING PROTEINS (CBPs)

Choline-binding proteins (CBPs) constitute a protein family characterised by their ability to interact non-covalently with choline molecules. Typically, they are associated to choline moieties present in the external cell surface of some bacteria. *Streptococcus pneumoniae* was the first microorganism in which this kind of proteins were described. At present, their presence has been reported in many other bacterial species, such as *Clostridium difficile*, *Streptococcus oralis*, *Pseudomonas aeruginosa*, and *Haemophilus influenzae*. (Gosink, Mann, Guglielmo, Tuomanen, & Masure, 2000; Hakenbeck, Madhour, Denapate, & Brückner, 2009; Beatriz Maestro, Santiveri, Jiménez, & Sanz, 2011)

Choline (2-hydroxy-*N,N,N*-trimethylethanamonium, $\text{HOCH}_2\text{CH}_2\text{N}^+(\text{CH}_3)_3$) is a small aminoalcohol with a net positive charge located on the nitrogen atom. This molecule is an essential nutrient for some Gram-positive microorganisms. These organisms take choline from the environment and incorporate it, in a phosphorylated form, into teichoic (TAs) and lipoteichoic acids (LTAs). TAs and LTAs are bacterial polysaccharides composed of glycerol phosphate or ribitol phosphate moieties bound by phosphodiester bonds. In the case of LTA, polysaccharide chains are attached to the cell membrane by a diacylglycerol residue. These polysaccharides are embedded in the thick peptidoglycan layer (a multi-layered murein) that envelopes the bacterial external cell surface, and whose basic function is to provide stiffness to the cell wall, attracting metal ions, such as sodium or magnesium (Figure 2.1). (P. García, Moscoso, Rodríguez-Cerrato, Yuste, & García, 2010; Beatriz Maestro et al., 2011; Pérez-Dorado, Galan-Bartual, & Hermoso, 2012)

The presence of choline associated to the TAs and LTAs is essential for the adequate functionality of many enzymes located in the polysaccharide layer surrounding the bacterial cell, such as hydrolases, involved in several cell processes. (P. García et al., 2010; Beatriz Maestro et al., 2011; Pérez-Dorado et al., 2012)

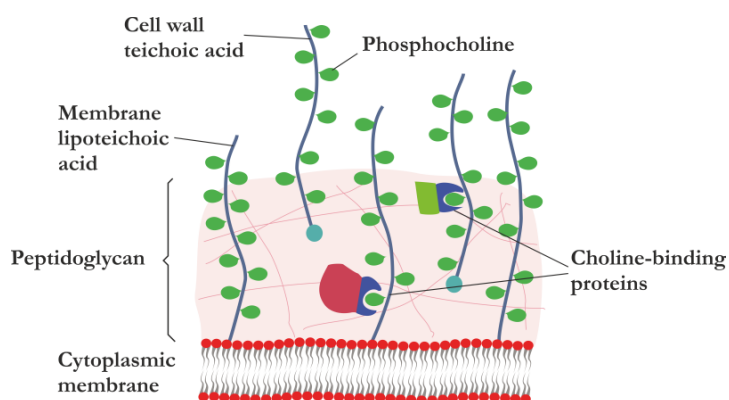


Figure 2.1. Representation of the peptidoglycan layer. CBPs are present in the peptidoglycan layer associated to choline moieties from TAs and LTAs.

CBPs are characterised by a modular structure in which two different domains can be clearly distinguished: a functional domain, responsible for the catalytic activity; and a choline-binding domain (CBD). Occasionally, the domains are separated by a flexible Pro-rich segment. (Gosink et al., 2000; Hermoso et al., 2003; Beatriz Maestro et al., 2011; Pérez-Dorado et al., 2012) The CBD is formed by choline-binding repeats (CBRs) that are sequences ordered in tandem and composed of approximately 20 highly conserved amino acids. Interaction between choline molecules and CBPs is defined by the residues present in the CBRs, being many of them aromatic. Aromatic side chains from some amino acids (2 to 4) interact with the charged region of choline, the ammonium centre, leading to a stabilization of the net positive charge by π -cation interactions. (Cheng, Goldstein, Gershenson, Stec, & Roberts, 2013; P. García et al., 2010)

Up to now, the proteins of CBP family are only known to be expressed by bacteria. Activity and molecular function of these proteins can be diverse. Some examples of CBP activities are:

- Adhesins: molecules located in the cell surface which facilitate bacterial cell adhesion to other cells or structures, a basic process to infect and colonize the host. (Pérez-Dorado et al., 2012)
- Hydrolases: enzymes that hydrolyse substrates, generally, molecules sited on the cell surface, such as envelope components. (Díaz, López, & García, 1991; E. García, García, Ronda, García, & López, 1985; Pérez-Dorado et al., 2012)
- Inhibitors: some CBPs act inhibiting other enzymes, as part of a regulation mechanism of certain processes. (Pérez-Dorado et al., 2012)

By means of the different activities that can exhibit, CBPs are involved in diverse biological functions, habitually related to the bacterial virulence. For this reason, many CBPs are considered virulence factors, that is, molecules expressed by pathogenic microorganisms that facilitate any event necessary for the persistence and expansion of the bacteria in the host organism. Some of those events include autolysis, cell adhesion, effects on immune host system, etc. (Fernández-Tornero, López, García, Giménez-Gallego, & Romero, 2001; B. Maestro & Sanz, 2005; Mellroth et al., 2012; Pérez-Dorado et al., 2012)

1.1 PNEUMOCOCCAL AUTOLYSIN LYTA

Streptococcus pneumoniae is one of the major opportunistic human pathogens and the most common source of acute respiratory infectious diseases. It is a Gram-positive microorganism that tends to colonize the upper airways, often settling in the nasopharynx. In general, it is a relatively benign organism when non-previous problems exist in the respiratory tract; nevertheless, in some conditions it can provoke pneumonia, meningitis or septicaemia, particularly in vulnerable individuals, such as children, elderly people, or

immunodeficient patients. (Fernández-Tornero et al., 2001; Hermoso et al., 2003; Jedrzejewski, 2006; Pérez-Dorado et al., 2012)

Cell structure of *S. pneumoniae* includes a thick external layer (150–300 Å), called murein. This layer is composed of peptidoglycan strata that envelope the cytoplasmic membrane (see Figure 3.2 of Chapter 3). The function of the murein layer is to provide stiffness and to maintain the cell turgor. In this region, numerous proteins covalently and non-covalently associated to the cell wall are found. Based on the structural motif used for their attachment to the cell surface, these proteins can be classified into three main groups:

- LPXTG-motif proteins
- LXCC-motif proteins
- Choline-binding proteins (CBPs)

Sixteen different CBPs have been reported in *S. pneumoniae*, depending on the considered strain. All of them are associated to the cell surface by means of non-covalent interactions with choline residues present in TAs and LTAs that are attached to the murein layer. Most pneumococcal CBPs somehow participate in cell adhesion, whereas other ones show hydrolase activity. In any case, all of them are related to the bacterial virulence. Presently, due to the increase in the number of antibiotic-resistant *S. pneumoniae* strains, and to the improvable efficiency of the available vaccines, CBPs from this microorganism are considered as good pharmacological targets, since they play a key role in bacterial virulence. (Fernández-Tornero et al., 2001; Gosink et al., 2000; Jedrzejewski, 2006; Pérez-Dorado et al., 2012)

One of the CBPs expressed by *S. pneumoniae* is the amidase LytA. This one was the first protein of this class to be described and it is an enzyme catalysing the hydrolysis of the bonds linking the *N*-acetylmuramoyl-L-alanine units present in the peptidoglycan backbone. This hydrolysis triggers the global degradation of the cell wall, and with that, the cell lysis. (B. Maestro & Sanz, 2005; Mellroth et al., 2012; Pérez-Dorado et al., 2012; Tettelin et al., 2001)

The role of LytA in *S. pneumoniae* cell functions is not totally clear yet. Some of the functions proposed in which LytA seems to be implied are:

- Mediation in the cell lysis process, which leads to the release of virulence factors, such as pneumolysin. This toxin is capable of forming transmembrane pores in the cells of the endothelium and epithelium, causing a tissue damage that permits the bacterial spread over the host organism. (Fernández-Tornero et al., 2001; B. Maestro & Sanz, 2005; Mellroth et al., 2012; Pérez-Dorado et al., 2012)
- Elicitation of a fratricidal action involving the lysis of non-competent pneumococcal cells. (Mellroth et al., 2012)

- Participation in evasion mechanisms or interference with the host immune system response. (Mellroth et al., 2012)
- Segregation of daughter cells at the end of cell division. (B. Maestro & Sanz, 2005)

All these proposed functions require the recognition of the choline molecules present in the cell wall by the choline-binding module (CBD) of the protein. (Beatriz Maestro & Sanz, 2007)

LytA, as all the CBPs, possesses a modular structure divided into two domains. On one hand, there is the biologically functional domain that shows amidase activity and is located at the N-terminal region of the protein. On the other hand, there is the choline-binding domain (CBD), occupying the C-terminal region (Figure 2.2). (B. Maestro & Sanz, 2005; Mellroth et al., 2012; Pérez-Dorado et al., 2012)

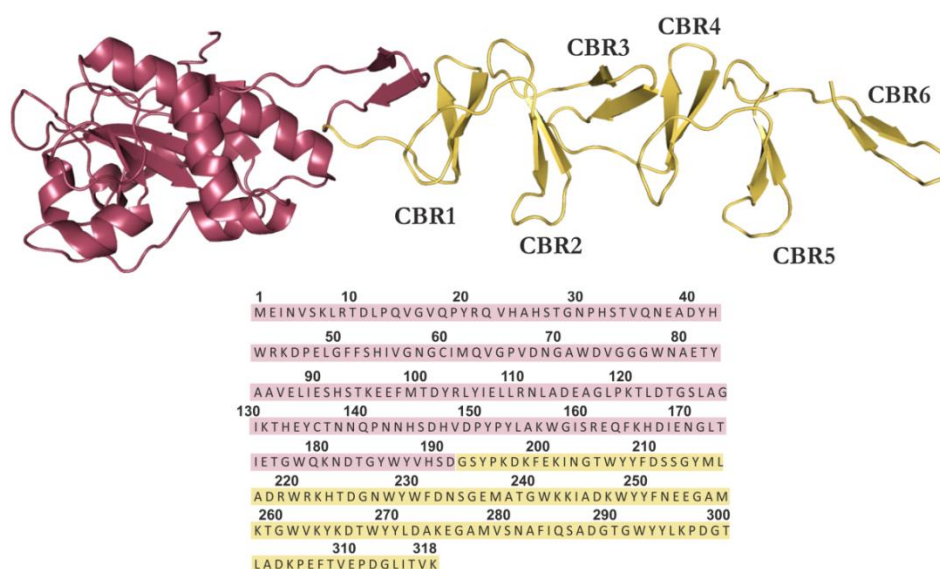


Figure 2.2. X-ray crystallographic structure of pneumococcal LytA (PDB code: 4X36). Both N-terminal catalytic domain (coloured in raspberry) and C-terminal choline-binding domain (CBP)(coloured in light yellow) are shown (top). Sequence is indicated below, with the two domains highlighted following the same colour code.

Generally, LytA performs its catalytic function as a V-shaped homodimer. The β -hairpin located at the C-terminal end is responsible for the dimerization, which is favoured upon the choline binding. In the monomeric state, the enzyme shows a decrease of the catalytic efficiency over 90 %. Choline binding not only induces dimerization, but also it provides thermal stability to the molecule. It is thought that this is due to a conformational change, which protects hydrophobic residues that, in absence of choline, are exposed to the solvent. (Fernandez-Tornero, García, Lopez, Gimenez-Gallego, & Romero, 2002; B. Maestro & Sanz, 2005)

LytA C-terminal region (CLytA) is composed of six imperfect repeats of a 20-amino acid sequence, known as P-motifs, cell wall-binding repeats (CBW), or choline-binding

repeats (CBRs). These repeats as a whole constitute the choline-binding domain (CBD). Structurally, this domain is made up of seven β -hairpins arranged in a left-handed superhelix (it belongs to the solenoid protein structural family) (Figure 2.2). In the first reported structures for this domain, which were initially used in this work, only six β -hairpins were identified. (Fernández-Tornero et al., 2001) For this reason, the numbering of the CBRs utilized in this chapter corresponds to the first reported structures, in which CBR1 is in fact the second β -hairpin of this domain. The first published structure containing the seven β -hairpins dates from 2014. (Mellroth et al., 2014) The β -hairpins of this domain are formed by two antiparallel β -strands connected by a small turn, they all have the same length at the β -strands and their composition is similar (five residues, mostly hydrophobic). Each β -hairpin is linked to the next one by a segment of 8–10 residues with non-regular structure. (Fernández-Tornero et al., 2001; Hernández-Rocamora et al., 2009; Pérez-Dorado et al., 2012)

Recognition of choline by CBPs usually occurs in CBR consensus motifs, typically $\text{GWXK-X}_{4-5}\text{-WYY-}\Phi\text{-X}_{3-5}\text{GXM}_{2,3}$, where X is any residue and Φ is a hydrophobic amino acid. 3D arrangement of these canonical motifs consists of a β -hairpin (12–14 residues) followed by a loop (8–10 residues). CBRs are generally organized in a $\beta\beta$ -3 solenoid superhelix (according to the Kobe-Kajava classification, (Kobe & Kajava, 2000)) and choline molecules are recognised at the interface between two consecutive repeats. LytA CDB is classically considered to have four choline-binding sites, although some authors suggest the presence of additional binding sites (Figure 2.3A). (Mellroth et al., 2014) Choline molecules are stabilised by cation- π interactions with three conserved residues: two tryptophans from one repeat and one tyrosine from the next repeat (Figure 2.3B). The side chains of these residues form a cavity that harbours the choline methyl groups located in the positively charged region of the molecule. A Met or Leu is often located at the bottom of the pocket. The net positive charge of the choline is stabilised by the π -electron system from the aromatic amino acids. (Cheng et al., 2013; Fernández-Tornero et al., 2001; Hernández-Rocamora et al., 2009; B. Maestro & Sanz, 2005; Pérez-Dorado et al., 2012) Despite being the general scheme, the canonical binding site can show variations in the number of aromatic residues present in the site. (Galán-Bartual, Pérez-Dorado, García, & Hermoso, 2015)

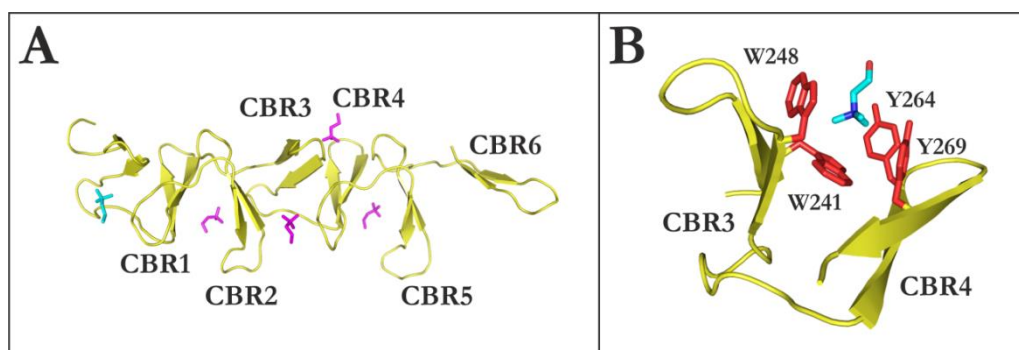


Figure 2.3. Choline-binding sites in the choline-binding domain (CBD) of pneumococcal autolysin LytA. **A.** Crystal structure of LytA CBD (PDB code: 1HCX) showing the choline-binding sites. Choline molecules are represented in magenta for the classical binding sites, and cyan for the additional binding site reported by some authors. (Mellroth et al., 2014) CBRs are labelled. **B.** Detail of the choline binding site located between CBR3 and CBR4. Residues involved in the interaction with choline are shown in red sticks and labelled.

1.2 APPLICABILITY OF THE STUDY OF LYTA

Since LytA was the first CBP reported, it is one of the most deeply studied. However, its complete structural determination in absence of ligands is complicated due to the difficulties in obtaining crystals with enough quality for X-ray diffraction. In the case of NMR studies, complications arise from the poor solubility of the protein and its tendency to dimerize at high concentration in absence of choline. (Beatriz Maestro et al., 2011; B. Maestro & Sanz, 2005)

LytA CBD is considered the most representative. As it is located in the C-terminal region of the protein, it is usually named as CLytA. The interest of studying this CBD comes from various aspects. One of them is to better understand folding processes and stability of β -solenoid proteins. On the other hand, as LytA is a protein expressed by a pathogen microorganism and it is actively involved in numerous functions related to its virulence, the study of LytA, its structure and its interaction with choline, is of potential therapeutic interest and is helpful for the rational design of efficient inhibitors. (B. Maestro & Sanz, 2005)

Besides the biological and clinical interest, study of LytA and its CBD has relevance for technologic applications. The utilization of the LytA CBD in chromatography has been reported. A simple amine-rich stationary phase fixed to chromatographic resins to separate hybrid proteins with a LytA CBD tag has been proposed. The proteins going through the chromatographic column are retained by the interaction between the CBD tag and the amines in the stationary phase. Then, elution is carried out adding a choline solution, since choline interacts more strongly with the LytA CBDs attached to the proteins. (B. Maestro & Sanz, 2005; Beatriz Maestro et al., 2008)

1.3 PEPTIDES DERIVED FROM LYTA

Given the importance of studying LytA and bearing in mind the mentioned difficulties, a very useful and simple approach employed in many studies is to work with peptide fragments from LytA, principally from its C-terminal domain (CLytA). Conformational investigations on protein fragments can provide data about the early stages of the protein global folding process. Many of these data are not accessible by kinetic studies of the whole protein folding reaction. (Blanco, Rivas, & Serrano, 1994)

Detection of conformations analogous to the native state of a protein in some of its fragments indicates that those structures can be present in the unfolded protein in native conditions. These structures can drive and facilitate the folding process, since they impose some restraints. Many isolated peptides in aqueous solution show secondary structure elements (α -helices, β -strands, β -turns...) which reproduce the folding pattern observed in the corresponding sequences of the whole, folded protein. This means that, in native conditions, many isolated protein sequences are capable of adopting specific conformations and originate the secondary structures that are part of the native protein. Therefore, it can be

concluded that the aforementioned elements of secondary structure can act as initiators of the protein global folding process. (Blanco et al., 1994)

Considering all the above mentioned, study of LytA, and more specifically, of its CBD (CLytA), can be addressed from the investigation of peptide fragments corresponding to sequences taken from that region. Dr. J. M. Sanz and collaborators followed this approach and studied a 14-residue peptide derived from the sequence of the CBR1 of CLytA (Figure 2.2). (Beatriz Maestro et al., 2011) This fragment corresponds to the β -hairpin core of the first choline-binding repeat of LytA CBD (Figure 2.4). They demonstrated the ability of that small peptide to fold autonomously in aqueous solution to adopt a stable native-like β -hairpin structure. They also confirmed the ability of this peptide to bind choline, the natural ligand of the original CBD. These results pointed to the idea of a cooperative action of the CBRs to give rise to a strong binding of choline in the whole domain. These observations also paved the way for developing new robust β -hairpin-forming peptides of biotechnological interest.

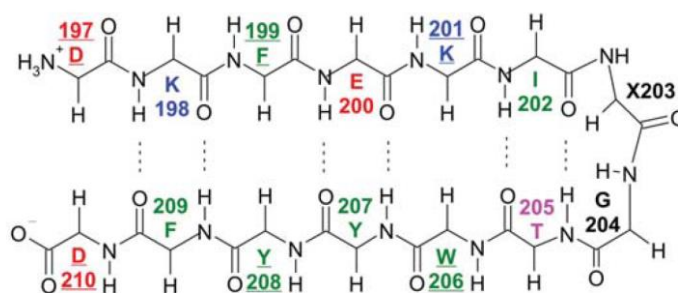


Figure 2.4. Schematic representation of peptide LytA₁₉₇₋₂₁₀. This peptide, whose sequence corresponds to the β -hairpin core of CBR1 from pneumococcal LytA was studied in a previous work. The β -hairpin shown in the representation is the native structure, which is also maintained in aqueous solution. (Beatriz Maestro et al., 2011)

This chapter includes a detailed study of new peptides derived from pneumococcal LytA. The analysis of one of these peptides, LytA₂₃₉₋₂₅₂, in the presence of detergent micelles revealed an unexpected conformational switch behaviour. This observation led us to extend the approach followed for the study of these peptides, and not only the choline-binding ability was examined, but also the conformational switching properties were deeply investigated.

2. CHAMELEON SEQUENCES

Secondary structure elements have been always viewed as the fundamental building blocks of proteins. It is well known that only local sequence composition and position of the amino acids is not enough to define (and predict) the secondary structure. The importance of non-local interactions has been brought to light since the mid-1980s, when identical amino acid sequences showing different secondary structure elements were identified within

proteins. (Kabsch & Sander, 1984) This kind of sequences, called chameleon sequences (ChSeqs), are defined as sequence strings of identical amino acids that can adopt different conformations in protein structures (Figure 2.5). (Ghozlane, Joseph, Bornot, & De Brevern, 2009; Wenlin Li, Kinch, Karplus, & Grishin, 2015; Minor & Kim, 1996)

ChSeqs were discovered in proteins, and as a result of several studies, many chameleon sequences of different lengths were reported. These sequences usually have no strong preference for either α - and β -conformation, and thus they can show both structures depending on the protein. ChSeqs described so far vary from 6 to 10 residues long, or even 11 residues in a designed sequence inserted in a bigger domain. Bioinformatics research has led to the confection of a large database containing thousands of ChSeqs. Some authors have defined two different classes of chameleon sequences known as HS sequences (with helix and strand conformations) and HE sequences (with helix and sheet/hairpin conformations). (Ghozlane et al., 2009; Guo, Jaromczyk, & Xu, 2007; Wenlin Li et al., 2015; Takano et al., 2007)

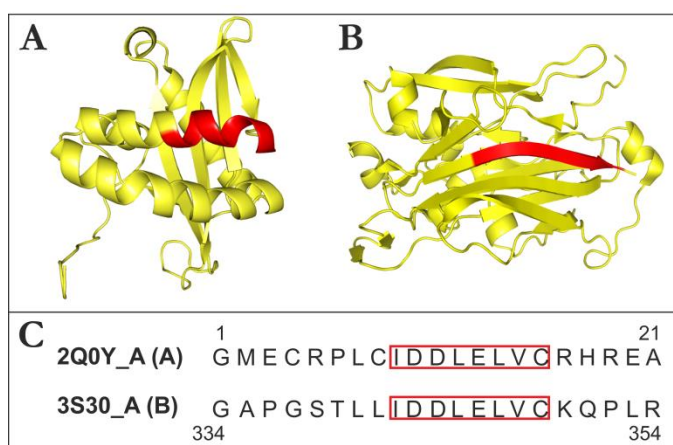


Figure 2.5. Chameleon sequences. Two identical sequences (**C**, boxed in red) adopt different secondary structures in two different proteins. In the N-acetyltransferase depicted in **A** (PDB code: 2Q0Y), the sequence forms a helical structure (in red). In the case of the putative glycoside hydrolase shown in **B** (PDB code: 3S30), the same sequence adopts an extended structure (in red).

The study of chameleon sequences has been of great interest in the last decades, because they seem to disagree with the well-known Anfinsen's dogma (also called "thermodynamical hypothesis"). This postulate states that the native conformation of a protein is determined by the amino acid sequence. Chameleon sequences show different conformations in distinct proteins, in spite of having identical primary structures. This demonstrates that the environment, the non-local interactions, are determining for the secondary and tertiary structure. Actually, Anfinsen indicated in his hypothesis that the amino acid sequence determines the protein conformation in a given environment, and thus, chameleon sequences do not call the hypothesis into question, rather they have evidenced the great importance of the environment and the non-local interactions to define the structure. (Anfinsen, 1973; Epstein, Goldberger, & Anfinsen, 1963; Guo et al., 2007)

3. CONFORMATIONAL SWITCHES

As mentioned before, chameleon sequences are identical amino acid stretches adopting different secondary structures in function of the protein of which they form part. However, conformational switches are defined as delimited protein regions capable of adopt different secondary structures depending on changes in the environment, such as pH variations, ligand binding, or posttranslational modifications. Structural changes of conformational switches are reversible. (Ambroggio & Kuhlman, 2006)

Conformational switches can be found everywhere in nature and they are usually related to the regulation of key biological processes. Conformational switches can involve transitions from disordered to ordered states, transitions between different ordered states, or changes in the oligomerization state. Changes can affect to the whole protein or only to a portion of it. (Ambroggio & Kuhlman, 2006)

3.1 FEATURES OF THE CONFORMATIONAL SWITCHING

Some features to be considered in conformational switches are the following: (Bryan & Orban, 2010)

- Conformational stability

A conformational state corresponding to an energy well deeper than $-5 \text{ kcal}\cdot\text{mol}^{-1}$ is favoured, as usually are native states of proteins, and thus, any alternative folding or unfolding process is not prone to occur. However, many circumstances can result in a stability diminution of the native state, increasing the possibility of adopting alternative folds. Alternative topologies distinct to the native state are usually driven by weak propensities persisting in the unfolded state. These propensities are dictated by the polypeptide nature and can arise from steric repulsions, buried hydrophobic surface residues, and main chain hydrogen bonds.

- Disordered segments

Disordered segments in the polypeptide chain can hide latent fold propensities which manifest themselves only under certain conditions, such as specific structural contexts, or changes in the solvent properties.

- Binding–surface changes

The establishment of new binding interactions can trigger a fold switching, because folding and binding are usually two thermodynamically related processes. Protein association or small ligand binding events can generate a conformational switching, which stabilizes the complex, and it usually leads to a rearrangement of the tertiary or quaternary interactions.

- Reversibility

Conformational switches are able to undergo changes in their structure when variations in the environment take place but, generally, these changes are reversible and the recovering of the initial conditions leads to the adoption of the original structural features.

3.2 CLASSES OF CONFORMATIONAL SWITCHES

Conformational switching can be triggered by a variety of causes, involving changes in the molecular interactions which maintain the native fold. In this regard, a series of different families of conformational switches can be found:

- Temperature switches

It has been established that thermal activation can control and direct the conformation of systems with inherent structural ambiguity. Generally, temperature has been exploited for generating $\alpha \rightarrow \beta$ switches. Several temperature-dependent switches based on the α -helical coiled coil folding motif have been described. (Ciani, Hutchinson, Sessions, & Woolfson, 2002; Kammerer et al., 2004; Kammerer & Steinmetz, 2006) Furthermore, several non-coiled coil-based systems that undergo temperature-induced conformational switching have been also reported, such as the conversion of a helix-turn-helix conformation to a β -sheet rich amyloid structure described by Teplow and co-workers. (Fezoui et al., 2000) It is also well known that amyloid formation, which involves the transition to β structures, is favoured with the increase of temperature. (Zhang & Rich, 1997)

- Polarity switches

Since Coulombian interactions are a major contributor to protein folding and stability, conformational switching can sometimes be triggered by the modulation of the charges present in the amino acid side chains, or by changes in the polarity, ionic strength or pH of the solvent. For example, coiled coil-based peptides which undergo conformational switching upon variation in the ionic strength or pH have been reported. (Weijun Li, Nicol, & Szoka Jr, 2004) In some cases, His residues are used to get this effect, as this amino acid possesses exceptional acid-base properties. Histidine is hydrophobic above pH~6 (pK_a of His is ~ 6), and thus can be part of the apolar core of folded peptides. However, below pH~6, His gets protonated and promotes the destabilization of the hydrophobic core (Pagel & Koksche, 2008) Changes in polarity of the solvent caused, for instance, by the mixture with other types of solvent, are known to induce structural transition in peptides. This is the case of many peptides which show random coil conformations in aqueous solution and become helical in the presence of TFE. (Hansen, Ruizendaal, Löwik, & van Hest, 2009; Luo & Baldwin, 1997; Marinelli, Castillo, & Ventura, 2013)

- Photoswitches

Some designed peptides are able to experiment conformational switching thanks to the presence of moieties, called chromophores, which change their configuration or constitution upon absorption of light. This strategy is advantageous because it does not require the addition of reactants and the response is fast. Chromophores are sensitive to certain wavelengths and, when irradiated, they usually experience a photoisomerization which forces the conformational switching in the peptide. Photoswitchable peptides with transitions from an unstructured state or a β -sheet conformation to α -helix upon irradiation have been reported. (Bredenbeck, Helbing, Kumita, Woolley, & Hamm, 2005; Cerpa, Cohen, & Kuntz, 1996; Hansen et al., 2009; Lungu et al., 2012; Nguyen, Gorbunov, & Stock, 2006)

- Redox switches

Since some amino acids have redox properties, such as cysteine and methionine, changes in the oxidation state can produce conformational transitions. The most usual way to induce structural switching by redox perturbations is to modify the hydrophobic/hydrophilic amino acid balance by oxidizing the sulphur moiety of Met or Cys side chains, generally to sulfoxides or sulphinic acid, respectively. In both cases, the final product has a higher polarity and is more hydrophilic. (Dado & Gellman, 1993; Hansen et al., 2009) Some approaches are also based on the formation/cleavage of Cys disulphide bonds. (Park & Raines, 2001)

- Metal-ion switches

Some peptides can bind to metal ions, and a specific secondary structure is stabilized upon the binding. For example, formation of α -helix is typically favoured if metal-binding groups are located in positions i , and $i+4$ or $i+5$; whereas β structures or turns can be induced when the binding sites are further apart of each other. (Albrecht & Stortz, 2005; Ghadiri & Fernholz, 1990) The most frequent metal-binding groups are N-heterocycles, such as imidazole, pyridine and bipyridine, and for this reason, His residues are common in metallopeptides. Sometimes other coordinating groups are introduced in the peptide structure, such as crown ethers or phosphano serine. (Gilbertson, Chen, & McLoughlin, 1994; Rossi, Tecilla, Baltzer, & Scrimin, 2004) The effect of metal coordination in peptides can be affected by pH variations or changes in the redox conditions. (Hansen et al., 2009; Wang, Bergenfeld, Arora, & Canary, 2012)

- Other switches

Other strategies can be explored to produce new peptide switches. One interesting example is switches based on X-N acyl migration. These switches are constructed with two peptide segments coupled via a switchable building block employing an ester (X=O) or thioester (X=S) bond, and a protected α -amino group. When the switchable building block is off, both peptide segments adopt their native conformations independently. The system can go

to the on state by a cleavage of the α -amino protecting group mediated by photolysis, enzymatic activity, pH variation, etc. Subsequently, an X–N acyl migration occurs and a new peptide bond is created. As a result, both peptide segments are no longer uncoupled and adopt the specific conformation of the full-length peptide. (Hansen et al., 2009; Mutter et al., 2004; Pagel & Koksche, 2008; Sohma et al., 2005)

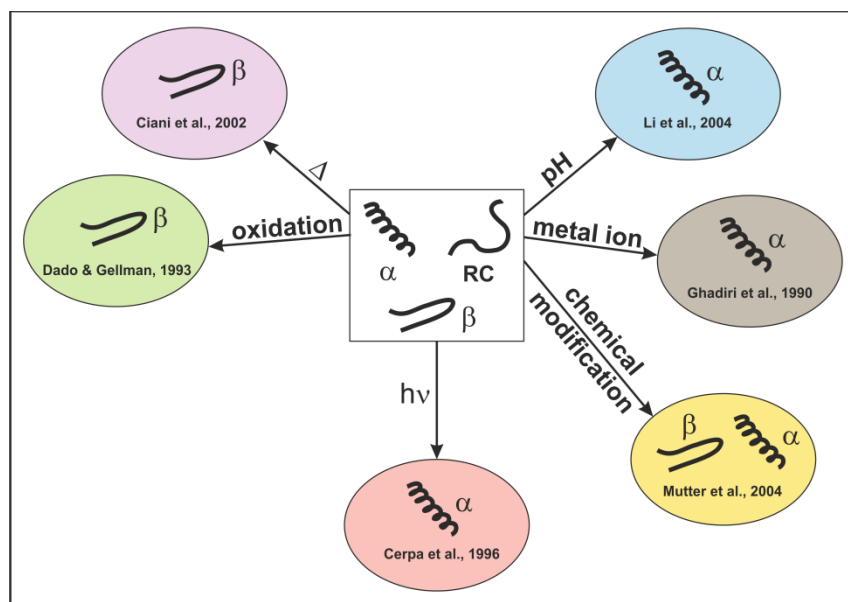


Figure 2.6. Scheme of the different kinds of conformational switches. Arrows show the change produced and the triggering agent is indicated. References for an example of each kind are included.

3.3 RELEVANCE OF THE CONFORMATIONAL SWITCHING

The study of sequences that behave as conformational switches has been a major focus in the last years since they are a useful tool to unveil the relationship between the primary structure and the conformational and physicochemical properties, and to understand the impact of this features on biological processes. (Saucède et al., 2006)

It has been demonstrated the importance of the role of conformational transitions in peptides related to human diseases. Probably, the most emblematic example is the soluble amyloid β precursor forming aggregated β -sheet type forms which are crucial for the deposition of amyloid plaques in Alzheimer's disease. The transformation of PrP^c form of the prion protein to the pathologic PrP^{sc} is considered to have its origin in α -helix to β -sheet transitions. In general, many studies have been realized on switch peptides related to amyloid structures formation in order to understand the dynamics of the transition and ultimately facilitate the rational design of therapeutic agents. (Fezoui et al., 2000; Saucède et al., 2006)

Conformational switches have also been proposed as the basis for developing novel biosensors. In fact, nature uses this kind of transitions to transduce stimuli (such as changes in the environment properties, presence of certain molecules, etc.) into specific biochemical outputs. In that respect, various designs have been reported, some based on naturally occurring switches and others designed *de novo*. (Cerasoli, Sharpe, & Woolfson, 2005; Pandya et al., 2004; Vallée-Bélisle & Plaxco, 2010)

OBJECTIVES

In this chapter, a series of peptides derived from the choline-binding domain of pneumococcal autolysin LytA have been investigated. The main motives inspiring this study were to get a deeper knowledge of the folding process of this kind of proteins; to find the minimal sequence able to maintain the choline-binding ability; and to obtain valuable information about a protein involved in the virulence of *S. pneumoniae*, a common human pathogen.

The specific objectives proposed in this chapter are:

- To find whether minimal peptides encompassing sequences from LytA choline-binding domain are capable of maintaining their native structure in different solvent conditions and their choline-binding ability.
- To check the effect of different solvent conditions on the peptide structures.
- To explore the behaviour of the peptides in membrane-like media, such as detergent micelles or lipid vesicles.
- To propose explanations for the observed results.

MATERIALS AND METHODS

1. PEPTIDE SYNTHESIS

LytA_{239–252}, LytA_{259–272} and LytA_{239–272} peptides, and the five designed LytA_{239–252} variants (K3W5–LytA_{239–252}, W5K10–LytA_{239–252}, S3S10–LytA_{239–252}, I3V10–LytA_{239–252} and I5Y6T11T13–LytA_{239–252}) were synthesized by Caslo Aps (Lingbym, Denmark). SESYV11 and SESYW11 peptides were purchased from DiverDrugs (Barcelona, Spain) and Lipotec (Barcelona, Spain), respectively. In all cases, Fmoc–solid phase synthesis protocols were followed and purification was performed by reverse–phase HPLC using C18 columns until reaching 95 % or more purity. The identity of the peptides was confirmed by an LC–MS system equipped with an 1100 Series LC-system from Agilent (Santa Clara, CA, USA) coupled to an ESI micro–TOF mass spectrometer from Bruker Daltonics (Bremen, Germany).

LytA_{239–252} (TGWKKIADKWYYFN): RP–HPLC: t_R = 9.99 min, purity 95.60 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 25–40 % B buffer gradient in 15 min). HRMS: Theoretical molecular mass: 1820.10; found: 1821.65 [M + H]⁺.

LytA_{259–272} (TGWVKYKDTWYYLD): RP–HPLC: t_R = 10.35 min, purity 95.02 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 25–40 % B buffer in 15 min). HRMS: Theoretical molecular mass: 1838.07; found: 1838.70 [M + H]⁺.

LytA_{239–272} (TGWKKIADKWYYFNEEGAMKTGWVKYKDTWYYLD): RP–HPLC: t_R = 11.61 min, purity 95.28 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 25–45 % B buffer in 20 min). HRMS: Theoretical molecular mass: 4285.89; found: 4286.56 [M + H]⁺.

K3W5–LytA_{239–252} (TGKKWIADKWYYFN): RP–HPLC: t_R = 10.56 min, purity 99.62 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 22–37 % B buffer gradient in 15 min). HRMS: Theoretical molecular mass: 1820.10; found: 1820.17 [M + H]⁺.

W5K10–LytA_{239–252} (TGWKWIADKKYYFN): RP–HPLC: t_R = 10.78 min, purity 98.68 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 22–37 % B buffer gradient in 15 min). HRMS: Theoretical molecular mass: 1820.10; found: 1821.11 [M + H]⁺.

S3S10–LytA_{239–252} (TGSKKIADKSYYFN): RP–HPLC: t_R = 11.99 min, purity 99.29 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 12–30 % B buffer gradient in 18 min). HRMS: Theoretical molecular mass: 1621.82; found: 1623.78 [M + H]⁺.

I3V10–LytA_{239–252} (TGIKKIADKVYYFN): RP–HPLC: $t_R = 9.92$ min, purity 98.13 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 18–33 % B buffer gradient in 15 min). HRMS: Theoretical molecular mass: 1659.96; found: 1660.17 [M + H]⁺.

I5Y6T11T13–LytA_{239–252} (TGWKIYADKWTYTN): RP–HPLC: $t_R = 9.87$ min, purity 98.56 % (buffer A: 0.05 % TFA in H₂O/CH₃CN 98:2; buffer B: 0.05 % TFA in H₂O/CH₃CN 1:9; linear 22–37 % B buffer gradient in 15 min). HRMS: Theoretical molecular mass: 1746.96; found: 1747.70 [M + H]⁺.

SESYV11 (SESYNSDGTVTVTE): RP–HPLC: $t_R = 14.7$ min, purity 95.8 % (buffer A: 0.1 % TFA in H₂O; buffer B: 0.07 % TFA in CH₃CN; linear 15–25 % B buffer gradient in 30 min). HRMS: Theoretical molecular mass: 1600.72; found: 1624.11 [M + Na]⁺.

SESYW11 (SESYNSDGTWTVTE): RP–HPLC: $t_R = 10.9$ min, purity 95.5 % (buffer A: 0.1 % TFA in H₂O; buffer B: 0.07 % TFA in CH₃CN; linear 21 % B buffer gradient in 20 min). HRMS: Theoretical molecular mass: 1687.72; found: 1688.85 [M + H]⁺.

2. C–LYTA DOMAIN PURIFICATION

LytA C–terminal domain (residues 192–318, corresponding to the choline–binding domain) was expressed using the overproducing *Escherichia coli* strain RB791 harbouring pCE17 plasmid (Sánchez-Puelles, Sanz, Garcia, & Garcia, 1990) and purified by the group of Dr. J. M. Sanz (Universidad Miguel Hernández, Elche, Spain), by affinity chromatography. Purified samples were subsequently dialyzed at 5 °C against 20 mM sodium phosphate buffer pH 7.0, plus 50 mM NaCl, to remove the choline used for elution, and stored at –20 °C. Protein concentration was determined spectrophotometrically as described previously (Sánchez-Puelles et al., 1990), by using a molar extinction coefficient at 280 nm of 62 540 M^{–1}·cm^{–1}. (Pace, Vajdos, Fee, Grimsley, & Gray, 1995)

3. PREPARATION OF SMALL UNILAMELLAR VESICLES (SUVs)

For vesicle preparation, different lipids were used, all of them purchased from Avanti Polar Lipids, Inc. (Alabaster, AL, USA):

- **DMPC** (1,2–dimyristoyl–sn–glycero–3–phosphocholine) (PC 14:0/14:0, purity >99 %)
- **DMPG** (1,2–dimyristoyl–sn–glycero–3–phospho–(1'–rac–glycerol), sodium salt) (PG 14:0/14:0, purity >99 %)
- **POPE** (1–palmitoyl–2–oleoyl–sn–glycero–3–phosphoethanolamine (PE 16:0/18:1 (9Z), purity > 99%)

- **POPG** (1–palmitoyl–2–oleoyl–sn–glycero–3–phospho–(1′–rac–glycerol), sodium salt) (PG 16:0/18:1 (9Z), purity > 99%)

All lipid powders were dissolved in chloroform/methanol 50:50 (v/v) to obtain 5 mg·mL⁻¹ stock solutions. Aliquots of the stock solutions were mixed in glass vials and thoroughly vortexed to obtain DMPC:DMPG 3:1 (molar ratio) and POPE:POPG 2:1 (molar ratio) mixtures. Subsequently, the organic solvents were removed under a gentle stream of nitrogen, followed by overnight vacuum. The lipid film formed at the bottom of the vials was dispersed by addition of 10 mM phosphate buffer at pH 7.0, and 10 min of vigorous vortexing. The suspensions of multilamellar vesicles (MLVs) were then homogenized by ten freeze–thaw cycles followed by 1 min vortexing after each cycle. Small unilamellar vesicles (SUVs) were then formed by sonication of MLVs for 16 min at 35°C in a strong ultrasonic bath (UTR200, Hielscher, Germany).

4. CIRCULAR DICHROISM (CD)

CD spectra were recorded with Jasco J–815 spectropolarimeters (Tokyo, Japan) equipped with either a Peltier PTC–423S system (samples in water and in detergent micelles) or a water–thermostatted rectangular cell holder (samples in vesicles). Peptide concentrations were determined from the 280 nm UV absorbance by using the corresponding molar extinction coefficients (Table 2.1). (Pace et al., 1995)

Table 2.1. List of molar extinction coefficients of the peptides.

Peptide	Molar extinction coefficient (M ⁻¹ ·cm ⁻¹)
LytA _{239–252}	13,980
LytA _{259–272}	15,470
LytA _{239–272}	29,450
K3W5–LytA _{239–252}	13,980
W5K10–LytA _{239–252}	13,980
S3S10–LytA _{239–252}	2,980
I3V10–LytA _{239–252}	2,980
I5Y6T11T13–LytA _{239–252}	13,980

For samples in aqueous solution and in detergent micelles, the peptide concentration was approximately 30 μM and the cuvette path lengths were 0.1 cm for far–UV region (250–195 nm) and 1.0 cm for near–UV region (320–250 nm). Samples were centrifuged 5 min prior to CD measuring, although no visible precipitate was seen. All measurements were carried out in triplicate at 5 and 25 °C in the presence of 20 mM glycine buffer at pH 3.0 or of 20 mM sodium phosphate buffer at pH 7.0. Isothermal wavelength spectra for these samples were acquired at a scan speed of 50 nm·min⁻¹ with a response time of 2 s and averaged over at least six scans.

Samples in vesicles were prepared by adding an aliquot of a 0.3 mM peptide stock solution in water to either DMPC:DMPG (3:1) or POPE:POPG (2:1) vesicle dispersion in 10 mM phosphate buffer at pH 7.0. The final peptide concentration was adjusted to obtain a

peptide-to-lipid molar ratio (P/L) of 1:50, and it was around 28 μM . CD spectra for these samples were measured by using a quartz glass cell (Suprasil, Hellma, Mülheim, Germany) of 1 mm path length between 260 and 185 nm at 0.1 nm intervals. Spectra were recorded at 30 °C (i.e., above the phase-transition temperature of the lipids). Three repeat scans at a scan-rate of 10 nm·min⁻¹, 8 s response time and 1 nm bandwidth were averaged for each sample and for the baseline of the corresponding peptide-free sample.

After subtracting the baseline spectra from the sample spectra, CD data were processed with the adaptative smoothing method, which is part of the Jasco Spectra Analysis software. Molar ellipticities ($[\theta]$) were expressed in unit of deg·cm²·dmol⁻¹, using the residue concentration of peptide.

For CD-monitored thermal denaturation experiments, the sample was layered with mineral oil to avoid evaporation, and the heating rate was 60 °C·h⁻¹. Thermal scans were fitted by least squares to the Gibbs-Helmholtz equation (Eq. 2.1) in which $\Delta G^0(T)$ is the free energy of the transition at a temperature T , ΔH_m is the van't Hoff enthalpy, T_m is the midpoint of denaturation (in Kelvin) and ΔC_p is the difference in heat capacity between the native and denatured states.

$$\Delta G^0 = \Delta H_m \cdot \left(1 - \frac{T}{T_m}\right) - \Delta C_p \cdot \left[(T_m - T) + T \cdot \ln \frac{T}{T_m}\right] \quad (2.1)$$

Stabilisation free energies (ΔG^0) were calculated from the CD titration traces (Eq. 2.2) in which K_{eq} is the equilibrium constant between the initial and final states, $[\theta]_I$ and $[\theta]_F$ are the ellipticities of the initial and final state, respectively, and $[\theta]_x$ is the experimental ellipticity at a given temperature.

$$\Delta G^0 = -RT \ln K_{eq} = -RT \ln \frac{[\theta]_I - [\theta]_x}{[\theta]_x - [\theta]_F} \quad (2.2)$$

For choline titration, independent peptide samples were prepared in the presence of different ligand concentrations, and incubated for 5 min prior to recording the wavelength spectra. Binding was analysed according to a Langmuir analysis (Eq. 2.3), considering only one binding site per peptide, in which $\Delta[\theta]_{293}$ is the change in ellipticity at 293 nm at each point, $\Delta[\theta]_{293}(\text{max})$ is the change in ellipticity at ligand saturation, and K_d is the dissociation constant.

$$\Delta[\theta]_{293} = \frac{\Delta[\theta]_{293}(\text{max}) \cdot [\text{choline}]}{K_d + [\text{choline}]} \quad (2.3)$$

5. NMR SAMPLE PREPARATION

NMR samples were prepared by solving each lyophilised peptide (1–2 mg) in 0.5 mL of solvent (see Table 2.2). Peptide concentrations were 0.5–1.0 mM, except where another concentration is indicated. pH was adjusted to 3.0 by adding minimal amounts of NaOD or DCl, measured with a glass micro-electrode and not corrected for isotopic effects. Approximate peptide/detergent ratios are indicated in each case. Peptide and detergent were equimolar in the samples at sub-micellar detergent concentrations. All the samples were placed in 5 mm NMR tubes, and contained DSS as internal reference for ^1H chemical shifts.

Table 2.2. List of solvents used in NMR sample preparation.

Solvent
$\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
D_2O (99.96 % purity)
30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
30 mM $[\text{D}_{38}]$ -DPC in D_2O
15 mM $[\text{D}_{38}]$ -DPC + 15 mM DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
0.5 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
0.5–0.6 mM $[\text{D}_{38}]$ -DPC in D_2O
30 mM $[\text{D}_{25}]$ -SDS in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
30 mM $[\text{D}_{25}]$ -SDS in D_2O
15 mM $[\text{D}_{25}]$ -SDS + 15 mM SDS in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
0.2 mM $[\text{D}_{25}]$ -SDS in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
0.2 mM $[\text{D}_{25}]$ -SDS in D_2O
30 % vol. $[\text{D}_3]$ -TFE + 70 % vol. $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v)
30 % vol. $[\text{D}_3]$ -TFE + 70 % vol. D_2O

6. NMR SPECTRA ACQUISITION

NMR spectra were recorded with a Bruker Avance-600 spectrometer operating at a proton frequency of 600.1 MHz and equipped with a cryoprobe, the temperature of which was calibrated by using a methanol sample. 1D ^1H NMR spectra were acquired by using 32 K data points, which were zero-filled to 64 K data points prior to Fourier transformation. Phase-sensitive two-dimensional correlated spectroscopy (COSY), total correlated spectroscopy (TOCSY) and nuclear Overhauser enhancement spectroscopy (NOESY) spectra were recorded by standard techniques using the time-proportional phase increment mode. Water signal was suppressed by either presaturation or by using a 3–9–19 pulse sequence. TOCSY spectra were obtained by using 60 ms DIPSI2 with z filter spin-lock sequence. NOESY mixing time was 150 ms. ^1H – ^{13}C heteronuclear single quantum coherence (HSQC) spectra were recorded at ^{13}C natural abundance. Acquisition data matrices had 2048×512 points in t_2 and t_1 , respectively. Data were processed with the standard TOPSPIN program (Bruker Biospin, Karlsruhe, Germany). The 2D data matrices were multiplied by a square-sine-bell window function with the corresponding shift optimised for every spectrum and zero-filled to 2×1 K complex matrices prior to Fourier transformation. Baseline correction was applied in both dimensions. ^{13}C δ -values were indirectly referenced by using the IUPAC–IUB recommended ^1H : ^{13}C chemical shift ratio (0.25144953). (Markley et al., 1998)

7. NMR SPECTRA ASSIGNMENT

^1H NMR signals of the studied peptides in each solvent condition were assigned by analyses of the 2D NMR spectra using the SPARKY software (Goddard & Kneller, 2004) and the standard sequential assignment strategy. (Wuthrich, 1986; Wüthrich, Billeter, & Braun, 1984) The ^{13}C resonances were identified on the basis of the correlations between the protons and the bound carbon atoms present in the ^1H - ^{13}C -HSQC spectra. These chemical shifts are listed in Tables A1, A2, and A3 (see “Appendices”).

8. STRUCTURE CALCULATION

Structure calculation was done by following a two-step protocol. First, we applied the standard iterative procedure for automatic NOE assignment of the CYANA 2.1 program, which performs seven cycles of combined automated NOE assignment and structure calculation of 100 conformers per cycle. (Güntert, 2004) As experimental input data, we used the lists of: i) assigned chemical shifts, ii) NOE integrated cross-peaks present in 150 ms NOESY spectra acquired in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v and in pure D_2O , and iii) φ and ψ dihedral angle restraints, which were derived from ^1H and ^{13}C chemical shifts using TALOS+ webserver. (Shen, Delaglio, Cornilescu, & Bax, 2009) Integration of NOE cross-peaks was performed by the automatic integration subroutine of SPARKY software. (Goddard & Kneller, 2004) For every calculation, the list of upper limit distance constraints resulting from the last automatic cycle was checked by inspection of the corresponding NOESY spectra, and ambiguous constraints were removed or relaxed to generate the final list used as input for a standard simulated annealing CYANA 2.1 calculation of 100 conformers. The final ensembles of the 20 lowest target function structures were visualised and examined by using programs MOLMOL (Koradi, Billeter, & Wüthrich, 1996) and PyMol (Schrodinger, 2015); and their quality was assessed by using PROCHECK/NMR as implemented at the Protein Structure Validation Suite server (PSVS server: <http://psvs-14-dev.nesg.org/>).

9. FLUORESCENCE MEASUREMENTS

Fluorescence measurements were carried out at 25 °C with a PTI–QuantaMaster fluorimeter (Birmingham, NJ, USA), model QM–62003SE, using a 5×5 mm path-length cuvette and a peptide concentration of 1 μM . Buffer was 20 mM glycine buffer at pH 3.0. Tryptophan emission spectra were obtained by using an excitation wavelength of 280 nm, with excitation and emission slits of 1.0 nm and a scan rate of 60 $\text{nm}\cdot\text{min}^{-1}$. The critical micelle concentration (cmc) of DPC in 20 mM glycine buffer at pH 3.0 and 25 °C was determined according to the procedure of Chattopadhyay and London (Chattopadhyay & London, 1984), using DPH as a fluorescence probe. The cuvette path length was 10×10 mm, and excitation and emission slits were set to 1 nm. Excitation wavelength was 360 nm.

For acrylamide quenching experiments, independent peptide samples at about 30 μM were incubated for 5 min with different acrylamide concentrations in the presence or absence of 30 mM of DPC, and the wavelength spectrum was recorded. For each sample, a blank

without peptide was subtracted from the recorded spectrum. Experiments were repeated at least three times. Data were analysed with the Stern–Volmer equation (Eq. 2.4), (Stern & Volmer, 1919) in which F_0 and F are the fluorescence intensities at 340 nm in the absence and presence of quencher, respectively, K_{SV} is the Stern–Volmer constant and $[Q]$ is the quencher concentration.

$$\frac{F_0}{F} = 1 + K_{SV} \cdot [Q] \quad (2.4)$$

For $\text{LytA}_{259-272}$ and $\text{LytA}_{239-272}$, choline titrations monitored by fluorescence were performed in order to achieve a good calculation of the dissociation constant (K_d). For that purpose, independent peptide samples were prepared in the presence of different ligand concentrations, and incubated for 5 min prior to recording the fluorescence spectra. A sample with 1000 mM NaCl was used as a control. Binding was analysed according to a Langmuir analysis (Eq. 2.5), considering only one binding site per peptide, in which $\Delta F_{330/350}$ is the increase of 330/350 fluorescence intensity at each point, $\Delta F_{330/350}(\text{max})$ is the increase of 330/350 fluorescence intensity at ligand saturation, and K_d is the dissociation constant.

$$\Delta F_{330/350} = \frac{\Delta F_{330/350}(\text{max}) \cdot [\text{choline}]}{K_d + [\text{choline}]} \quad (2.5)$$

10. EFFECT OF PARAMAGNETIC COMPOUNDS ON NMR SPECTRA

Samples of 0.5–1 mM $\text{LytA}_{239-252}$ in 30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v) pH 3.0 were titrated with three paramagnetic compounds: one hydro-soluble, MnCl_2 , and two liposoluble, 5-doxyl-stearic acid (free radical) and methyl-16-doxyl-stearate (free radical). On the other hand, samples of 0.5–1 mM $\text{LytA}_{259-272}$ and $\text{LytA}_{239-272}$ in 30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v) pH 3.0 were titrated only with the hydro-soluble paramagnetic compound MnCl_2 . Titrations were performed by adding aliquots (5–30 μL) from stock solutions of the paramagnetic agents, and monitored by 2D ^1H - ^1H -TOCSY spectra acquired at 25 °C at each titration point. The stock solutions were 10–40 mM MnCl_2 in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v) pH 3.0, 13 mM 5-doxyl-stearic acid in deuterated methanol ($[\text{D}_4]$ -MeOH), and 12.6 mM methyl-16-doxyl-stearate in $[\text{D}_4]$ -MeOH.

RESULTS AND DISCUSSION

1. STRUCTURAL CHARACTERIZATION OF $\text{LytA}_{239-252}$ AND $\text{LytA}_{259-272}$

$\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ encompass the sequences of the β -hairpin core of CBR3 and CBR4 from LytA choline-binding domain, respectively. These sequences were selected because they possess the most stable β -hairpin turn type from those present in the protein domain, excluding the previously studied CBR1. (Beatriz Maestro et al., 2011) In addition, the selection of two sequences from consecutive CBRs allows the construction of a longer peptide containing both hairpins, which corresponds to the theoretical minimal choline-binding unit (Figures 2.2 and 2.3).

1.1. CD STUDY OF $\text{LytA}_{239-252}$ AND $\text{LytA}_{259-272}$ IN AQUEOUS SOLUTION

Far and near-UV CD spectra were recorded for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ peptides in aqueous solution at pH 3.0 (20 mM HCl-glycine buffer). The strong positive bands at 227 nm observed in the far-UV spectra of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ (Figure 2.7A and 2.7C (solid lines), respectively), together with the significant near-UV CD signals (Figure 2.7B and 2.7D, (solid lines)), resemble those observed for the full-length C- LytA (Figure 2.46B, solid line, page 94) (B. Maestro & Sanz, 2005; Sánchez-Puelles et al., 1990; Sanchez-Puelles, Sanz, Garcia, & Garcia, 1992) that have been described as arising from aromatic rings in rigid conformations. Therefore, these results suggest that peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ form a well-ordered, native-like structure in aqueous solution.

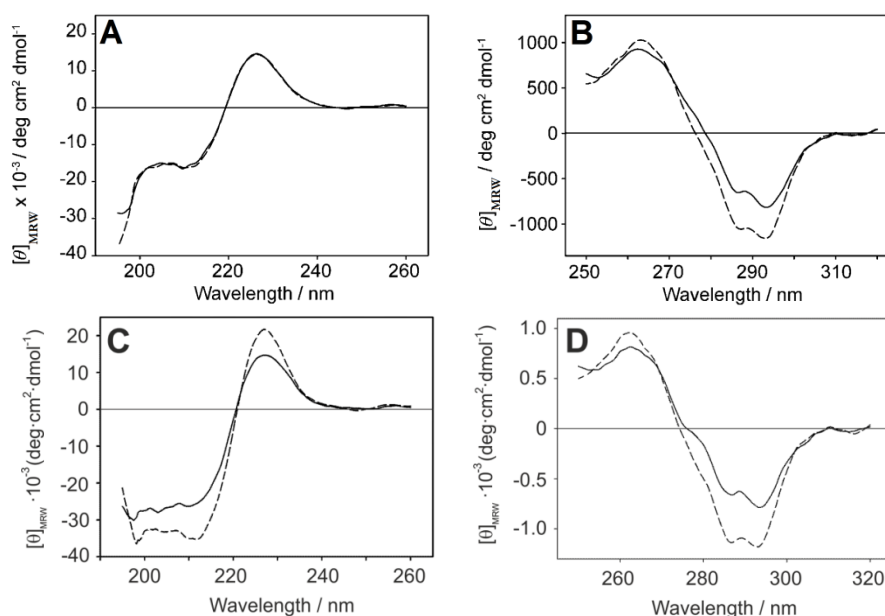


Figure 2.7. CD spectra acquired for 27.5 μM $\text{LytA}_{239-252}$ (top) and $\text{LytA}_{259-272}$ (bottom) aqueous samples. **A** and **C** correspond to the far-UV region and **B** and **D** to the near-UV region. Solid lines are from samples in aqueous solution (20 mM HCl-glycine buffer, pH 3.0), and dashed lines represent samples in the same solvent plus 500 mM choline. All spectra were recorded at 5 $^{\circ}\text{C}$, and mean residue ellipticities are shown.

The ability of these peptides to bind choline was also examined by CD analysis. As observed in Figures 2.7B and 2.7D, the near-UV bands at 293 nm (attributable to Trp side chains) and 286 nm (from Tyr and Trp) become more intense in the presence of the ligand (500 mM). This suggests that the peptides are able to bind choline even if the complete choline-binding site is not present. The far-UV CD spectrum of $\text{LytA}_{239-252}$ is not affected by choline (Figure 2.7A), whereas the spectrum of $\text{LytA}_{259-272}$ shows an increase of the intensity, as observed in the near-UV CD spectra (Figure 2.7C). The far-UV CD spectrum of $\text{LytA}_{239-252}$ is the only one unaffected by choline, as changes have been verified in the corresponding spectra of the previously analysed $\text{LytA}_{197-210}$ (Beatriz Maestro et al., 2011) and the full length C-LytA. (B. Maestro & Sanz, 2005; Sánchez-Puelles et al., 1990; Sanchez-Puelles et al., 1992) The lack of change in the far-UV CD spectrum of $\text{LytA}_{239-252}$ can be explained by the fact that, as demonstrated later by NMR, this peptide has already acquired completely the secondary structure in the absence of ligand, meanwhile $\text{LytA}_{259-272}$, the free C-LytA module (B. Maestro & Sanz, 2005) and $\text{LytA}_{197-210}$ (Beatriz Maestro et al., 2011) are only partially folded in solution and need choline additionally to fully acquire structure.

Thermal stability of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ was examined by analysing the CD spectra before and after heating, and after cooling the sample again, both in the absence and in the presence of 500 mM choline. For both peptides, heating led to featureless CD spectra, both in the presence and absence of choline, pointing to a massive loss of structure, which was reversible upon cooling (Figure 2.8).

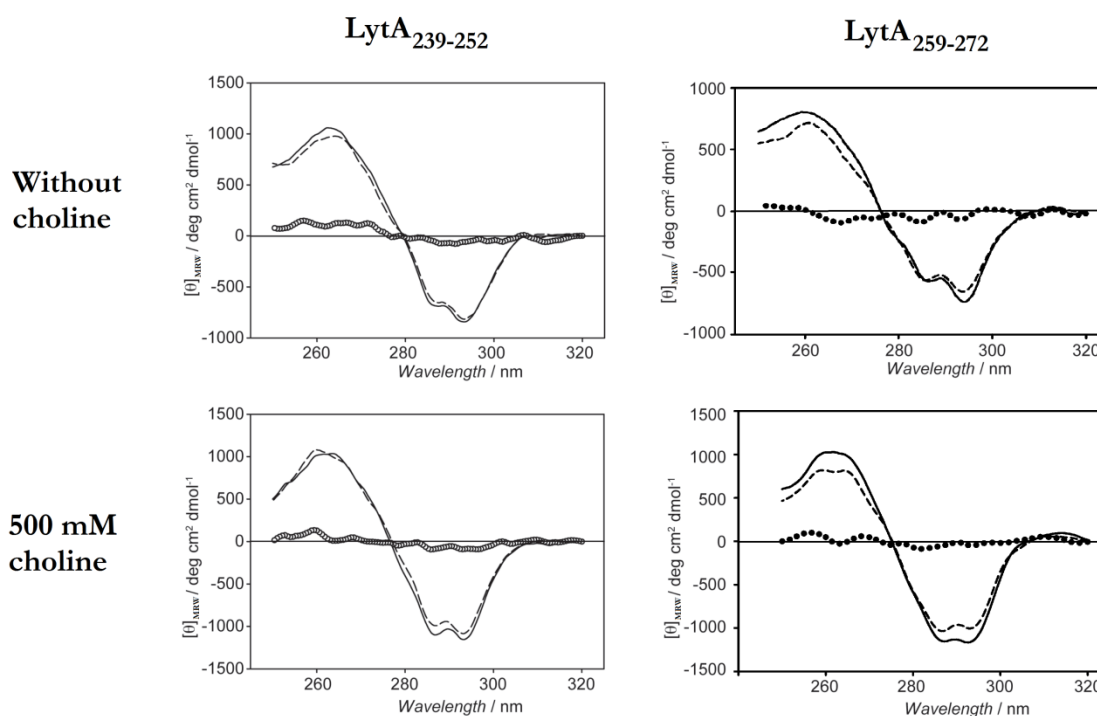


Figure 2.8. Thermal unfolding of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ monitored by near-UV CD. Left panels show the CD spectra for $\text{LytA}_{239-252}$ in the absence (**top**) and the presence (**bottom**) of 500 mM choline. The same is shown for $\text{LytA}_{259-272}$ in the right panels. In each panel a superposition of three CD spectra is represented: samples at ~ 5 °C (solid lines), samples heated at ~ 90 °C (dotted lines), and samples cooled at ~ 5 °C after heating (dashed lines).

A deep study of the thermal stability of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ was performed by monitoring the temperature dependence of the molar ellipticity at 293 nm and 294 nm, respectively, in the range 5–90 °C (Figure 2.9). Transition data were fitted to the Gibbs–Helmholtz equation (see “Material and methods” section, Eq. 2.1) assuming an approximated ΔC_p of $1.6 \text{ kJ}\cdot\text{mol}^{-1}$, the value reported for the 12-residue, tryptophan zipper, trpzip4 hairpin. (Cochran, Skelton, & Starovasnik, 2001) However, data fitting using this parameter was very poor (data not shown). Therefore, because the thermal transitions show very little cooperativity, as expected from the lack of sizeable hydrophobic cores and tight packing around the aromatic residues (see below), the contribution of ΔC_p was assumed negligible, which is an approximation already followed for other β -hairpin peptides. (Honda, Kobayashi, & Munekata, 2000; Santiveri et al., 2002; Xu, Oyola, & Gai, 2003) The thermodynamic parameters for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ calculated by using this approach in the absence and the presence of 500 mM choline (in HCl–glycine buffer at pH 3.0), are shown in Table 2.3.

Table 2.3. Thermodynamic parameters calculated for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in HCl–gly buffer at pH 3.0.

	$\text{LytA}_{239-252}$		$\text{LytA}_{259-272}$	
	0 mM choline	500 mM choline	0 mM choline	500 mM choline
$\Delta H_m (\text{kJ}\cdot\text{mol}^{-1})$	38 ± 4	50 ± 4	44 ± 5	52 ± 4
$T_m (\text{K})$	$321 \pm 2 (48^\circ\text{C})$	$326 \pm 1 (53^\circ\text{C})$	$308 \pm 2 (35^\circ\text{C})$	$320 \pm 1 (47^\circ\text{C})$
$\Delta G (25^\circ\text{C}) (\text{kJ}\cdot\text{mol}^{-1})$	2.5 ± 0.4	4.2 ± 0.4	1.4 ± 0.3	3.6 ± 0.4

The stabilities of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ at 25 °C are intermediate between those of the 15-residue SESYW11 hairpin ($0.1 \text{ kJ}\cdot\text{mol}^{-1}$), (Santiveri et al., 2002) and the tryptophan zippers trpzip3 ($4.6 \text{ kJ}\cdot\text{mol}^{-1}$) and trpzip4 ($6.3 \text{ kJ}\cdot\text{mol}^{-1}$), although in the latter cases the Trp/Trp stacking contributes decisively to stability. (Cochran et al., 2001) For $\text{LytA}_{259-272}$, the stability is lower than that observed for $\text{LytA}_{239-252}$ ($1.1 \text{ kJ}\cdot\text{mol}^{-1}$ lower, Table 2.3).

Results for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in the presence of 500 mM choline (Table 2.3) demonstrate that interaction with choline increases the stability of the peptides. It is noteworthy that, in both peptides, the two unfolding traces (this is in the presence and absence of choline), converge at around 70 °C (340 K) (Figure 2.9A), indicating that the peptides are competent to bind choline up to these high temperatures. In both peptides, there is certain cooperativity in the presence of choline, pointing to the presence of a folded core that is lost in a two-state transition.

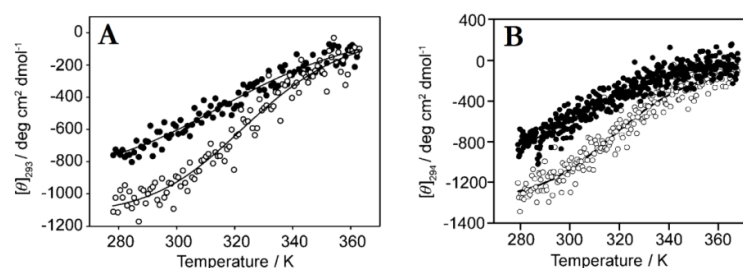


Figure 2.9. Thermal unfolding of peptides $\text{LytA}_{239-252}$ (A) and $\text{LytA}_{259-272}$ (B) monitored by CD, in the presence (○) and absence (●) of 500 mM choline. Solid lines represent the fits to the Gibbs–Helmholtz equation (Eq. 2.1)

The affinity of $\text{LytA}_{239-252}$ for choline was calculated by recording near-UV CD spectra at 25 °C and different ligand concentrations. The plot of the change in molar ellipticity at 293 nm against choline concentration was fitted to Eq. 2.3 (see “Materials and methods” section), assuming one binding site, so that the dissociation constant (K_d) was calculated as 80 ± 10 mM (Figure 2.10).

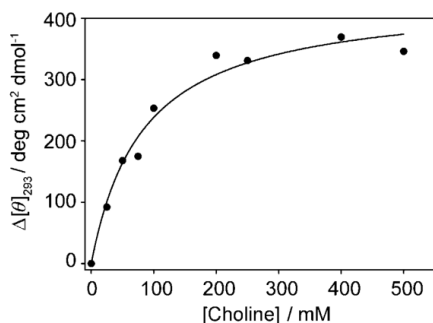


Figure 2.10. CD-monitored titration of $\text{LytA}_{239-252}$ with choline ($\Delta[\theta]_{293} = [\theta]_{293}^{\text{no choline}} - [\theta]_{293}^{\text{choline}}$). Data were fitted to a simple binding model (Eq. 2.3).

In the case of $\text{LytA}_{259-272}$, an equivalent procedure based on a titration monitored by CD was first followed to calculate the affinity for choline, but the fitting showed a poor quality. To solve this situation, a titration monitored by fluorescence was employed instead. $\text{LytA}_{259-272}$ was titrated with increasing amounts of choline (0–1000 mM), and the curve representing the augmentation of the fluorescence intensity at 330/350 against choline concentration was fitted to Eq. 2.3. The dissociation constant (K_d) for $\text{LytA}_{259-272}$ was calculated as 294 ± 28 mM (Figure 2.11).

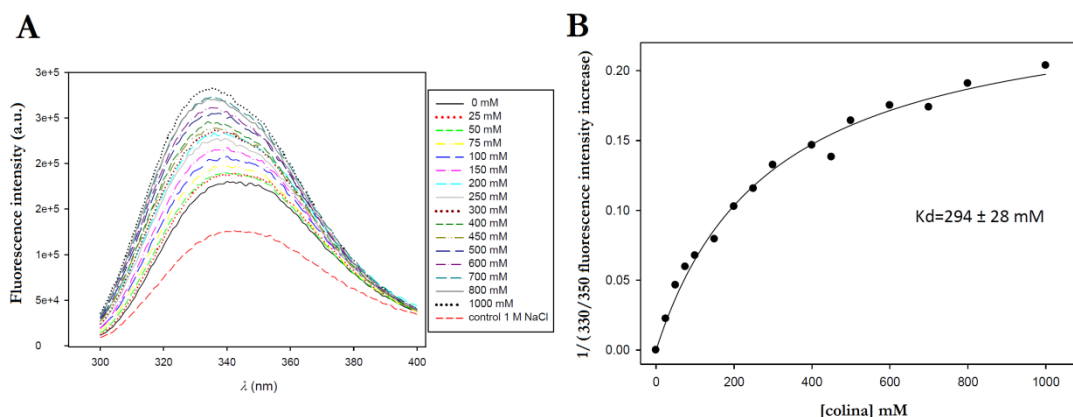


Figure 2.11. Affinity of $\text{LytA}_{259-272}$ for choline by fluorescence spectroscopy. **A.** Titration of a $\text{LytA}_{259-272}$ sample with choline (0–1000 mM), in HCl–Gly buffer, at pH 3.0 and 25 °C. A sample with 1000 mM NaCl was used as a control. **B.** Fitting of the 330/350 fluorescence intensity increase to Eq. 2.5 to obtain the dissociation constant (K_d) for $\text{LytA}_{259-272}$.

1.2. NMR STRUCTURE CALCULATIONS IN AQUEOUS SOLUTION

To determine the structures adopted by peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in aqueous solution, 1D and 2D NMR spectra were recorded using ~ 1 mM samples. The ^1H and ^{13}C chemical shifts were assigned by following a standard strategy (see “Materials and methods” section). The non-sequential NOE cross-peaks observed in 2D ^1H – ^1H –NOESY

spectra include those characteristic of antiparallel β -sheets (Figure 2.1A, B, C and D); that is, those between the H_α protons of residues facing each other in non-hydrogen-bonded sites, and between amide protons of residues facing each other in hydrogen-bonded sites (see Figure 2.12). The presence of these NOEs shows that peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ in aqueous solution adopt β -hairpin structures, and that the β -strand registers are native-like. Also, numerous non-sequential NOEs involving side chain protons (Table 2.4) are indicative of native-like β -hairpins. Formation of β -hairpin structures is further confirmed by the plot of $\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of peptide sequences; that is, two stretches of positive $\Delta\delta H_\alpha$ and $\Delta\delta C_\beta$ values, and negative $\Delta\delta C_\alpha$ corresponding to residues at N- and C-terminal strands, which are separated by $\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ values of the corresponding opposite sign at the turn region (Fesinmeyer et al., 2005; Santiveri et al., 2002) (Figure 2.12E and F, and Figures A1 and A2 in “Appendices”). Based on the averaged $\Delta\delta H_\alpha$ values at the strand residues (+0.42 ppm and +0.30 at 25 °C, for LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂, respectively) and considering that the averaged $\Delta\delta H_\alpha$ value at protein β -strands is +0.40 ppm, (Wishart, Sykes, & Richards, 1991) the β -hairpin populations formed in aqueous solution at pH 3.0 and 25 °C are approximately 100 % and 76 % for LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂, respectively. This demonstrates that LytA₂₃₉₋₂₅₂ is a more robust hairpin than LytA₂₅₉₋₂₇₂, and the previously studied LytA₁₉₀₋₂₁₀ peptide (63 % β -hairpin population in aqueous solution at 15 °C). (Beatriz Maestro et al., 2011)

Table 2.4. Summary of non-sequential cross-strand NOEs involving side chain protons observed for the peptides LytA₂₃₉₋₂₅₂ (Zamora-Carreras et al., 2015), and LytA₂₅₉₋₂₇₂ in aqueous solution.

Side chains from non-HB residues	Peptide	
	LytA ₂₃₉₋₂₅₂	LytA ₂₅₉₋₂₇₂
Facing residues 5/10	K5/W10	K5/W10
No. NOEs	15	20
Facing residues 3/12	W3/Y12	W3/Y12
No. NOEs	8	10
Facing residues 1/14	T1/N14	T1/D14
No. NOEs	3	3
Diagonal residues 5/12	K5/Y12	K5/Y12
No. NOEs	0	0
Diagonal residues 3/10	W3/W10	W3/W10
No. NOEs	11	4
Diagonal residues 1/12	T1/Y12	T1/Y12
No. NOEs	1	1
Side chains from HB residues		
Facing residues 6/9	I6/K9	Y6/T9
No. NOEs	2	4
Facing residues 4/11	K4/Y11	V4/Y11
No. NOEs	0	7
Diagonal residues 6/11	I6/Y11	Y6/Y11
No. NOEs	11	2
Diagonal residues 4/13	K4/F13	V4/L13
No. NOEs	12	0

*Numbering 1 to 14 is used to simplify the data, correspondence to protein numbering: 1 \equiv 239, 14 \equiv 252 for LytA₂₃₉₋₂₅₂; 1 \equiv 259, 14 \equiv 272 for LytA₂₅₉₋₂₇₂.

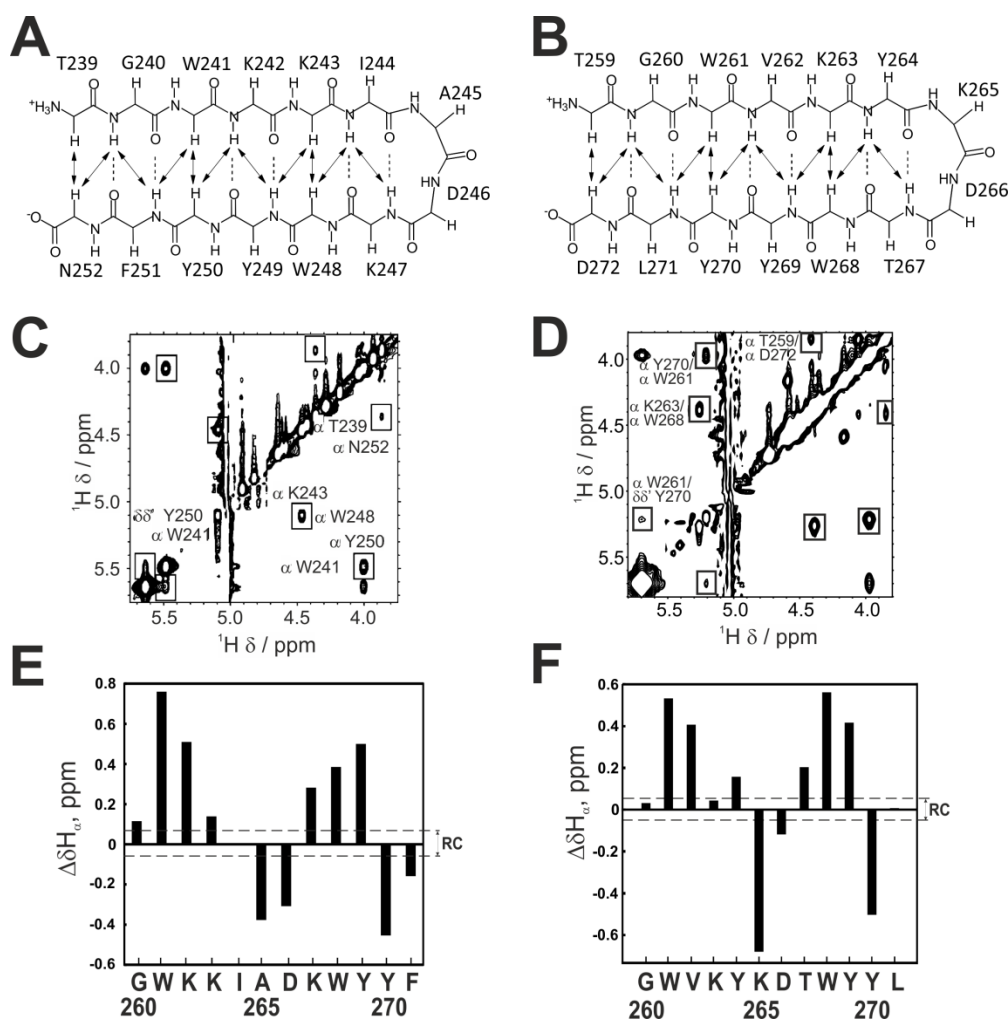


Figure 2.12. NMR data for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. **A** and **B**. Schematic representations of the 2:2 β -hairpins formed by $\text{LytA}_{239-252}$ (**A**) and $\text{LytA}_{259-272}$ (**B**) in aqueous solution. Double arrows indicate the NOEs observed in 2D NOESY spectra. **C** and **D**. 2D NOESY spectra (H_α region) of $\text{LytA}_{239-252}$ (**C**) and $\text{LytA}_{259-272}$ (**D**) in D_2O at pH 3.0 and 5 °C. Non-sequential NOEs are boxed and labelled at one of the diagonal sides. **E** and **F**. Bar plots of $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_\alpha^{\text{observed}} - \delta H_\alpha^{\text{RC}}$) as a function of sequence for peptides $\text{LytA}_{239-252}$ (**E**) and $\text{LytA}_{259-272}$ (**F**) in H_2O/D_2O 9:1 v/v at pH 3.0 and 25 °C. $\delta H_\alpha^{\text{RC}}$ values were taken from (Wishart et al., 1995) Values for the two Gly H_α protons are plotted. The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

Table 2.5. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ peptides in aqueous solution.

	$\text{LytA}_{239-252}$	$\text{LytA}_{259-272}$
Total number of restraints		
Upper limit distances	172	203
φ and ψ dihedral angles	24	24
Pairwise RMSD (Å)		
Backbone atoms	0.3 ± 0.1	0.09 ± 0.04
All heavy atoms	1.0 ± 0.1	0.8 ± 0.2

*RMSD were calculated only with residues 241-250 and 261-270, as the amino acids located close to the termini are more flexible.

To obtain further details of the features of these structures, structure calculations were performed on the basis of the distance restraints derived from the complete sets of observed NOEs and the dihedral angle restraints obtained from the $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts by using the program TALOS+ (Shen et al., 2009) (Table 2.5).

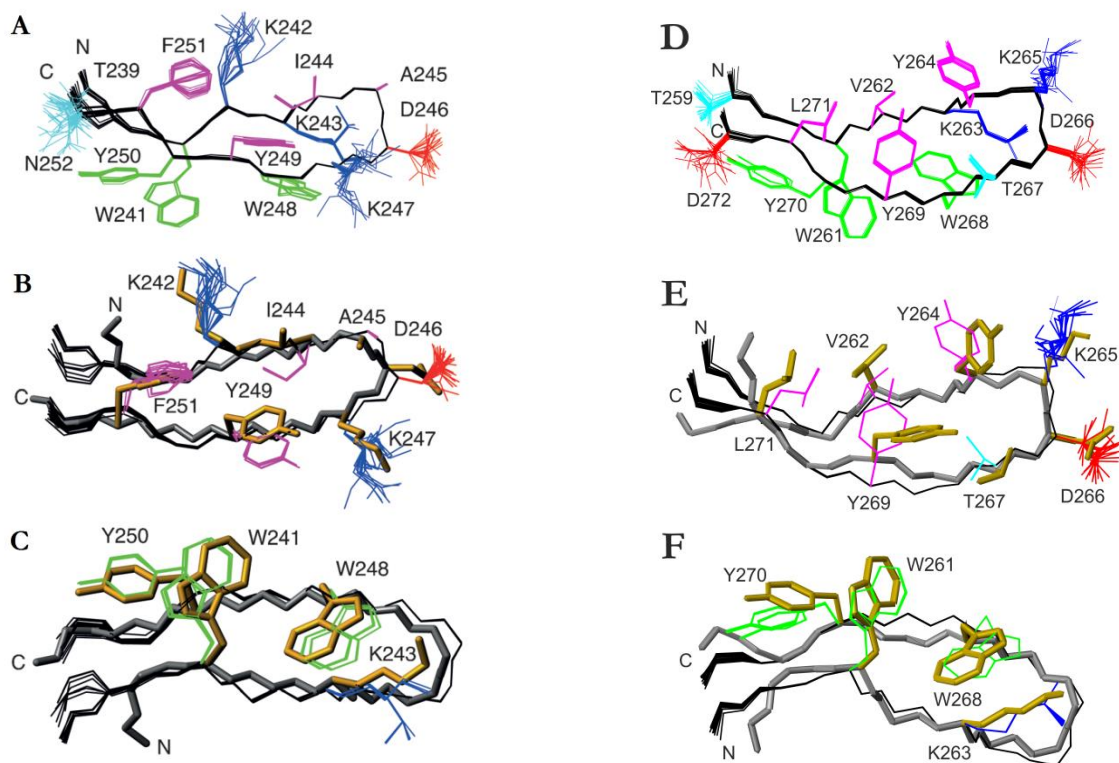


Figure 2.13. Structures of $\text{LytA}_{239-252}$ (left) and $\text{LytA}_{259-272}$ (right) in aqueous solution: **A** and **D** Ensemble of the 20 lowest target function structures overlaid onto the backbone atoms (black). Side chains colour code: blue for positively charged; red for negatively charged; cyan for polar; green and magenta for aromatic and hydrophobic side chains pointing downwards and upwards, respectively. **B**, **C**, **E**, and **F** show backbone atoms (in black) overlaid onto the corresponding atoms in the crystalline C-LytA structure (PDB code: 1GVM, in thick grey) for $\text{LytA}_{239-252}$ (**B** and **C**) and $\text{LytA}_{259-272}$ (**E** and **F**). Side chains of residues at the turn region and at hydrogen-bonded sites are shown in **B** and **E**, and those of residues at non-hydrogen-bonded sites in **C** and **F**. Side chains are coloured following the same colour code mentioned before, and those of C-LytA in gold. The amino and carboxylate termini are labelled by “N” and “C”, respectively.

The calculated structures (Figure 2.13) are well defined, as indicated by the small pairwise RMSDs presented by the backbone atoms (Table 2.5). Both $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ show structures in aqueous solution that are very similar to their native structures. Backbone RMSD values were calculated by fitting each calculated conformer to the native structure and obtaining the next average values: $0.738 \pm 0.008 \text{ \AA}$ and $0.7887 \pm 0.0003 \text{ \AA}$, respectively), as can be appreciated in Figure 2.13B, C, E and F. In addition, the side chains are also ordered in all the residues, except for the N- and C-terminal ones, since their χ_1 dihedral angles show very little variation among the 20 calculated structures in both structures (χ_1 variation range $< 4^\circ$ for most residues), and display orientations quite similar to those in the native protein (sidechain RMSD: $1.53 \pm 0.06 \text{ \AA}$ for $\text{LytA}_{239-252}$, and $1.73 \pm 0.04 \text{ \AA}$ for $\text{LytA}_{259-272}$) (Figure 2.13).

1.3. EFFECT OF DPC MICELLES

In an attempt to check whether regions of LytA could interact with the cell membrane to translocate from the cytoplasm to the cell wall without a signal peptide (López, García, García, & García, 1997), the structure of LytA_{239–252} and LytA_{259–272} was studied in the presence of DPC micelles, since they represent a simple membrane model, commonly used for solution NMR studies. (Mäler, 2013; Sanders & Sönnichsen, 2006) First, far-UV CD spectra for both peptides in the presence of 30 mM DPC (peptide/detergent ratio \approx 1:1000) were recorded (Figure 2.14).

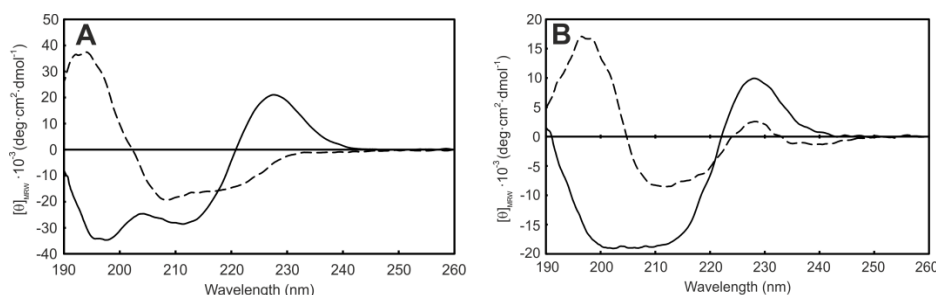


Figure 2.14. Far-UV CD spectra of LytA_{239–252} (A) and LytA_{259–272} (B) in the presence of DPC micelles (30 mM DPC, peptide/detergent ratio \approx 1:1000). For each peptide, the spectra in the absence (solid lines) and the presence of DPC micelles (dashed lines) are shown. All spectra were acquired at 30 °C in HCl–Gly buffer at pH 3.0.

Addition of detergent cause noticeable changes in the aromatic-dominated spectra of the peptides. In LytA_{239–252}, a minimum at 209 nm, a shoulder at 222 nm, and a maximum at 197 nm, characteristic of α -helices with some contribution of β -structures, are observed (Figure 2.14) (Johnson Jr, 1988). In the case of LytA_{259–272}, the minimum at 209 nm and the shoulder at 222 nm, are less obvious pointing to α -helical spectrum with a contribution of β -structures larger than the observed for LytA_{239–252}. Even the maximum at 227 nm observed in aqueous solution is not totally absent in the presence of DPC micelles for LytA_{259–272}, which is in agreement with a bigger presence of residual β -structures.

A titration of LytA_{239–252} with increasing concentrations of DPC is shown in Figure 2.15. Transition from the β -hairpin to the α -helix occurs cooperatively and independently of the monitored wavelength. Moreover, the CD spectral transition occurs with an isodichroic point at 217 nm (Figure 2.15A). These two facts suggest that the structural conversion takes place between two states and without any detectable intermediates.

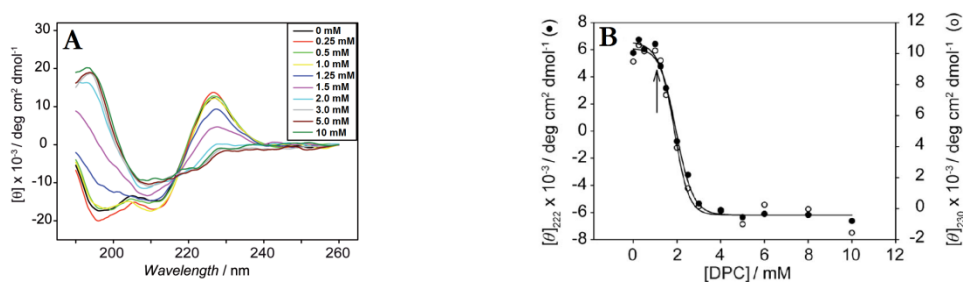


Figure 2.15. A. Titration of LytA_{239–252} with DPC monitored by CD. **B.** Evolution of ellipticity at 222 nm (●) and 230 nm (○) was followed. The arrow indicates the critical micelle concentration (cmc) of DPC.

When subjected to a thermal scan, the helical structure of $\text{LytA}_{239-252}$ in 30 mM DPC is gradually lost in a non-cooperative way (Figure 2.16, inset), indicating the lack of a detectable hydrophobic core, although unfolding was reversible upon cooling (Figure 2.16).

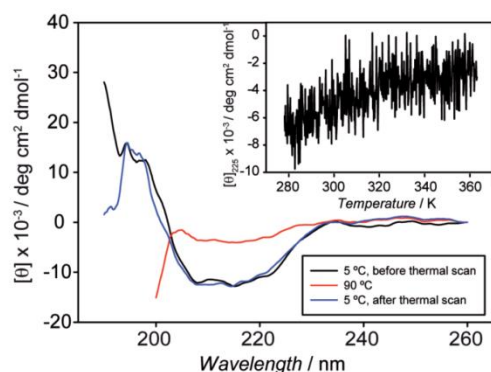


Figure 2.16. Thermal unfolding of $\text{LytA}_{239-252}$ in the presence of 30 mM DPC. Far-UV spectra recorded at 5 °C (black line), then at 90 °C (red line), and finally at 5 °C after cooling down the sample (blue line) are shown. **Inset:** temperature dependence of the ellipticity at 225 nm.

In order to determine whether DPC monomers or micelles are responsible for the structural transition observed in these peptides, the critical micelle concentration (cmc) of DPC was determined under the experimental conditions used here (20 mM HCl-Gly buffer, pH 3.0), giving a value of 1.2 mM (Figure 2.17). As depicted in Figure 2.15B, this concentration is right at the onset of the cooperative transitions observed. Therefore, interaction of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ with DPC micelles may drive the conformational change of the peptides to a largely helical structure.

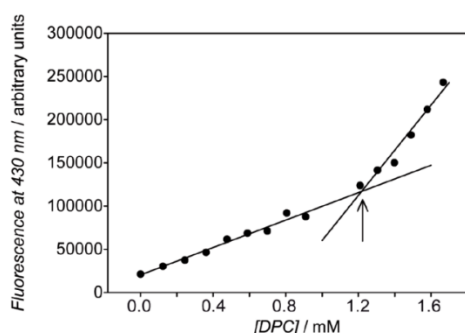


Figure 2.17. Determination of the critical micelle concentration (cmc) of DPC in 20 mM HCl-Gly buffer, pH 3.0, using DPH as probe (see Materials and methods). The arrow points to the determined cmc value.

Next step was the characterization of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in the presence of DPC by NMR spectroscopic analysis. First, 1D and 2D NMR spectra of the peptides in the presence of DPC in a concentration below the cmc, were acquired (0.45 mM for $\text{LytA}_{239-252}$ and 0.6 mM for $\text{LytA}_{259-272}$). Under these conditions, the NMR spectra were essentially identical to those in aqueous solution (see 2D ^1H - ^{13}C -HSQC spectra, Figure 2.18). The profile of conformational shifts (Figure 2.18B and D) provides further confirmation that peptide $\text{LytA}_{239-252}$ in 0.45 mM DPC and peptide $\text{LytA}_{259-272}$ in 0.6 mM DPC adopt β -hairpin structures.

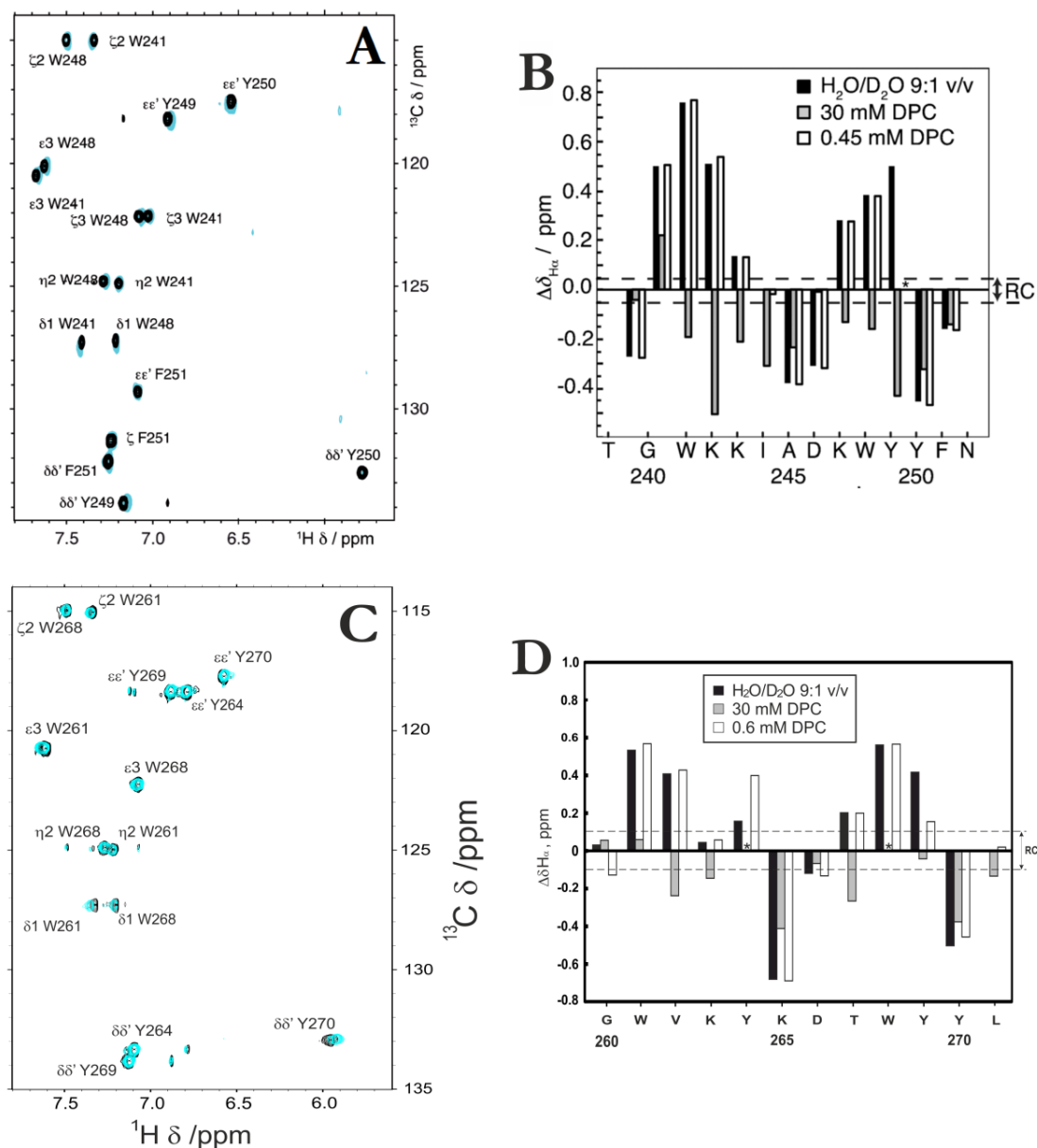


Figure 2.18. Effect of DPC below cmc on LytA₂₃₉₋₂₅₂ (top) and LytA₂₅₉₋₂₇₂ (bottom) peptides monitored by NMR. **A** and **C** Superposition of two 2D ^1H - ^{13}C -HSQC spectral regions showing the cross-peaks corresponding to the aromatic rings of Trp, Tyr, and Phe residues recorded at 25 °C and pH 3.0. Black spectra correspond to samples in D₂O, cyan spectra correspond to samples in the presence of DPC under the cmc (0.5 mM for LytA₂₃₉₋₂₅₂ and 0.6 mM for LytA₂₅₉₋₂₇₂). **B** and **D** Plots of $\Delta\delta_{\text{H}_\alpha}$ ($\Delta\delta_{\text{H}_\alpha} = \delta_{\text{H}_\alpha}^{\text{observed}} - \delta_{\text{H}_\alpha}^{\text{RC}}$, ppm) as a function of peptide sequence in H₂O/D₂O 9:1 v/v (black bars), 30 mM DPC (>cmc) (grey bars), and DPC below cmc (white bars). The N- and C-terminal residues are not shown. $\delta_{\text{H}_\alpha}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The asterisk means that the corresponding δ_{H_α} values were not determined. Measured at 25 °C and pH 3.0. Dashed lines indicate the random coil (RC)

As occurs in the case of CD spectra, NMR spectra of peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ in 30 mM DPC are very different from those in aqueous solution, as observed in the 1D ^1H NMR and 2D ^1H - ^{13}C -HSQC spectra shown in Figure 2.19. The observed differences look larger than would be expected to be observed based on the effect of solvent

on chemical shifts. Indeed, the profiles of conformational shifts in the presence of DPC micelles are completely different from those in aqueous solution (Figure 2.18). It is noticeable that those residues with positive $\Delta\delta H_\alpha$ values in aqueous solution have negative values in the presence of DPC micelles. Instead of a profile characteristic of β -hairpin structures, as observed in aqueous solution, the profile observed in the presence of DPC micelles is that typical of helices; that is, negative $\Delta\delta H_\alpha$ (Figure 2.18B and D) and $\Delta\delta C_\beta$ values, and positive $\Delta\delta C_\alpha$ values for residues 241–251 in LytA_{239–252}, and residues 262–271 in LytA_{259–272} (see “Appendices”, Figures A1 and A2).

In the case of LytA_{239–252}, further confirmation about the formation of a helix in DPC micelles comes from the set of NOEs, which include those characteristic of helices; i.e., $\alpha N(i,i+3)$, $\alpha\beta(i,i+3)$ and strong sequential $NN(i,i+1)$ (Figure 2.19C). In LytA_{259–272}, very few non-sequential NOE signals involving H_α protons were observed. However, $(i,i+3)$ NOE signals between nuclei from the side chains were identified, confirming the helical structure.

These results, in accordance with the CD data, confirm that the formation of helix in peptides LytA_{239–252} and LytA_{259–272} is induced by DPC micelles. However, the peptide/detergent ratio was different at a concentration of DPC under cmc (ca. 1:1), than at 30 mM DPC (1:30 or 1:60). Therefore, as an additional check that the conformational change occurs in the presence of DPC micelles and not by interaction with the DPC monomer, 1D 1H NMR spectra at a peptide/detergent ratio of 1:30, but at a sub-micellar DPC concentration (0.6 mM DPC and 0.02 mM peptide) were acquired. As seen in Figure 2.19A and D, 1D 1H NMR spectra acquired under these conditions are essentially identical to those recorded in aqueous solution, except for the signal-to-noise ratio. These samples were prepared by a 1:50 dilution in water of an aliquot of a ~ 1 mM LytA_{239–252} and LytA_{259–272} samples in 30 mM DPC, pH 3.0, in which peptides form α -helices. Hence, the fact that once diluted to a sub-micellar DPC concentration, their 1D 1H NMR spectra are identical to that of LytA_{239–252} and LytA_{259–272} in aqueous solution, in which the peptides form β -hairpin structures, provides evidence for the reversibility of the α -helix to β -hairpin transition, and confirms the role of DPC micelles on the peptide conformational change.

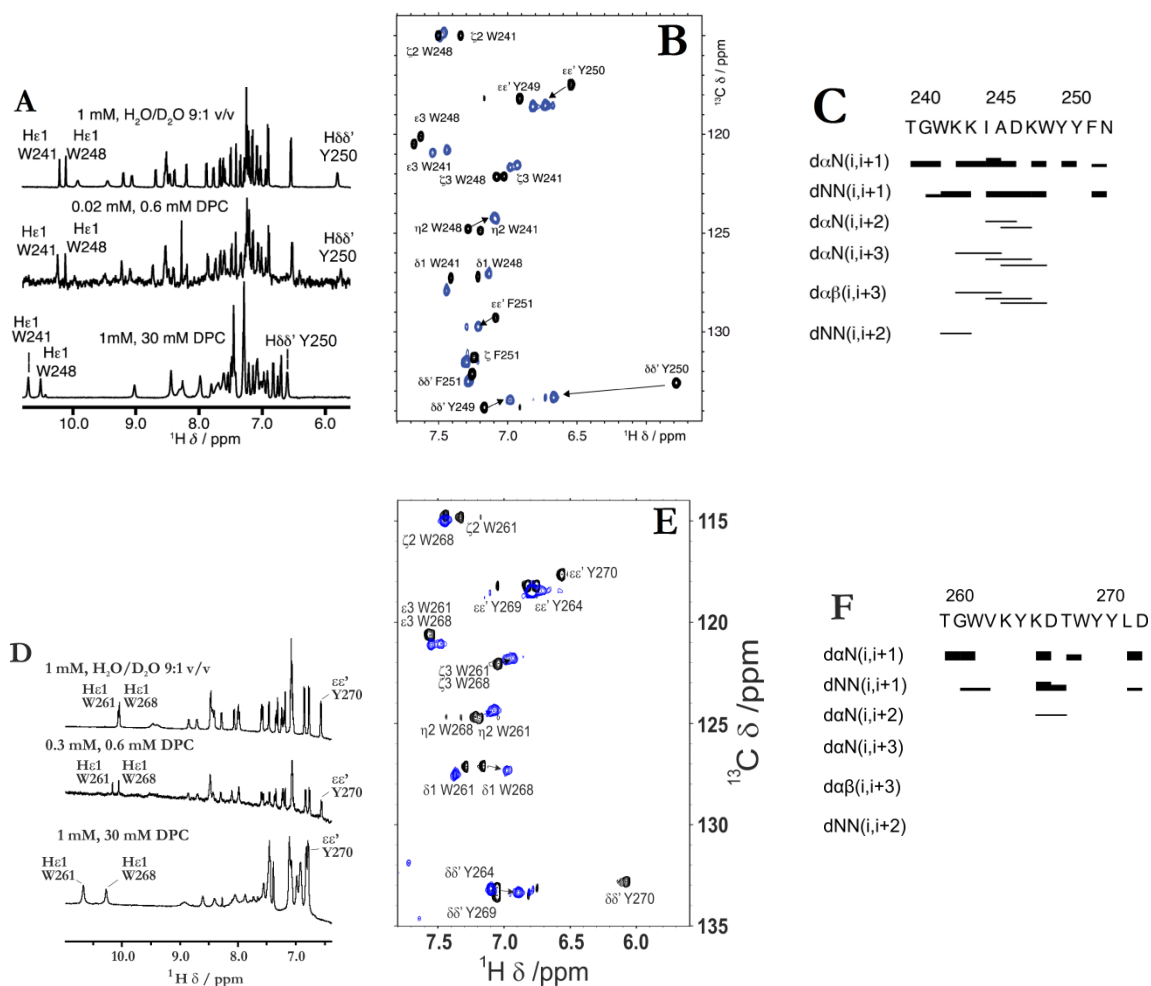


Figure 2.19. Effect of DPC above cmc on LytA_{239–252} (top) and LytA_{259–272} (bottom) peptides monitored by NMR. **A.** and **D.** Selected regions of the 1D ¹H NMR spectra of LytA_{239–252} and LytA_{259–272}, respectively, at 1 mM concentration in 30 mM DPC in H₂O/D₂O 9:1 v/v (bottom spectrum), at 0.02 mM concentration in 0.6 mM DPC in H₂O/D₂O 9:1 v/v (middle spectrum), and at 1 mM concentration in H₂O/D₂O 9:1 v/v (top spectrum), all spectra measured at pH 3.0 and 25 °C. **B.** and **E.** Superposition of two 2D ¹H–¹³C–HSQC spectral regions showing the cross-peaks corresponding to the aromatic rings of Trp, Tyr, and Phe residues recorded at 25 °C and pH 3.0. Black spectrum corresponds to samples in D₂O, cyan spectra corresponds to samples in the presence of 30 mM DPC. **C.** and **F.** NOE summaries for LytA_{239–252} and LytA_{259–272} in 30 mM DPC (thickness of the lines is proportional to NOE signal intensities).

Bearing in mind that the populations of the helix form estimated from the magnitude of the $\Delta\delta H_{\alpha}$ values and the averaged $\Delta\delta H_{\alpha}$ value at protein α -helices (-0.39 ppm)(Vila, Ponte, Suau, & Rico, 2000) were high (62 % for LytA_{239–252}, and 47 % for LytA_{259–272}, at 35 °C), the calculation of the peptide structures under these conditions were executed following the same protocol that in aqueous solution (see pages 62–63). The resulting structures are depicted in Figure 2.20. Helical backbones are well defined, as well as most side chains (see RMSD values in Table 2.6), which exhibit small ranges of variation ($< 10^{\circ}$) for the χ_1 dihedral angles in most residues (see “Appendices”, Tables A15 and A16). The packing of side chains in these α -helix structures (Figure 2.20) and in the β -hairpins formed in aqueous solution (Figure 2.13) is quite different. For example, in the case of LytA_{239–252} the

pairs W241/Y250 and I244/Y249 are close in the β -hairpin structure, but far away in the α -helix. Similarly, in $\text{LytA}_{259-272}$ the pairs W261/Y270 and Y264/Y269 are near in the β -hairpin structure, but distant in the α -helix (see Figures 2.13A, B and 2.13D, E, vs Figures 2.20A and 2.20C).

Table 2.6. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ peptides in 30 mM $[\text{D}_{38}]$ -DPC.

	$\text{LytA}_{239-252}$	$\text{LytA}_{259-272}$
Total number of restraints		
Upper limit distances	188	166
φ and ψ dihedral angles	24	22
Pairwise RMSD (Å)		
Backbone atoms*	0.3 ± 0.1	0.3 ± 0.2
All heavy atoms*	1.1 ± 0.3	0.8 ± 0.2

*RMSD were calculated only with residues 241-250 and 261-270, as the amino acids located close to the termini are more flexible.

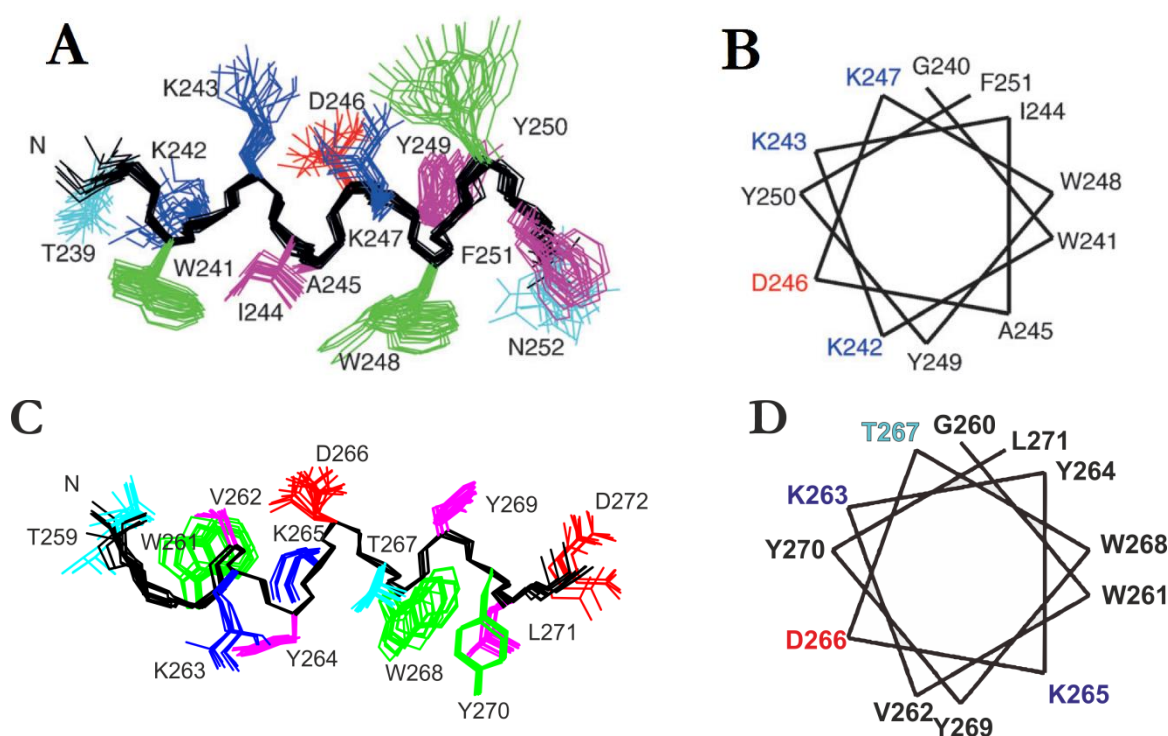


Figure 2.20. Structures of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in DPC micelles. **A** and **C**. Ensembles of the 20 lowest target function structures of $\text{LytA}_{239-252}$ (**top**) and $\text{LytA}_{259-272}$ (**bottom**), respectively, overlaid onto the backbone atoms (black). Side chains are coloured as in Figure 2.13. “N” indicates the amino end. **B** and **D**. Helical wheel representations of the side chain distribution. Positively charged residues are in blue, negatively charged in red, polar in cyan, and aromatic and hydrophobic in black.

1.4. CHARACTERIZATION OF THE INTERACTION WITH DPC MICELLES

DPC micelles are relatively spherical entities with a radius of approximately 18.6–23.3 Å, formed by 44–61 monomers per micelle. (Abel, Dupradeau, & Marchi, 2012) The length of the helices formed by $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in the presence of DPC micelles was measured from the N to C-end distances in the calculated structures by using MOLMOL (Koradi et al., 1996), resulting in approximately 22–23 Å. The peptide helices could, in principle, be lying on the micelle surface, or be immersed either totally or partially into the micelles. In the case of $\text{LytA}_{239-252}$, the former hypothesis is supported by the helical-wheel analysis shown in Figure 2.20B that clearly depicts amphipathic helix with hydrophobic and polar faces noticeably segregated. On the other hand, peptide $\text{LytA}_{259-272}$ also shows a rather segregated hydrophobic and polar faces, being the hydrophobic patch smaller, compared to $\text{LytA}_{239-252}$, because of the presence of a charged Lys (Lys265) in it (Figure 2.20D). The helix formed by $\text{LytA}_{259-272}$ can be also totally or partially immersed into the micelles, but its lower amphipathicity compared to the helix formed by $\text{LytA}_{239-252}$ may affect to the interaction.

To get experimental data allowing to discern among these possibilities, the effects of a hydro-soluble relaxation agent on the NMR signals of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in 30 mM $[\text{D}_{38}]$ -DPC, and also of two lipo-soluble relaxation agents in the case of $\text{LytA}_{239-252}$ were examined. Hydro-soluble paramagnetic compounds should affect the signals corresponding to residues lying outside the micelle, whereas lipo-soluble agents would affect those of residues buried inside the micelle. (Mäler, 2013) Upon titration of $\text{LytA}_{239-252}$ with the hydro-soluble MnCl_2 , a decrease of the intensities of the α -NH cross-peaks of residues at the N-terminal moiety was observed, but remain visible in the 2D ^1H - ^1H -TOCSY spectrum, whereas those of residues Tyr249–Asn252 at the C-terminal half disappear (Figure 2.21A). Similarly, the addition of hydro-soluble MnCl_2 to $\text{LytA}_{259-272}$ in 30 mM $[\text{D}_{38}]$ -DPC caused the broadening of signals corresponding to residues located in the C-terminus. Thus, signals from residues Leu271 and Asp272 disappear upon the addition of MnCl_2 , as seen in Figure 2.22. These results suggest that, in the two peptides, the C-terminal segment either lies outside or points outwards from the micelle.

In the case of lipo-soluble methyl-16-doxyl-stearate (free radical), which is a probe for the deepest micelle core, the α -NH cross-peaks that remain observable at the 2D ^1H - ^1H -TOCSY spectrum of $\text{LytA}_{239-252}$ are Gly240, Lys243, Ile244 (very weak), Lys247 and Asn252 (Figure 2.21C). These same α -NH cross-peaks plus those of Tyr250 and Phe251, although very weak (Figure 2.21B), persist upon titration with the lipo-soluble 5-doxyl-stearic acid (free radical). These persistent signals should correspond to residues outside the micelle or close to the surface of the micelle. Interestingly, the side chains for most of these residues are located at the same side of the α -helix (Figure 2.20B). Nevertheless, the distinction between residues inside and outside the micelle is not accurate because of the dynamic character of the peptide/micelle complex (see above), and so some signals are mostly unaffected by both hydro- and lipo-soluble compounds (Figure 2.21).

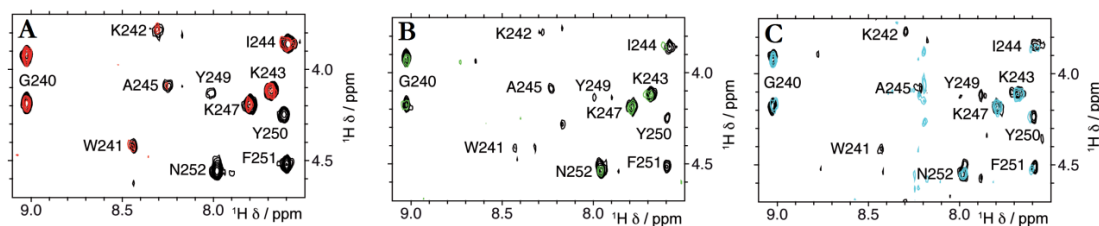


Figure 2.21. Effect of paramagnetic compounds in 2D ^1H - ^1H -TOCSY spectra of 0.5 mM $\text{LytA}_{239-252}$ in 30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v, pH 3.0 at 25 °C. **A.** Hydro-soluble MnCl_2 . Black spectrum = 0 mM relaxing agent; red spectrum = 1.4 mM relaxing agent. **B.** Lipo-soluble 5-doxyl-stearic acid (free radical). Black spectrum = 0 mM relaxing agent; green spectrum = 0.37 mM relaxing agent. **C.** Lipo-soluble methyl-16-doxyl-stearate (free radical). Black spectrum = 0 mM relaxing agent; cyan spectrum = 0.74 mM relaxing agent.

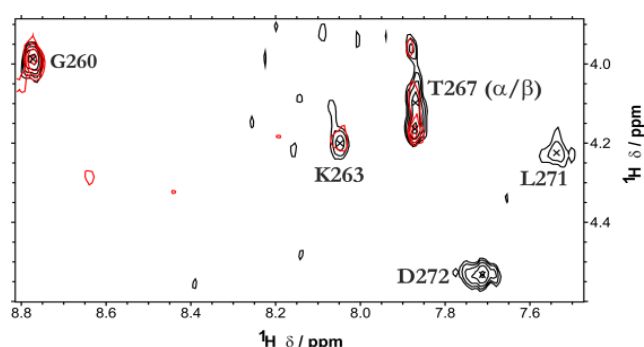


Figure 2.22. Effect of hydro-soluble paramagnetic MnCl_2 in 2D ^1H - ^1H -TOCSY spectra of 0.5 mM $\text{LytA}_{259-272}$ in 30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v, pH 3.0 at 25 °C. Black spectrum = 0 mM relaxing agent; red spectrum = 1.4 mM relaxing agent.

To better characterise how the two peptide helices interact with the micelles, additional 2D NOESY spectra of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in a mixture of 15 mM $[\text{D}_{38}]$ -DPC and 15 mM non-deuterated DPC were acquired. In these conditions, intermolecular NOE signals between peptide and micelles are observed (Figures 2.23 and 2.24). The most intense and numerous NOE signals were observed between the nuclei from the central region of the DPC aliphatic tail and the aromatic side chains from the peptides. This observation indicates that the aromatic hydrophobic side chains interact with the micelle core. In addition, some NOEs between aromatic residues with the choline methyl groups and with atoms from the polar head are observed. This could be explained by the dynamics of the interaction process, as the peptides are probably continuously moving from the solvent to the micelle and *vice versa*, establishing transient interactions with the DPC polar heads, which are solvent-exposed. Also, the fact that the aromatic residues at the C-terminal segment of the two peptides are those giving place to most intense NOEs with the choline moiety agrees with the peptide C-termini being close to the micelle surface, as indicated by the effect of relaxation agents.

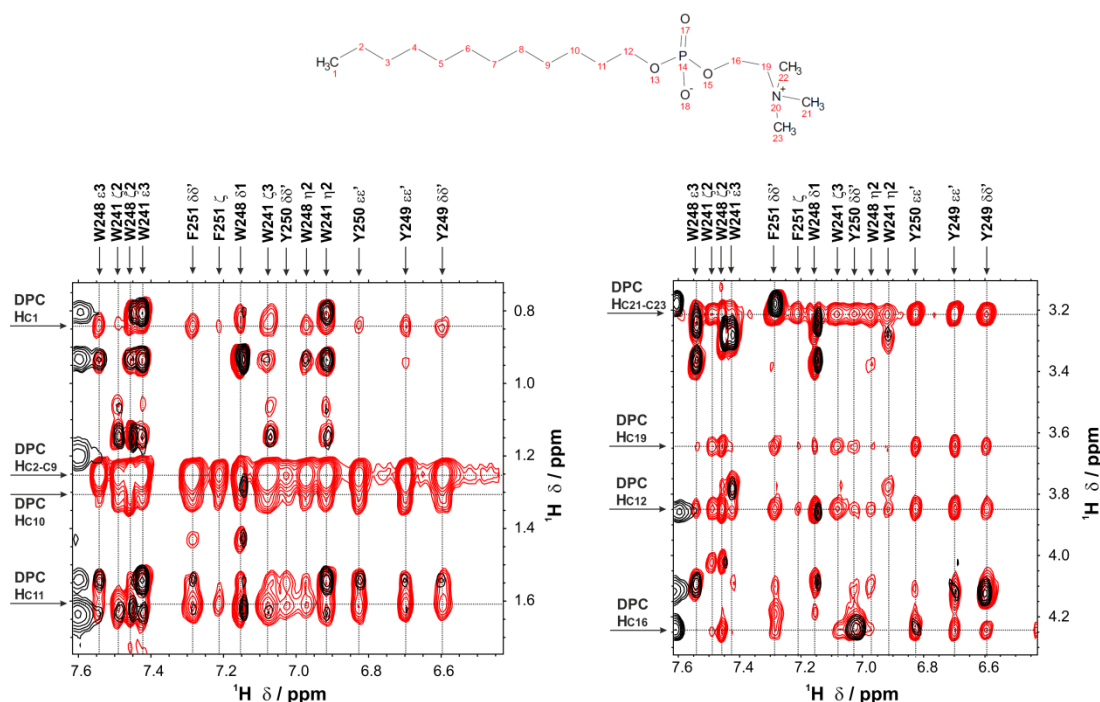


Figure 2.23. Characterization of the interaction between LytA_{239–252} and DPC micelles by NMR. The structure of a DPC molecule is shown **on top** (atoms are numbered). A superposition of two regions of the ^1H - ^1H -NOESY spectra from LytA_{259–272} samples in the presence of 30 mM [D₃₈]-DPC (black), and of 15 mM [D₃₈]-DPC and 15 mM non-deuterated DPC (red) is represented. Corresponding atoms from DPC molecule are indicated with arrows in the left vertical axis, and the correlated peptide atoms are indicated with arrows in the top horizontal axis. Measurements were made at 25 °C and pH 3.0.

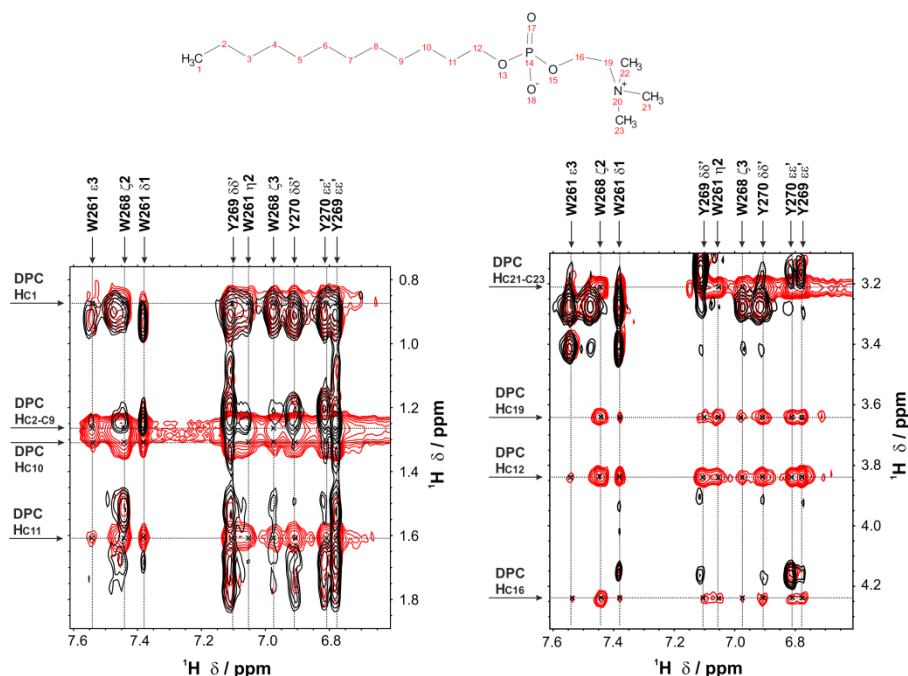


Figure 2.24. Characterization of the interaction between LytA_{259–272} and DPC micelles by NMR. The structure of a DPC molecule is shown **on top** (atoms are numbered). A superposition of two regions of the ^1H - ^1H -NOESY spectra from LytA_{259–272} samples in the presence of 30 mM [D₃₈]-DPC (black), and of 15 mM [D₃₈]-DPC and 15 mM non-deuterated DPC (red) is represented. Corresponding atoms from DPC molecule are indicated with arrows in the left vertical axis, and the correlated peptide atoms are indicated with arrows in the top horizontal axis. Measurements were made at 25 °C and pH 3.0.

In the case of $\text{LytA}_{239-252}$, to gain further information on the environment around the aromatic residues in the presence of DPC micelles, the intrinsic fluorescence spectra upon excitation at 280 nm were recorded (Figure 2.25A). In the absence of detergent, the emission spectrum of $\text{LytA}_{239-252}$ is dominated by tryptophan contributions, with a maximum at 340 nm, indicating a high solvent exposure. Addition of DPC micelles caused a blueshift in the spectrum maximum to 331 nm concomitant with an increase in fluorescence intensity. This indicates that the Trp residues are in a less polar environment and more buried from solvent in the presence of DPC micelles than in aqueous solution. Moreover, acrylamide quenching experiments in the absence and presence of detergent were carried out. Figure 2.25A shows that, for $\text{LytA}_{239-252}$, the quencher affects the Trp fluorescence to a much higher extent in aqueous solution than in the presence of DPC. Stern–Volmer analysis of the data (Figure 2.25B) yields quenching constants, K_{SV} , of $80 \pm 3 \text{ M}^{-1}$ and $10 \pm 1 \text{ M}^{-1}$ in the absence and the presence of DPC micelle, respectively.

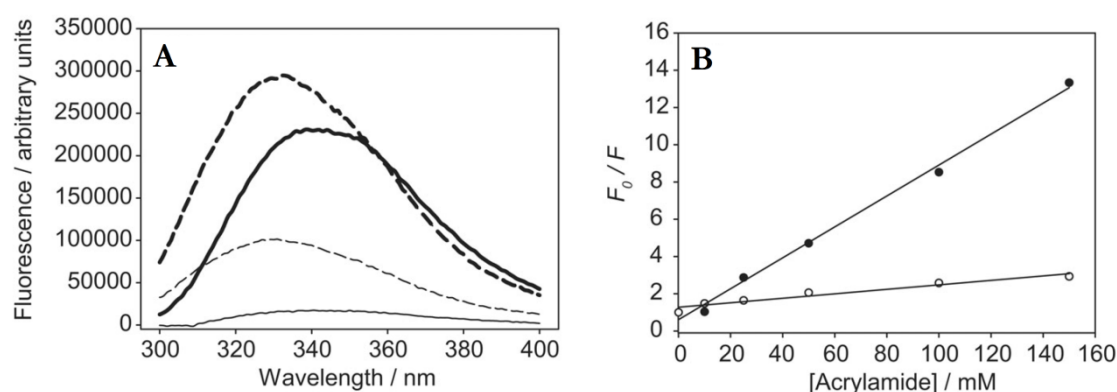


Figure 2.25. Intrinsic fluorescence of $\text{LytA}_{239-252}$. **A.** Wavelength spectra in 20 mM HCl–Gly buffer, pH 3.0 and 25 °C, upon excitation at 280 nm. Thick solid line, no addition; thin solid line, after addition of 150 mM acrylamide; thick dashed line, after addition of 30 mM DPC; thin dashed line, after addition of 30 mM DPC plus 150 mM acrylamide. **B.** Stern–Volmer plot for acrylamide quenching in the absence (●) and the presence (○) of DPC micelles (Eq. 2.4) F_0 and F represent the fluorescence intensity at 340 nm in the absence and the presence of the quencher, respectively.

These results indicate that there is a physical impediment for the quencher to reach the Trp residues when DPC micelles are formed. These differences cannot simply arise from the Trp side chains being less accessible in the helical conformation than in the β –hairpin, because the solvent–accessible areas of these residues in the two structures, calculated using MOLMOL (Koradi et al., 1996) are quite similar: 47 % for W241 and 43 % for W248 in the β –hairpin (Figure 2.13), and 46 % for W241 and 47 % for W248 in the α –helix for $\text{LytA}_{239-252}$ (Figure 2.20). Therefore, the fluorescence data suggest that the Trp side chains located in the hydrophobic face of the helix (Figure 2.20B) are immersed in the micelle. In fact, in contrast to most polar side chains, the indole rings are very ordered in the helix formed by $\text{LytA}_{239-252}$ in DPC (Figure 2.20A), indicating a rigid environment that restricts their fluctuation. Additional fluorescence experiments will be done in the future also with $\text{LytA}_{259-272}$.

On the whole, a picture that would fit both with fluorescence data, the effect of paramagnetic compounds on NMR spectra and the observed direct peptide/DPC NOEs is that the $\text{LytA}_{239-252}$ helix lies in a slightly tilted position relative to the micelle normal, probably quite close to the surface, and the N-terminus holds most interactions with the micelle (Figure 2.26). Furthermore, the hydrophobic face of the helix, which contains the Trp side chains, points towards the micelle centre, and the hydrophilic side, where Lys243 and Lys247 are placed, points to the micelle surface. For $\text{LytA}_{259-272}$, no fluorescence experiments have been performed yet, but the observed intermolecular peptide/DPC NOEs and the effect of the hydro-soluble paramagnetic MnCl_2 point out to the interaction taking place with a similar orientation to that proposed for $\text{LytA}_{239-252}$, with the C-terminus pointing outwards the micelle, and the N-terminus inserted in it (Figure 2.26).

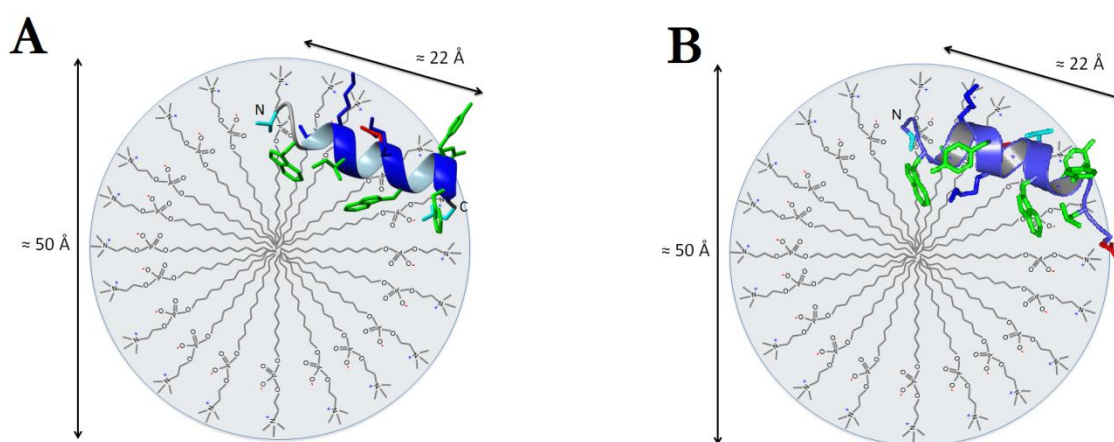


Figure 2.26. Cartoon representation of a hypothetical model for peptide/DPC micelle complexes. $\text{LytA}_{239-252}$ (A) and $\text{LytA}_{259-272}$ (B) helical peptides are displayed as blue ribbons, and side chains in sticks. Positively charged residues are in blue, negatively charged in red, Asn and Thr in cyan, and aromatic and hydrophobic in green. The amino and carboxylate termini are labelled by 'N' and 'C', respectively. The approximate lengths of the micelle diameter and peptide helices are indicated.

1.5. EFFECT OF TFE

TFE has been shown to stabilise β -hairpins (Santiveri, Pantoja-Uceda, Rico, & Jimenez, 2005), but it is most commonly known as a helix-inducer solvent. (Buck, 1998) Given that $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ spontaneously adopt native-like β -hairpin structures in aqueous solution, and α -helix structures in DPC micelles, it was intriguing the question of which of these structures the peptides would acquire in the presence of TFE. Thus, 1D and 2D NMR spectra of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ were recorded in 30 % vol. TFE and assigned their ^1H and ^{13}C resonances were assigned (see "Appendices", Tables A1 and A2).

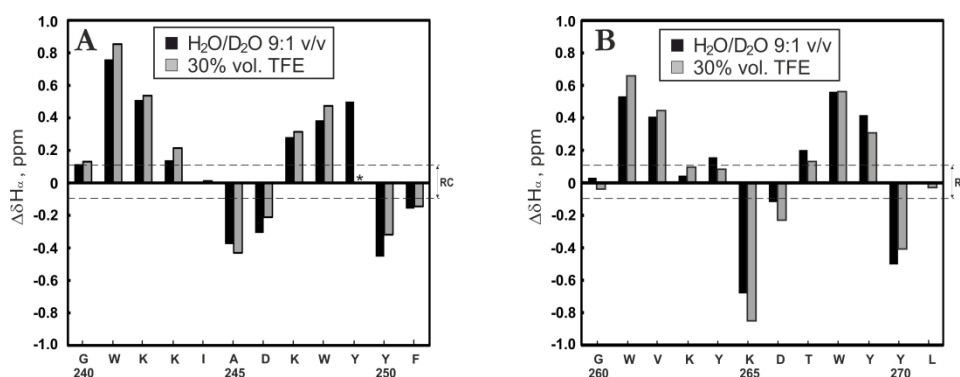


Figure 2.27. Bar plots of $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_{\alpha}^{\text{observed}} - \delta H_{\alpha}^{\text{RC}}$, ppm) as a function of sequence for LytA_{239–252} (A) and LytA_{259–272} (B) in 30 % vol. TFE (grey bars) and in aqueous solution (black bars), at pH 3.0 and 25 °C. $\delta H_{\alpha}^{\text{RC}}$ values are taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range, and the asterisk (*) indicates that the corresponding $\Delta\delta H_\alpha$ value was not determined.

The $\Delta\delta H_\alpha$ (Figure 2.27), $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ (see “Appendices”, Figures A1 and A2) values plotted as a function of sequence follow the same pattern as that in aqueous solution for both peptides, which indicates that merely inducing intramolecular H-bonding is not sufficient to convert the β -hairpins into an α -helices, and that an anisotropic environment such as that provided by detergent micelles is also necessary.

1.6. EFFECT OF NEGATIVELY-CHARGED SDS MICELLES

Given that DPC contains a positively charged choline unit, the possibility existed that the quaternary amine could emulate the role of its counterpart in the cell-wall teichoic acids and, specifically, could interact with the aromatic residues in the LytA_{239–252} and LytA_{259–272} CBR-derived peptides. To check this hypothesis, structural NMR studies of LytA_{239–252} and LytA_{259–272} in the presence of SDS were performed. Concentrations both below (0.2 mM [D₂₅]-SDS) and above (30 mM [D₂₅]-SDS) cmc, for which reported values are in the range 1–8 mM, were utilized (Mäler, 2013; Sanders & Sönnichsen, 2006). DPC and SDS have aliphatic chains of the same length (12 carbon atoms), but differ in their polar head groups: zwitterionic in DPC and negatively charged in SDS. As in the case of DPC, the NMR spectra of LytA_{239–252} and LytA_{259–272} at sub-micellar SDS concentrations are similar to those in pure aqueous solution, whereas they completely differ in the presence of SDS micelles (Figures 2.28). Similarly, the profiles of $\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ values in 30 mM SDS (Figures 2.28C and F) are very different to those observed in aqueous solution, and are very similar to those in 30 mM DPC; that is, they provide evidence that LytA_{239–252} and LytA_{259–272} in SDS micelles also adopt helical structures. The set of non-sequential NOEs confirms the formation of helix structures in SDS micelles, which have populations of 56 % and 48 % at 25 °C, respectively, as estimated from the averaged $\Delta\delta H_\alpha$ for residues 241–251 in LytA_{239–252}, and residues 261–271 in LytA_{259–272} (Vila et al., 2000). The structures in SDS micelles were calculated by following the same protocol as in aqueous solution and in DPC micelles (see “Materials and methods” section). The resulting α -helix formed by LytA_{239–252} is quite well defined (Table 2.7 and Figure 2.29), whereas helical LytA_{259–272} is less defined in first segment of the N-terminus (residues 259–263), and thus the structure is more variable there (Figure

2.29B). The helices formed by $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in both DPC and SDS micelles are quite similar, as deduced from the low RMSD values ($0.5 \pm 0.3 \text{ \AA}$ and $0.8 \pm 0.5 \text{ \AA}$ for the backbone, respectively). This suggests that the influence of the choline head group in DPC is not relevant to induce the hairpin to helix transition.

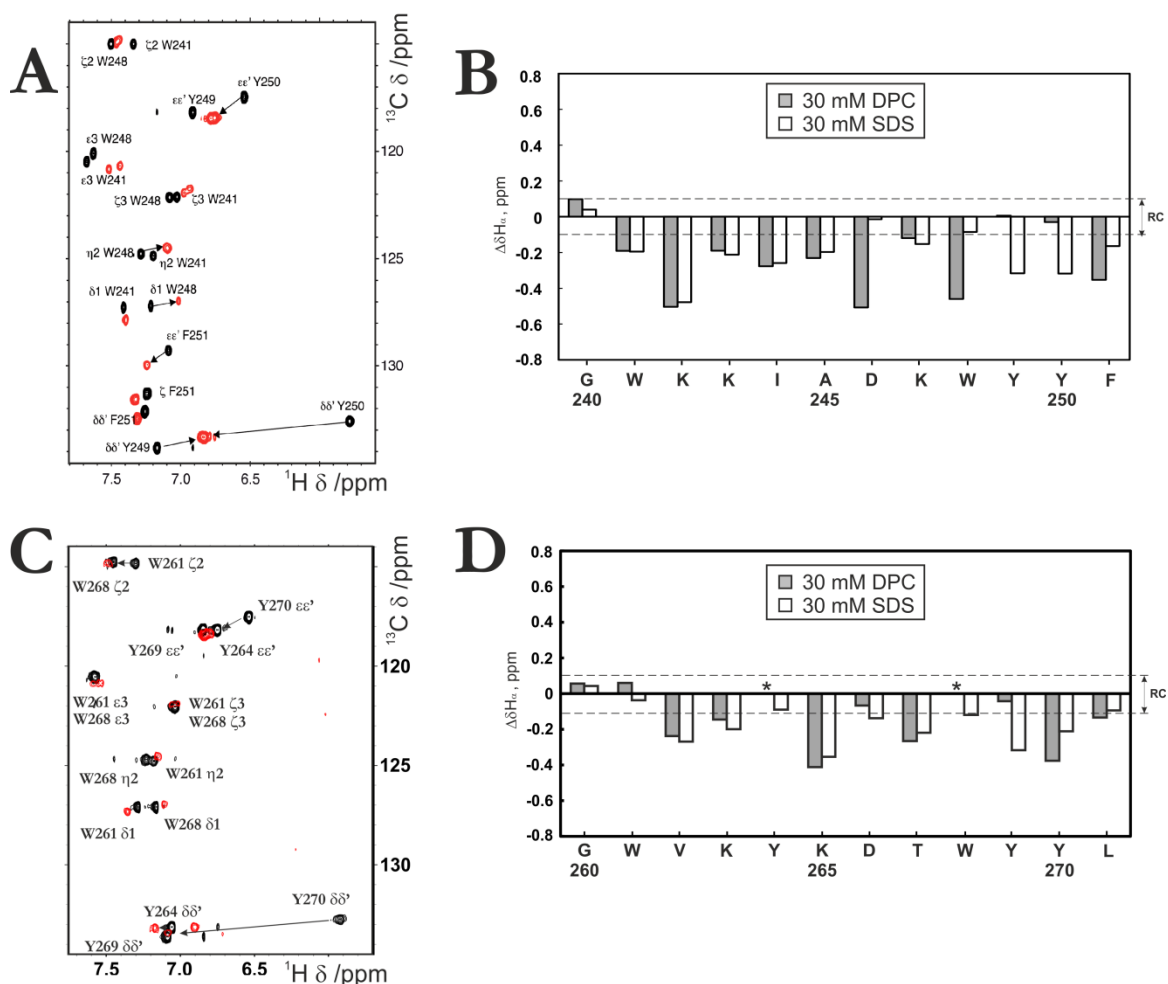


Figure 2.28. Effect of SDS on peptides $\text{LytA}_{239-252}$ (top) and $\text{LytA}_{259-272}$ (bottom). **A** and **C.** Superposition of 2D ^1H - ^{13}C -HSQC spectral regions showing the cross-peaks corresponding to the aromatic rings of Trp, Tyr, and Phe residues recorded at 25 °C and pH 3.0. Black spectra correspond to samples in D_2O , red spectrum corresponds to samples in the presence of 30 mM $[\text{D}_{25}]$ -SDS. **B** and **D.** Bar plots of $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_\alpha^{\text{observed}} - \delta H_\alpha^{\text{RC}}$, ppm) as a function of sequence for $\text{LytA}_{239-252}$ (top) and $\text{LytA}_{259-272}$ (bottom) in 30 mM $[\text{D}_{25}]$ -SDS (white bars) (25 °C) and in 30 mM $[\text{D}_{38}]$ -DPC (grey bars) (35 °C), at pH 3.0. $\delta H_\alpha^{\text{RC}}$ values are taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

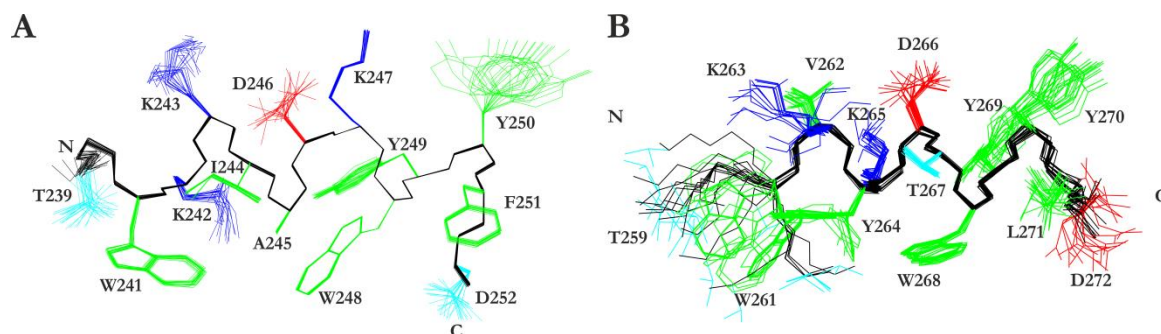


Figure 2.29. Structures of LytA₂₃₉₋₂₅₂ (A) and LytA₂₅₉₋₂₇₂ (B) in SDS micelles. Ensemble of the 20 lowest target function structures overlaid onto the backbone atoms (black). Positively charged residues are in blue, negatively charged in red, Asn and Thr in cyan, and aromatic and hydrophobic in green. The amino and carboxylate termini are labelled by 'N' and 'C', respectively.

Table 2.7. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ peptides in 30 mM [D₂₅]-SDS.

	LytA ₂₃₉₋₂₅₂	LytA ₂₅₉₋₂₇₂
Total number of restraints		
Upper limit distances	237	150
φ and ψ dihedral angles	23	20
Pairwise RMSD (Å)		
Backbone atoms*	0.06 ± 0.02	0.3 ± 0.2
All heavy atoms*	0.8 ± 0.3	1.6 ± 0.4

*RMSD were calculated only with residues 241-250 and 261-270, as the amino acids located close to the termini are more flexible.

To analyse the interaction of LytA₂₃₉₋₂₅₂ with the SDS micelles, an additional 2D NOESY spectra of the peptide in aqueous solution with a mixture of 15 mM [D₂₅]-SDS and 15 mM non-deuterated SDS was acquired, as it was done previously for the DPC micelles. Intermolecular NOE signals between peptide and micelles were detected (Figure 2.30). The only NOE signals were observed between the nuclei from the central region of the SDS aliphatic tail and the aromatic sidechains from the peptide. As it was confirmed in DPC micelles, the aromatic hydrophobic side chains interact with the micelle core.

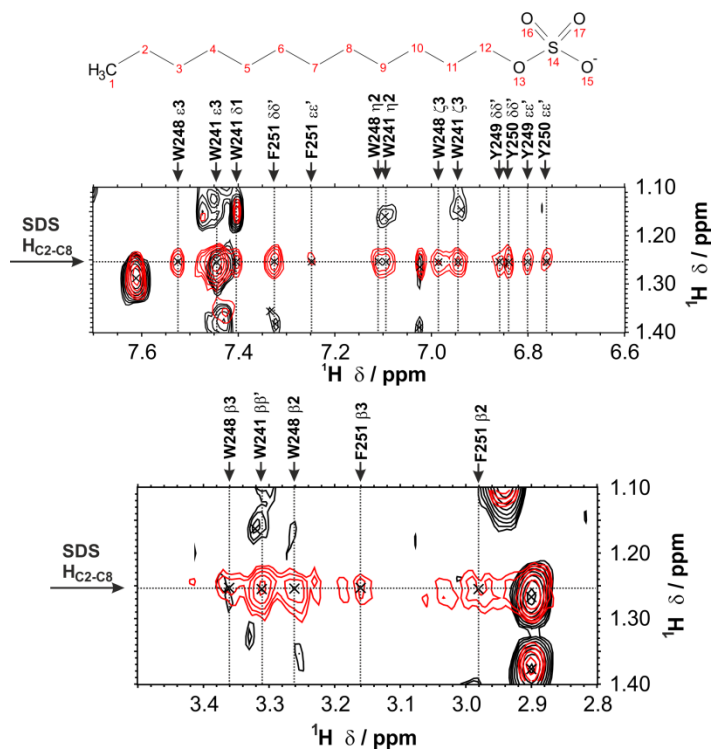


Figure 2.30. Characterization of the interaction between LytA_{239–252} and SDS micelles by NMR. The structure of a SDS molecule is shown **on top** (atoms are numbered). A superposition of two regions of the ¹H–¹H–NOESY spectra from LytA_{239–252} samples in the presence of 30 mM [D₂₅]-SDS (black), and of 15 mM [D₂₅]-SDS and 15 mM non-deuterated SDS (red) is represented. Corresponding atoms from SDS molecule are indicated with arrows, and the correlated atoms are labelled in the spectra. Measurements were made at 25 °C and pH 3.0.

1.7. BEHAVIOUR IN THE PRESENCE OF LIPID VESICLES

Many peptides have been reported to be α -helical in the presence of DPC; therefore, further experimental data concerning the importance of the choline head group of the phospholipid for helix formation by the peptides LytA_{239–252} and LytA_{259–272} were obtained. To this end, far-UV CD spectra in two types of small unilamellar vesicles (SUVs) were utilized: DMPC:DMPG (3:1) vesicles, formed by choline-phospholipids, and POPE:POPG (2:1) vesicles, formed by non-choline phospholipids. As seen in Figure 2.31, CD spectra of LytA_{239–252} in both types of vesicles exhibit a minimum at about 208 nm, a shoulder at about 222 nm, and a maximum at ~197 nm, which are characteristic of helical structures. The similarity between these spectra and that of DPC micelles (Figure 2.14) indicates that LytA_{239–252} forms the same helix structure in SUVs and in micelles. Interestingly, the fact that the change in the type of vesicles does not affect significantly to the CD spectra indicates that the choline head group of the phospholipids is not the driving force for the transition from native-like β -hairpin to amphipathic α -helix.

By contrast, LytA_{259–272} in the presence of both types of vesicles shows spectral features which are almost identical to those observed in aqueous solution (Figure 2.31, dotted line), *i.e.* positive bands at ~197 and ~227 nm, and broad negative bands from ~200–215 nm. These results point to the fact that LytA_{259–272} maintains its β -hairpin structure in the presence of SUVs, regardless of the presence of choline-phospholipids. Contrary to LytA_{239–252}, LytA_{259–272} is not capable of adopting a helical structure in the presence of SUVs.

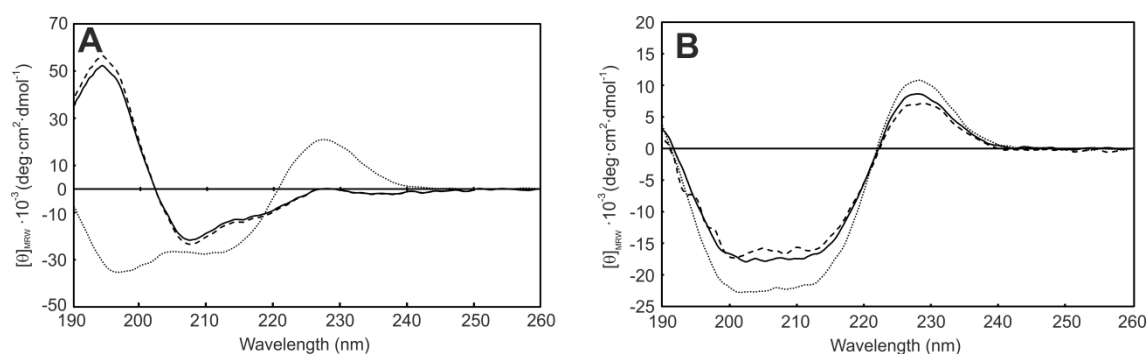


Figure 2.31. Far-UV CD spectra of LytA₂₃₉₋₂₅₂ (A) and LytA₂₅₉₋₂₇₂ (B) in vesicles: DMPC:DMPG (3:1) (solid line) and POPE:POPG (2:1) (dashed line). All samples were prepared in 10 mM phosphate buffer (pH 7.0) and a peptide/lipid ratio 1:50, and were measured at 30 °C. For a better comparison, CD spectra in aqueous solution (10 mM phosphate buffer, pH 7.0, 30 °C) are shown for both peptides (dotted lines).

Strikingly, the analogous behaviour of LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂, observed in all the other solvent conditions, is not reproduced in the presence of SUVs, as the latter keeps its original β -hairpin structure in the presence of vesicles (Figure 2.31B). There is no satisfactory explanation for this observation yet.

2. STRUCTURAL CHARACTERIZATION OF LYTA₂₃₉₋₂₇₂

To go in depth into the understanding of the behaviour of peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂, a new peptide was designed containing the sequences of both of them separated by the linker present in the native protein. This new 34-residue peptide, LytA₂₃₉₋₂₇₂, is interesting for two motives. On one hand, it serves to check if the individual behaviour of LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ is reproduced in this new, longer peptide. On the other hand, as it encompasses most of the essential elements constituting the theoretical minimal choline-binding unit, it is a good system to study their ability to bind choline. (Galán-Bartual et al., 2015)

2.1. CD STUDY IN AQUEOUS SOLUTION AND INTERACTION WITH CHOLINE

Firstly, the behaviour of LytA₂₃₉₋₂₇₂ in aqueous solution and its interaction with choline were studied by CD. Far and near-UV CD spectra acquired for LytA₂₃₉₋₂₇₂ in aqueous solution at pH 3.0 (20 mM HCl-glycine buffer) showed equivalent features than those from isolated LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ peptides. This is, a strong positive band at 227 nm in the far-UV CD spectrum (Figure 2.32A, solid line), together with the positive band at ~265 nm and negative minima at ~285–295 nm in the near-UV CD spectrum (Figure 2.32B, solid line), arising from aromatic rings in rigid conformations. These results indicate a remarkable structural similarity between LytA₂₃₉₋₂₇₂ and the isolated peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂, suggesting that the regions corresponding to the sequence of both peptides are adopting their native-like β -hairpin conformation observed previously.

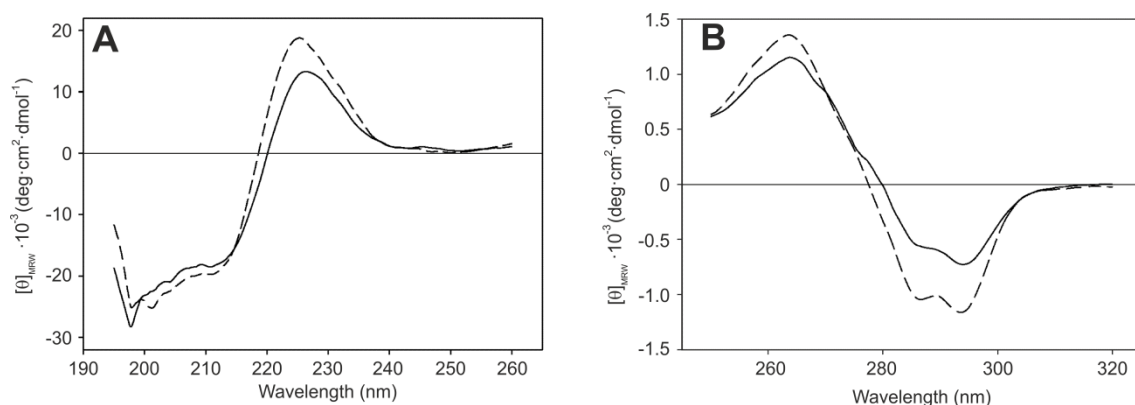


Figure 2.32. CD spectra acquired for a 27.5 μM $\text{LytA}_{239-272}$ sample. Far-UV (**A**) and near-UV (**B**) regions. Solid lines are from samples in aqueous solution (20 mM HCl-glycine buffer, pH 3.0), and dashed lines represent samples in the same solvent plus 500 mM choline. All spectra were recorded at 5 $^{\circ}\text{C}$, and mean residue ellipticities are shown.

In the near-UV region of the CD spectra, $\text{LytA}_{239-272}$ behaves like the shorter peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, since both display substantial changes upon addition of choline (500 mM). The increase of intensity is higher than those observed for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, pointing to a larger number of bound choline molecules (maybe two, instead of one), and/or a greater ligand-induced conformational change. The magnitude of the changes observed in the far-UV CD spectrum is smaller than those observed for $\text{LytA}_{259-272}$, but higher than those for $\text{LytA}_{239-252}$, for which no changes occurred upon addition of choline. It is reasonable to observe an averaged behaviour in this context, as the long peptide contains the two short peptides in its sequence.

Thermal stability of $\text{LytA}_{239-272}$ was examined by monitoring the temperature dependence of the molar ellipticity at 225 nm in the range 5–90 $^{\circ}\text{C}$, both in the absence and in the presence of 500 mM choline (Figure 2.33), as previously done for both $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. For this peptide, heating also led to featureless CD spectra, indicating a loss of structure, as observed for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. Data were fitted to the Gibbs–Helmholtz equation (Eq. 2.1), with the same considerations commented before for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. The thermodynamic parameters calculated for $\text{LytA}_{239-272}$, in the absence and the presence of 500 mM choline (in HCl-glycine buffer at pH 3.0) are shown in Table 2.8.

Table 2.8. Thermodynamic parameters calculated for $\text{LytA}_{239-272}$ in HCl-Gly buffer at pH 3.0.

	$\text{LytA}_{239-272}$	
	0 mM choline	500 mM choline
ΔH_m ($\text{kJ}\cdot\text{mol}^{-1}$)	45 ± 4	62 ± 3
T_m (K)	311 ± 1 (38 $^{\circ}\text{C}$)	320 ± 1 (47 $^{\circ}\text{C}$)
ΔG (25 $^{\circ}\text{C}$) ($\text{kJ}\cdot\text{mol}^{-1}$)	1.9 ± 0.2	4.3 ± 0.4

As seen for both $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, the two unfolding traces converge at $\sim 70^\circ\text{C}$ (Figure 2.33), indicating that this peptide containing the sequences of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ is also competent to bind choline up to these high temperatures. Stability of $\text{LytA}_{239-272}$ in the absence of choline shows a value intermediate to those from $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. In the presence of choline, stability is close to the values observed for $\text{LytA}_{259-272}$ (see Tables 2.3 and 2.8).

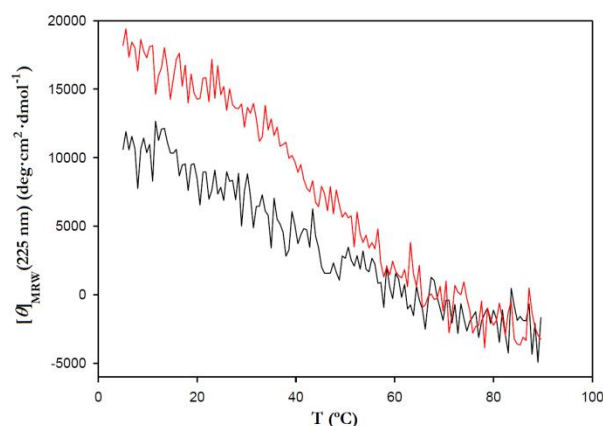


Figure 2.33. Thermal unfolding of peptide $\text{LytA}_{239-272}$ monitored by CD, in the presence (black line) and absence (red line) of 500 mM choline.

As done in the case of peptide $\text{LytA}_{259-272}$, the affinity of $\text{LytA}_{239-272}$ for choline was calculated by recording fluorescence spectra at 25°C at different ligand concentrations. The change in fluorescence intensity upon addition of choline can be observed in Figure 2.34A. To check whether the effect is due to unspecific electrostatic interactions (choline has a net positive charge), a control experiment with NaCl was performed. The different effect produced by the addition of 1 M choline (higher intensity and blueshift) and 1 M NaCl (lower intensity) demonstrates the specificity of the interaction with choline (Figure 2.34B).

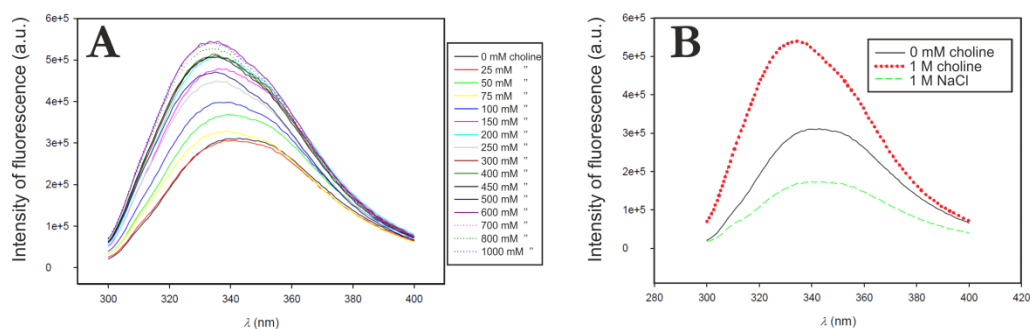


Figure 2.34. Affinity of $\text{LytA}_{239-272}$ for choline by fluorescence spectroscopy. A) Titration of a $\text{LytA}_{239-272}$ sample with choline (0–1000 mM), in HCl – Gly buffer, at pH 3.0 and 25°C . B) Experiment to determine the specificity of the interaction. $\text{LytA}_{239-272}$ sample with 1000 mM choline and $\text{LytA}_{239-272}$ sample with 1000 mM NaCl were compared.

The dissociation constant was calculated fitting the representation of the inverse of $\Delta F_{330/350}$ against choline concentration to Eq. 2.5 (Figure 2.35). The value obtained was 184 ± 15 mM, showing that the affinity of $\text{LytA}_{239-272}$ for choline is lower than that of $\text{LytA}_{239-252}$ (80 ± 10 mM), and similar to $\text{LytA}_{259-272}$ (294 ± 28 mM). Considering the binding of one molecule by $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, an additional binding site is not detected in $\text{LytA}_{239-272}$.

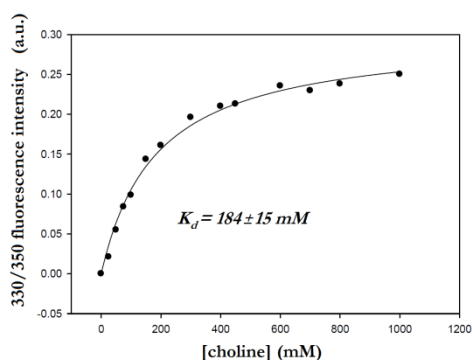


Figure 2.35. Affinity of LytA_{239–272} for choline by fluorescence spectroscopy. Fitting of the 330/350 fluorescence intensity increase to Eq. 2.5 to obtain the dissociation constant (K_d).

2.2. NMR STRUCTURE CALCULATION IN AQUEOUS SOLUTION

The structure adopted by peptide LytA_{239–272} in aqueous solution was determined by NMR. 1D and 2D NMR spectra were recorded using a ~1 mM sample. Following a standard strategy, ¹H and ¹³C chemical shifts were assigned. In the regions corresponding to the sequences of LytA_{239–252} and LytA_{259–272}, non-sequential NOE cross-peaks characteristic of antiparallel β -sheets were observed in 2D ¹H–¹H-NOESY spectra (Figure 2.36A and 2.36B); that is, those between the H _{α} protons of residues facing each other in non-hydrogen-bonded sites, and between amide protons of residues facing each other in hydrogen-bonded sites. The presence of these NOEs shows that peptide LytA_{239–272} maintains the β -hairpin structures observed in the isolated peptides LytA_{239–252} and LytA_{259–272}, connected by a mostly unstructured linker. Formation of β -hairpin structures was confirmed by the plot of $\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ as a function of peptide sequences; that is, two stretches of positive $\Delta\delta H_\alpha$ and $\Delta\delta C_\beta$ values, and negative $\Delta\delta C_\alpha$ separated by $\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ values of the corresponding opposite sign at the turn regions, plus a random coil stretch located in the linker segment (Fesinmeyer et al., 2005; Santiveri et al., 2002) (Figure 2.36C and Figure A3 in “Appendices”).

Based on the averaged $\Delta\delta H_\alpha$ values at the strand residues in both β -hairpin regions (+0.19 ppm and +0.32 ppm for 239–252 and 259–272 segments, respectively, at 35 °C) and considering that the averaged $\Delta\delta H_\alpha$ value at protein β -strands is +0.40 ppm, (Wishart et al., 1991) the β -hairpin populations formed in aqueous solution at pH 3.0 and 35 °C are approximately 48 % and 81 % for LytA_{239–252} and LytA_{259–272} regions, respectively. Surprisingly, within LytA_{239–272}, the β -hairpin formed by the segment corresponding to LytA_{239–252} peptide is less populated than when isolated (48 % vs 100 %, respectively), pointing to the fact that it is maybe less stabilized when forming part of the longer peptide. In the case of the population of the β -hairpin formed by the sequence of LytA_{259–272}, it is almost identical when forming part of LytA_{259–272} and isolated (76 % vs 81 %), showing that stability of this β -hairpin is almost unaffected by the rest of the peptide.

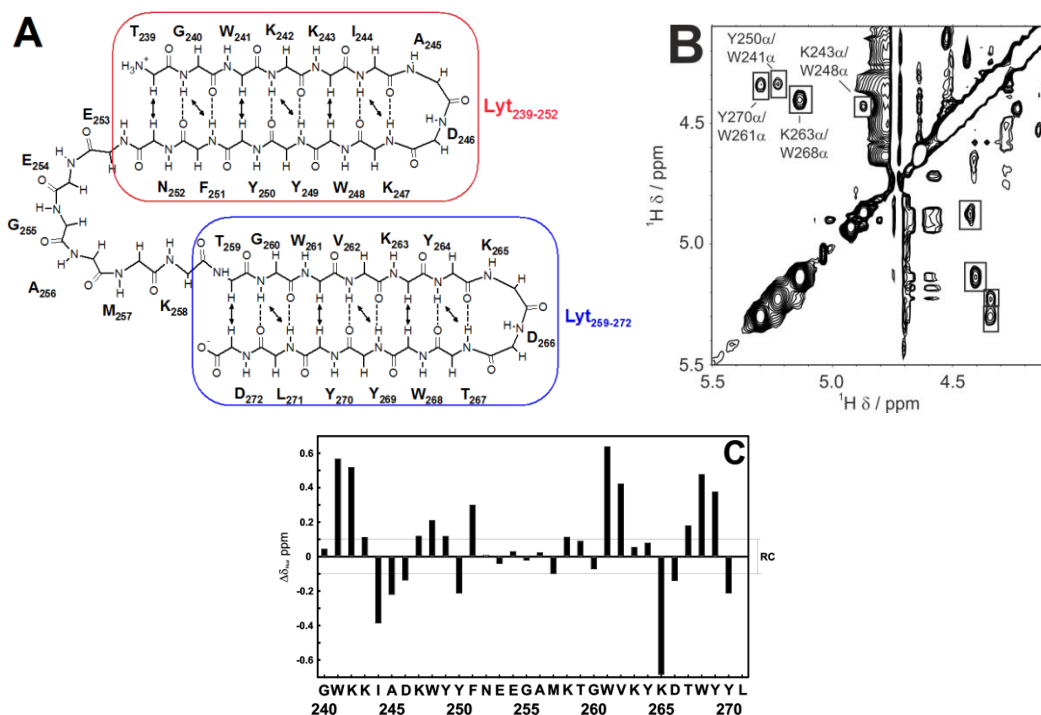


Figure 2.36. NMR data for LytA_{239–272}. **A.** Schematic representation of the two 2:2 β -hairpin formed by LytA_{239–272} in aqueous solution. Double arrows indicate the NOEs observed in 2D NOESY spectra. Fragments corresponding to LytA_{239–252} (red box) and LytA_{259–272} (blue box) are indicated. **B.** 2D NOESY (H_α region) spectra of LytA_{239–272} in D₂O at pH 3.0 and 35 °C. Non-sequential NOEs are boxed and labelled at one of the diagonal sides. **C.** Bar plot of $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_{\alpha, \text{observed}} - \delta H_{\alpha, \text{RC}}$) as a function of sequence for peptide LytA_{239–272} in D₂O at pH 3.0 and 35 °C. $\delta H_{\alpha, \text{RC}}$ values were taken from (Wishart et al., 1995) The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

Structure calculations were performed on the basis of the distance restraints derived from the complete sets of observed NOEs and the dihedral angle restraints obtained from the $^1H_\alpha$, $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts by using the program TALOS+ by following the same protocol used for the shorter peptides (see pages 62–63) (Table 2.9, and Table A.17 of the Appendices). (Shen et al., 2009)

Table 2.9. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for LytA_{239–272} peptide in aqueous solution.

	LytA _{239–272}
Total number of restraints	
Upper limit distances	278
φ and ψ dihedral angles	64
Pairwise RMSD (Å)	
Backbone atoms	
Residues 239–252*	0.3 ± 0.1
Residues 259–272*	0.6 ± 0.4
All heavy atoms	
Residues 239–252*	1.5 ± 0.4
Residues 259–272*	1.5 ± 0.5

*RMSD were calculated only with residues 241–250 and 261–270, as the amino acids located close to the termini are more flexible.

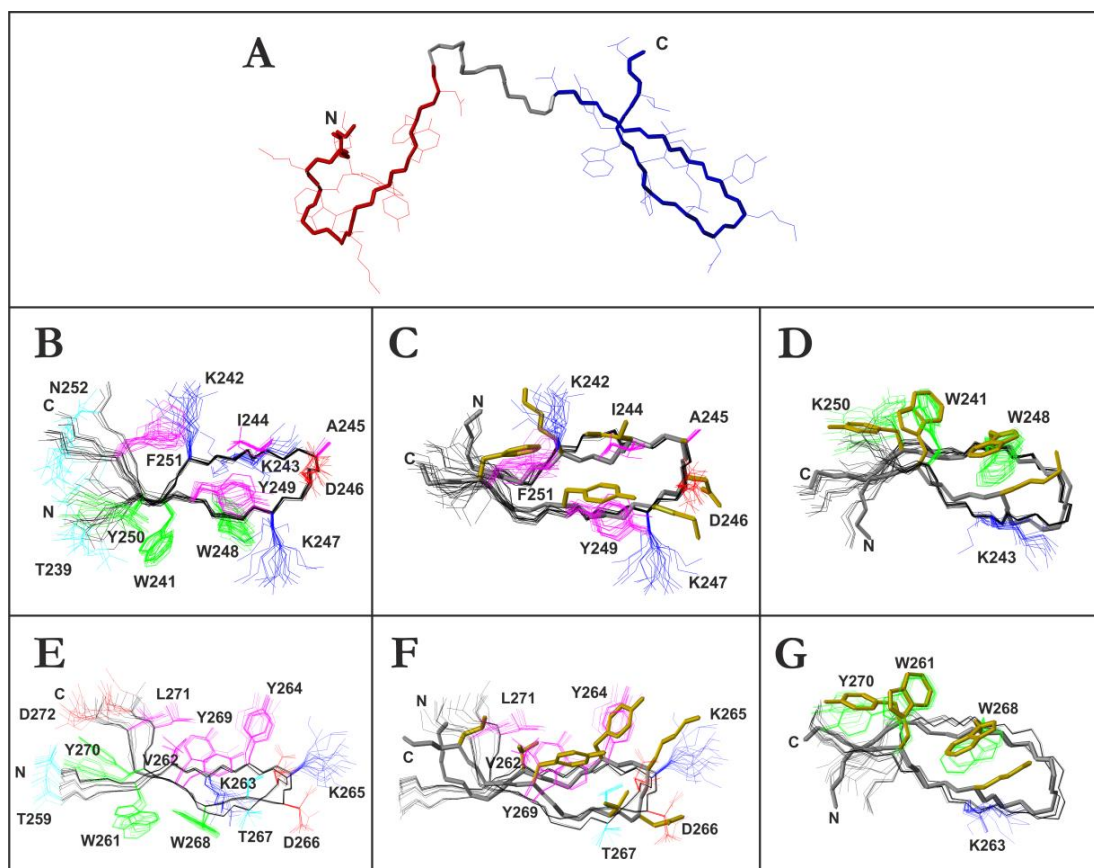


Figure 2.37. Structure of LytA₂₃₉₋₂₇₂ in aqueous solution: **A.** A representative conformer from the 20 calculated structures for the peptide (colour code: LytA₂₃₉₋₂₅₂ region in red, LytA₂₅₉₋₂₇₂ region in blue, linker segment in grey). The ensemble is not represented, as the relative position of the two β -hairpins is not fixed. **B** and **E.** Ensemble of the 20 lowest target function structures for LytA₂₃₉₋₂₅₂ region and LytA₂₅₉₋₂₇₂ region (respectively) overlaid onto the backbone atoms (black). Side chains colour code: blue for positively charged; red for negatively charged; cyan for polar; green and magenta for aromatic and hydrophobic side chains pointing downwards and upwards, respectively. **C, D, F,** and **G** show backbone atoms (in black) overlaid onto the corresponding atoms in the crystalline C-LytA structure (PDB code: 1GVM, in thick grey) for LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ regions. Side chains of residues at the turn region and at hydrogen-bonded sites are shown in **C** and **F**, and those of residues at non-hydrogen-bonded sites in **D** and **G**. Side chains are coloured following the same colour code mentioned before, and those of C-LytA in gold. The amino and carboxylate termini are labelled by "N" and "C", respectively.

In the calculated structures, the regions corresponding to the β -hairpins are well defined (Figure 2.37A), as indicated by the small pairwise RMSDs presented by the backbone atoms (Table 2.9). They are also very similar to the native structures, as it can be appreciated in Figure 2.37 (backbone RMSD: 1.33 ± 0.01 and 1.18 ± 0.08 , for LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ regions, respectively). However, the relative position of the two β -hairpins is not fixed, as there are no restrictions for their movement around the linker segment.

2.3. EFFECT OF DPC AND SDS MICELLES

In order to check if the β -hairpin to α -helix transition occurs in the context of the long LytA₂₃₉₋₂₇₂, samples of this peptide in the presence of DPC and SDS micelles were

examined. First, far-UV CD spectra were acquired and almost the same results were observed for DPC and SDS micelle samples. They showed a drastic change in the curve shape, similar to those observed for isolated $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. The characteristic maximum at 227 nm seen for β -hairpin structures in aqueous solution (Figure 2.38, solid line) is lost and helical features appear in the spectra in the presence of micelles, such as the minimum at 208 nm, the shoulder at 222 nm, and the positive band at ~ 197 nm (Figure 2.38, dashed and dotted lines). These observations point to the fact that the β -hairpin to α -helix transition triggered by the presence of detergent micelles also takes place in the long $\text{LytA}_{239-272}$ peptide.

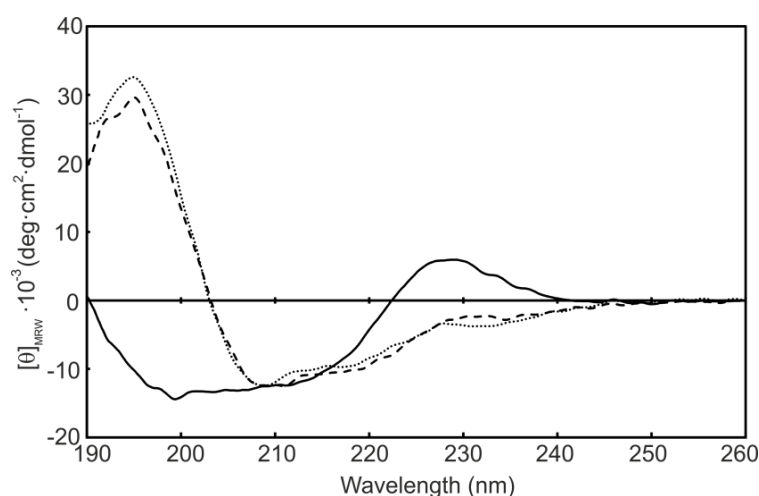


Figure 2.38. Far-UV CD spectra acquired for 27.5 μM $\text{LytA}_{239-272}$ samples. Solid line is from a sample in aqueous solution (20 mM HCl-glycine buffer, pH 3.0), dashed line represents a sample in the same solvent plus 30 mM DPC, and dotted line is from a sample in the same solvent plus 30 mM SDS. The spectra were recorded at 30 $^{\circ}\text{C}$, and mean residue ellipticities are shown.

NMR characterization of $\text{LytA}_{239-272}$ in the presence of DPC and SDS micelles was next carried out. As occurs in the case of peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ in 30 mM DPC and in 30 mM SDS, spectra of $\text{LytA}_{239-272}$ look very different in these conditions compared to those obtained in aqueous solution, as observed in the 2D ^1H - ^{13}C -HSQC spectra shown in Figure 2.39A and B. Conformational shift plots in the presence of DPC and SDS micelles are completely different from those in aqueous solution (Figure 2.39C), and compatible with a β -hairpin to α -helix transition. The $\Delta\delta\text{H}_{\alpha}$ profile observed in aqueous solution showed two β -hairpin regions (characterised by two positive stretches from the β -strands separated by a small stretch of negative values, corresponding to the turn), and a poorly structured (low $\Delta\delta\text{H}_{\alpha}$ values) region between them (corresponding to the linker). In 30 mM DPC and in 30 mM SDS, the β -hairpin regions change to negative values, representative of helical structures. The linker segment also shows negative values, but with a smaller magnitude, pointing to the fact that the helical structure may be extended through this linker. The transition is also evident in the $\Delta\delta\text{C}_{\alpha}$ and $\Delta\delta\text{C}_{\beta}$ plots (see “Appendices”, Figure A3).

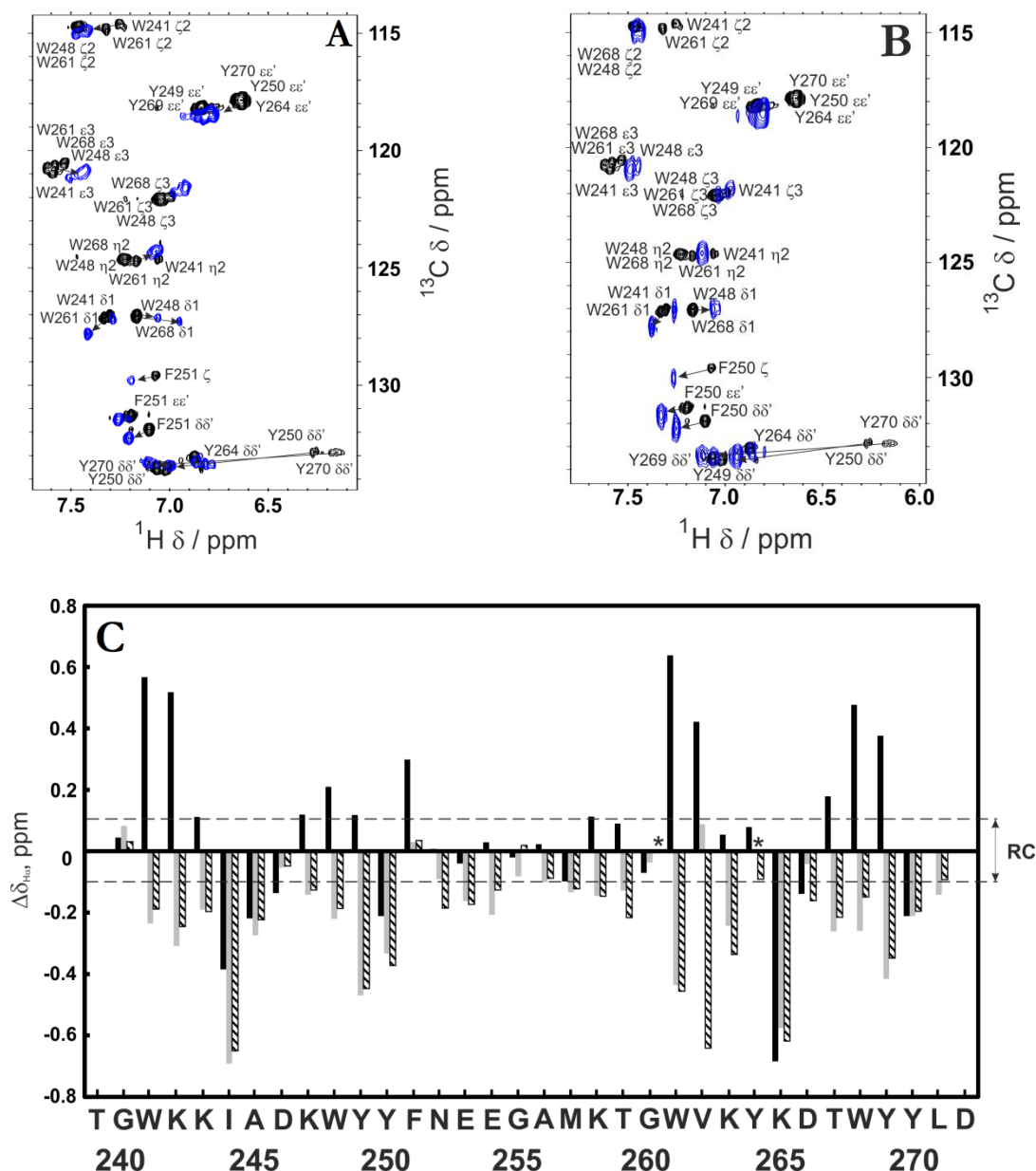


Figure 2.39. Effect of DPC and SDS above cmc on LytA₂₃₉₋₂₇₂ peptide monitored by NMR. Superposition of two 2D ^1H - ^{13}C -HSQC spectral regions showing the cross-peaks corresponding to the aromatic rings of Trp, Tyr, and Phe residues recorded at 35 °C and pH 3.0 for LytA₂₃₉₋₂₇₂ in 30 mM [D₃₈]-DPC (blue spectra) (A) and in 30 mM [D₂₅]-SDS (blue spectra) (B). Black spectra correspond to samples in D₂O. C. Plots of $\Delta\delta_{\text{H}\alpha}$ ($\Delta\delta_{\text{H}\alpha} = \delta_{\text{H}\alpha}^{\text{observed}} - \delta_{\text{H}\alpha}^{\text{RC}}$, ppm) as a function of LytA₂₃₉₋₂₇₂ peptide sequence in H₂O/D₂O 9:1 v/v (black bars), 30 mM DPC (>cmc) (grey bars), and 30 mM SDS (>cmc) (stripped bars). The N- and C-terminal residues are not shown. $\delta_{\text{H}\alpha}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The asterisk means that the corresponding $\delta_{\text{H}\alpha}$ values were not determined. Measured at 35 °C and pH 3.0. Dashed lines indicate the random coil range (RC).

Helix populations of the regions corresponding to peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ within peptide LytA₂₃₉₋₂₇₂ were estimated as explained before. For residues 239-252 (LytA₂₃₉₋₂₅₂ segment) helix population was estimated in 70 %, and for residues 259-272 (LytA₂₅₉₋₂₇₂ segment) helix population was estimated in 67 %. Comparing with the isolated peptides, an

increase in helix populations is observed (70 % vs 62 % for LytA_{239–252}; 67 % vs 47 % for LytA_{259–272}).

The formation of a helical structure in DPC and in SDS micelles was confirmed by the observation of a set of characteristic NOEs, including $\alpha\text{N}(i,i+3)$, $\alpha\beta(i,i+3)$ and strong sequential $\text{NN}(i,i+1)$ in both conditions (Figure 2.40A and B). NMR observations together with CD results prove that peptide LytA_{239–252} reproduces the β –hairpin to α –helix transition induced by DPC and SDS micelles observed in the isolated peptides LytA_{239–252} and LytA_{259–272}.

272*

To check that the conformational change occurs in the presence of DPC or SDS micelles and not by interaction with the monomers, 2D NMR spectra at a peptide/detergent ratio of around 1:1, but at sub-micellar DPC and SDS concentrations (0.6 mM detergent, and ~ 0.5 mM LytA_{239–272}) were acquired. The obtained spectra were essentially identical to those in aqueous solution (not shown). As observed for peptides LytA_{239–252} and LytA_{259–272}, reversibility of the conformational transition was proved also for LytA_{239–272}.

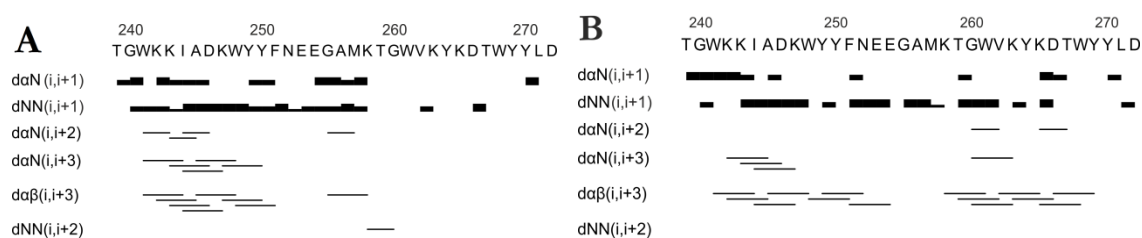


Figure 2.40. NOE summary for LytA_{239–272} in 30 mM DPC (**A**), and in 30 mM SDS (**B**) (thickness of the lines is proportional to NOE signal intensities).

All data point to the formation of an analogous helical structure both in the presence of DPC and SDS micelles. For this reason, NMR structure calculation was executed on LytA_{239–272} in the presence of 30 mM SDS micelles to get a representative structure, since the quality of the spectra was better in SDS micelles than in DPC micelles. The resulting structures are represented in Figure 2.41. As suggested by the conformational shifts (Figure 2.39C), a helical structure extended through the peptide full-length was obtained. Calculating the RMSD values for the regions corresponding to the sequences of LytA_{239–252} and LytA_{259–272} independently, a good fitting is verified, whereas the RMSD values for the full-length are significantly worse (Table 2.10). This indicates that the linker segment is worse defined than the other two regions, which is in accordance with the observation of negative $\Delta\delta H_x$ values lower than those from the rest of the peptide (Figure 2.39, stripped bars). This fact may favour the curvature observed in the full-length peptide.

Table 2.10. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for LytA_{239–272} peptide in SDS micelles.

	LytA _{239–272} in SDS micelles
Total number of restraints	
Upper limit distances	390
φ and ψ dihedral angles	61
Pairwise RMSD (Å)	
Backbone atoms	
Full-length (241–270)*	3 ± 1
Residues 239–252*	0.5 ± 0.3
Residues 259–272*	0.7 ± 0.3
All heavy atoms	
Full-length (241–270)*	4 ± 2
Residues 239–252*	1.4 ± 0.5
Residues 259–272*	1.5 ± 0.4

*RMSD were calculated only with residues 241–250 and 261–270, as the amino acids located close to the termini are more flexible. In the full-length, terminal residues are also omitted for RMSD calculation.

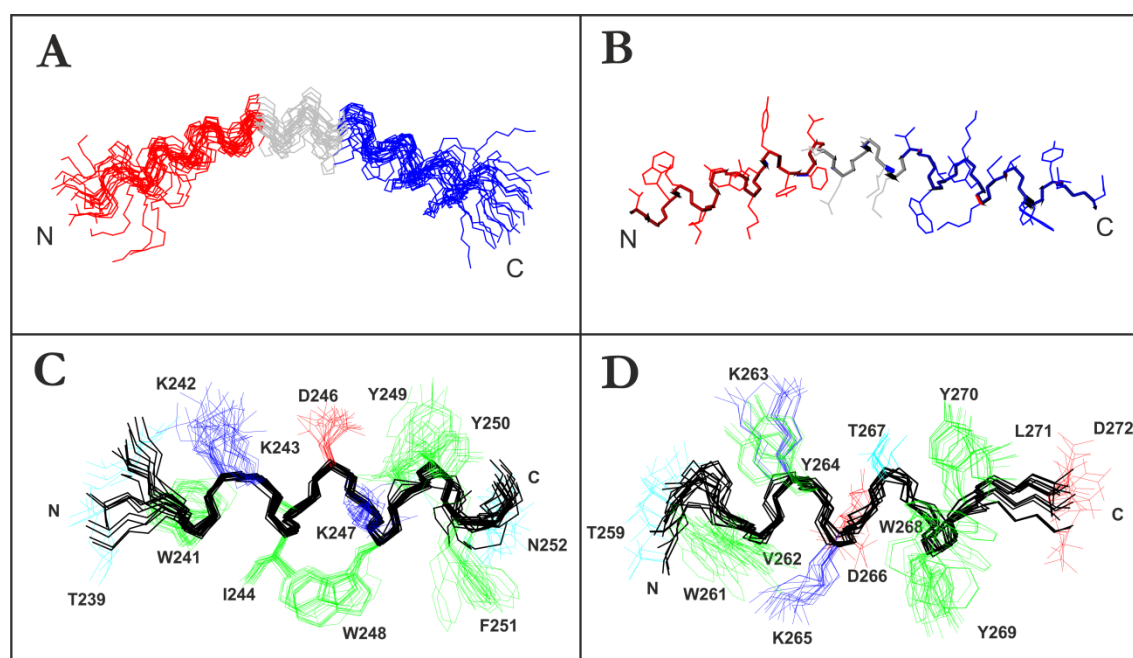


Figure 2.41. Structure of LytA_{239–272} in SDS micelles: **A.** Ensemble of the 20 lowest target function structures for LytA_{239–272} peptide (colour code: LytA_{239–252} region in red, LytA_{259–272} region in blue, linker segment in grey). **B.** A representative conformer from the 20 calculated structures for the LytA_{239–272} peptide is represented with the side chains for better visualization. **C** and **D.** Ensembles of the 20 lowest target function structures for LytA_{239–252} region and LytA_{259–272} region (respectively). Side chains colour code: blue for positively charged; red for negatively charged; cyan for polar; green for aromatic and hydrophobic side chains. The amino and carboxylate termini are labelled by “N” and “C”, respectively.

2.4. CHARACTERIZATION OF THE INTERACTION WITH DPC MICELLES

The effect of a hydro-soluble relaxation agent on the NMR signals of $\text{LytA}_{239-272}$ in 30 mM $[\text{D}_{38}]$ -DPC was analysed. Upon titration of $\text{LytA}_{239-272}$ with the hydro-soluble MnCl_2 , most α -NH cross-peaks are still present. Many of the α -NH cross-peaks from residues in the 260 – 272 were not visible even in the absence of the relaxing agents. However, the α -NH cross-peak from L271 clearly disappears, as well as the β -NH cross-peaks from D272 (Figure 2.42). These observations demonstrate that MnCl_2 affects the C-terminus of the helical $\text{LytA}_{239-272}$, suggesting that this region either lies outside or points outwards from the micelle. This orientation is the same observed for the isolated peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$.

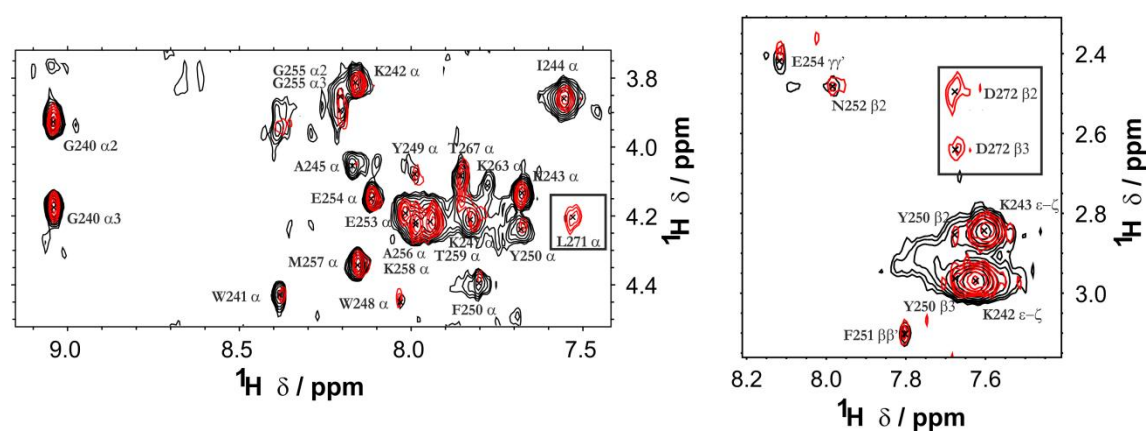


Figure 2.42. Effect of hydro-soluble paramagnetic MnCl_2 in 2D ^1H - ^1H -TOCSY spectra of 0.5 mM $\text{LytA}_{239-272}$ in 30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v, pH 3.0 at 35 °C. Red spectrum = 0 mM relaxing agent; black spectrum = 2.1 mM relaxing agent. Boxed signals disappear upon addition of MnCl_2 .

2.5. EFFECT OF TFE

The effect of TFE was also examined on peptide $\text{LytA}_{239-272}$ by CD and by NMR. The far-UV CD spectrum in the presence of 30 % vol. TFE shows a loss of intensity of the 227 nm band observed in aqueous solution (Figure 2.43), pointing to a decrease of the rigidity of the aromatic side chains. The broad negative band present in the aqueous sample became more similar to the negative band profile characteristic of helical structures, and a positive band appears around 195 nm. All these features taken together indicate that the aqueous structure composed of two well-defined β -hairpins is distorted in the presence of TFE, and the packing of the side chains in the hairpins is looser in this condition.

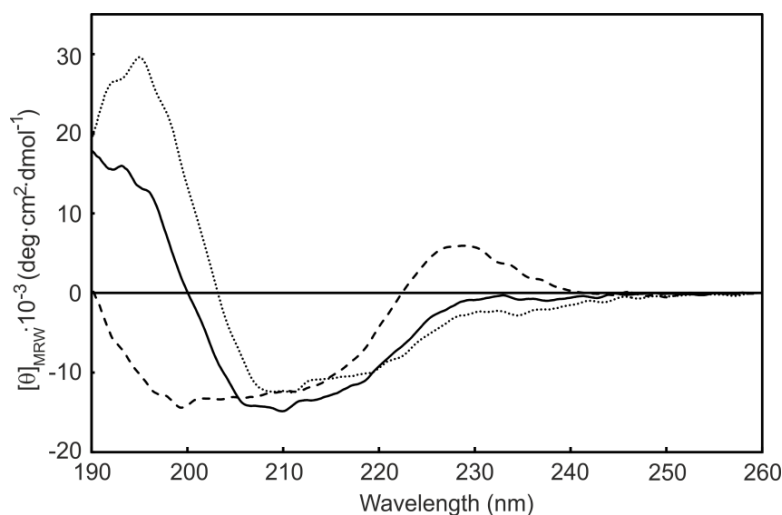


Figure 2.43. Far-UV CD spectra of LytA₂₃₉₋₂₇₂ in the presence of 30 % vol. TFE. The sample was prepared in HCl-Gly buffer (pH 3.0) and was measured at 30 °C. For a better comparison, CD spectra in aqueous solution (HCl-Gly buffer, pH 3.0) and in the presence of 30 mM DPC are also shown (dashed lined and dotted line, respectively).

To get more information about the secondary structure of LytA₂₃₉₋₂₇₂ in the presence of TFE, 1D and 2D NMR spectra were recorded, and their ¹H and ¹³C resonances were assigned. The $\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$ values plotted as a function of sequence (Figure 2.44) follow the same pattern as that in aqueous solution, showing that simply inducing intramolecular H-bonding is not sufficient to convert the β -hairpins into an α -helices, and that an anisotropic environment such as that provided by detergent micelles is also necessary. In this regard, LytA₂₃₉₋₂₇₂ behaves the same way as the isolated peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂.

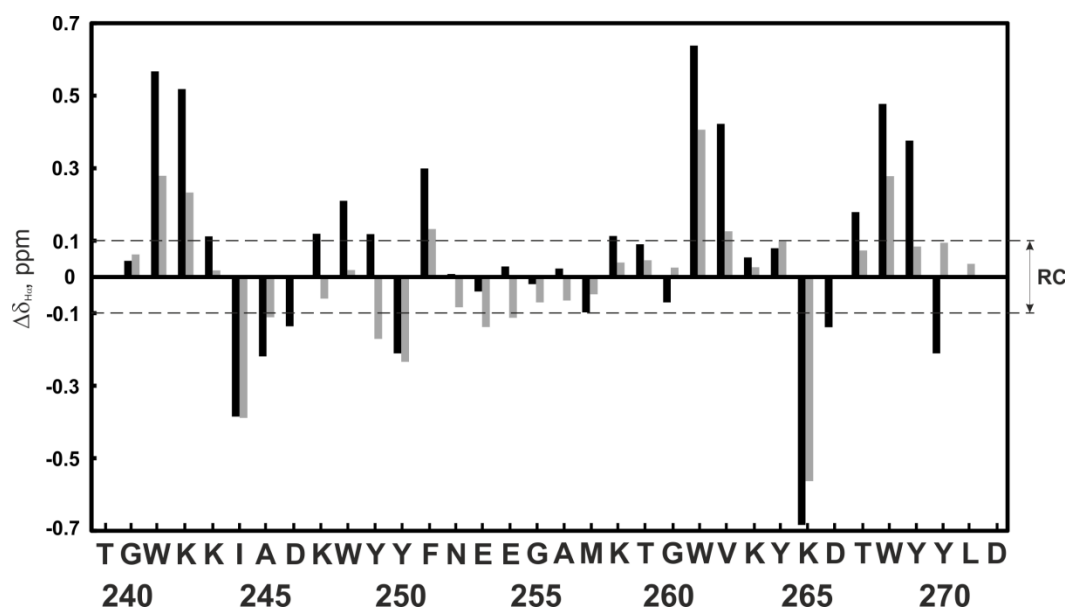


Figure 2.44. Plots of $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_{\alpha}^{\text{observed}} - \delta H_{\alpha}^{\text{RC}}$, ppm) as a function of LytA₂₃₉₋₂₇₂ peptide sequence in H₂O/D₂O 9:1 v/v (black bars), and in the presence of 30 % vol. TFE (grey bars). The N- and C-terminal residues are not shown. $\delta H_{\alpha}^{\text{RC}}$ values were taken from (Wishart et al., 1995). Measured at 35 °C and pH 3.0. Dashed lines indicate the random coil range (RC).

The $\Delta\delta H_\alpha$ plot shown in Figure 2.44 shows that LytA₂₃₉₋₂₇₂ possesses a similar secondary structure both in aqueous solution and in the presence of 30 % vol. TFE. However, some slight differences are detected. In the region corresponding to the first β -

hairpin (residues 239–252), the second β -strand seems to be less structured in the presence of TFE (see residues 247–249). In the linker segment (residues 253–258), the presence of TFE favours a tendency to helical structure (see also “Appendices”, Figure A3), which may be the origin of some of the helical features observed in the CD spectrum. Finally, in the region of the second β -hairpin (residues 259–272), $\Delta\delta H_\alpha$ values are rather conserved regardless of the presence of TFE.

Considering the CD and NMR data, it can be concluded that the presence of the TFE, which was irrelevant for the isolated peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, provokes some structural distortions in $\text{LytA}_{239-272}$. It is probable that the TFE, a known secondary structure stabilizer, exerts an influence on the disordered linker segment, favouring the appearance of some helical tendencies in it. This fact could affect the structure of the second strand of the first β -hairpin.

2.6. BEHAVIOUR IN THE PRESENCE OF LIPID VESICLES

Peptide $\text{LytA}_{239-272}$ was also examined in the presence of lipid vesicles. To compare with the results observed for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, two types of small unilamellar vesicles (SUVs) were used: DMPC:DMPG (3:1) vesicles, formed by choline-phospholipids, and POPE:POPG (2:1) vesicles, formed by non-choline phospholipids. As seen in Figure 2.45, CD spectra of $\text{LytA}_{239-272}$ in both types of vesicles exhibit a minimum at about 208 nm, a shoulder at about 222 nm, and a positive band at ~ 197 nm, which are characteristic of helical structures. The similarity between these spectra and that of DPC and SDS micelles (Figure 2.38) indicates that $\text{LytA}_{239-272}$ forms the same helix structure in SUVs and in micelles.

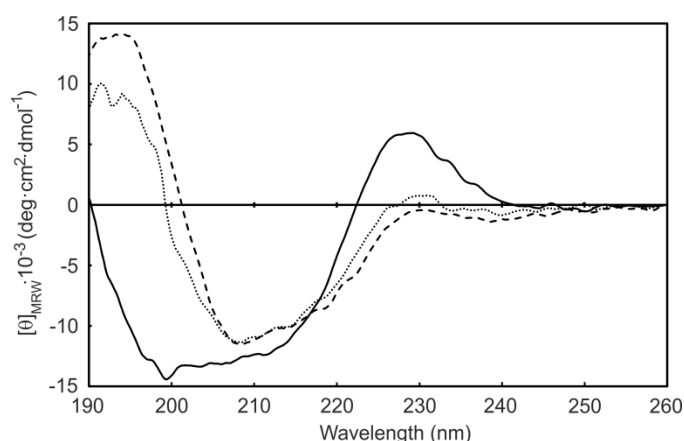


Figure 2.45. Far-UV CD spectra acquired for 12.8 μM $\text{LytA}_{239-272}$ samples. Solid line is from a sample in aqueous solution (10 mM phosphate buffer, pH 7.0), dashed line represents a sample in the same solvent plus DMPC:DMPG 3:1 (molar ratio) SUVs, and dotted line is from a sample in the same solvent plus POPE:POPG 2:1 (molar ratio) SUVs. The spectra were recorded at 30 $^{\circ}\text{C}$, peptide/lipid ratio was 1:50, and mean residue ellipticities are shown.

3. EFFECT OF DPC MICELLES ON THE FULL-LENGTH C-LYTA MODULE

The results shown so far demonstrate that some peptides derived from CBRs have the ability to interact with detergent micelles and undergo a dramatic conformational change. Nevertheless, CBRs are never found isolated in nature; they are arranged as linked units within the choline-binding modules, and usually display intramolecular interactions between

them. (B. Maestro & Sanz, 2005; Beatriz Maestro & Sanz, 2007; Medrano et al., 1996) Therefore, the question arose of whether the individual CBR propensities to become inserted into the micelles could be maintained in the framework of the full-length C-LytA module.

To answer the question, CD experiments were performed on the full-length C-LytA module in the presence of DPC and SDS micelles, and compared with the results obtained in aqueous solution. As shown in Figure 2.46A, at pH 7.0 and 25 °C, DPC micelles affect the far-UV CD spectrum of C-LytA, but SDS micelles clearly disrupt the anisotropic environment around the aromatic residues (loss of the positive band at 223 nm), while inducing an appreciable amount of α -helical structure (minimum at 208 nm and shoulder at 222 nm). At pH 3.0, where C-LytA is less stable (Hernández-Rocamora, Maestro, Mollá-Morales, & Sanz, 2008), DPC micelles are able to complete the hairpin to helix transition to SDS levels (Figure 2.46B). This suggests that insertion into the micelles requires some degree of flexibility in the protein to be accomplished and explains why SDS is more effective than DPC, because the former detergent is a strong denaturant that, in fact, has been described to fully unfold C-LytA at sub-micellar concentrations (B. Maestro & Sanz, 2005; Beatriz Maestro & Sanz, 2007). To investigate this hypothesis, the effect of DPC micelles on C-LytA at pH 7.0 at different temperatures was analysed. As shown in Figure 2.47, the CD spectrum at 5 °C is similar to that in the absence of micelles (Figure 2.46A), whereas, in contrast, at a physiological temperature in which C-LytA is more unstable (37 °C) (B. Maestro & Sanz, 2005; Sánchez-Puelles et al., 1990), a clear induction of α -helix can be seen, which is reversible upon cooling the sample. It can therefore be concluded that loosening the structure of the module either by temperature or pH greatly facilitates micelle insertion.

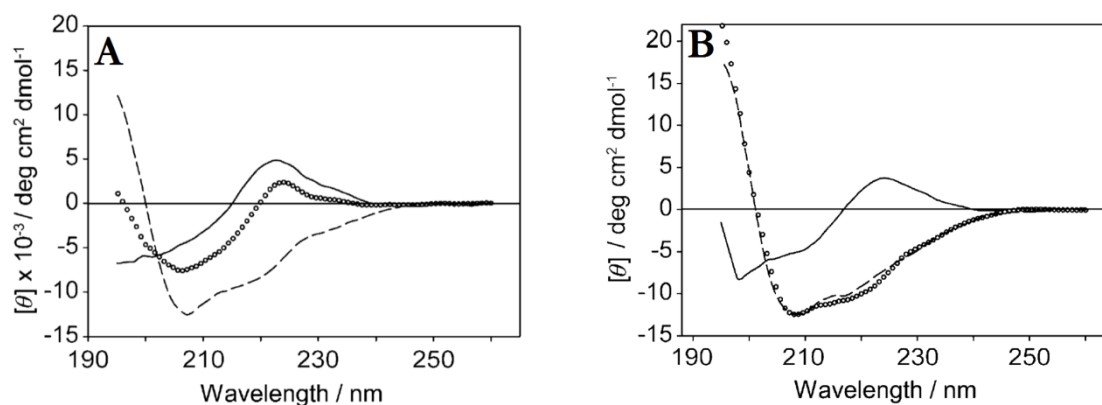


Figure 2.46. Effect of detergent micelles on the full-length C-LytA module: **A.** Experiments at pH 7.0 in the absence (solid line) and the presence of 30 mM DPC (circles) or 30 mM SDS (dashed line). **B.** Experiments at pH 3.0, same line scheme.

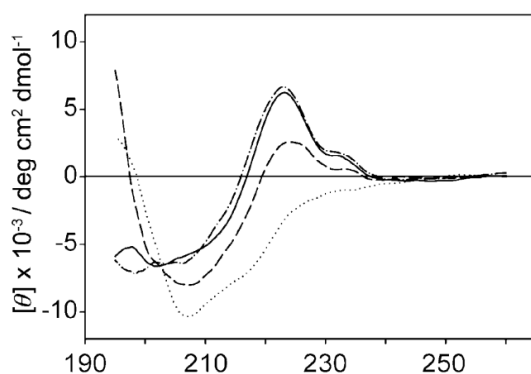


Figure 2.47. Effect of temperature on DPC-induced C-LytA conformational changes. Experiments at pH 7.0 and 30 mM DPC: 5 °C (solid line), 25 °C (dashed line), 37 °C (dotted line) and 5 °C recorded 16 h after heating (dashed-dotted line).

4. ARE DPC MICELLES GENERAL HELIX INDUCERS?

Most of the structures determined so far in the presence of DPC micelles correspond to cationic antimicrobial peptides and cell-penetrating peptides, and can be structurally classified as 1) α -helical-prone peptides, which are mainly unstructured in aqueous solution and become helical in the presence of the micelles (Mäler, 2013); 2) disulphide-rich peptides, some of which exhibit β -hairpin structures that are stabilised by one or more cross-strand disulphide bonds (Shenkarev et al., 2011; Usachev, Efimov, Kolosova, Filippov, & Klockhov, 2014), and 3) Trp-rich peptides with complex structural behaviours, such as indolicidin (Rozek, Friedrich, & Hancock, 2000) and puoroindoline derivatives. (Haney et al., 2013) None of these groups include linear peptides adopting β -hairpin structures in DPC micellar media. In fact, to our knowledge, only a linear octapeptide that adopts a β -hairpin structure in micelles has been reported (Mahajan & Bhattachariya, 2013), although this peptide contains a D-Pro-Gly turn sequence, which is known to nucleate β -hairpin structures (Santiveri, Santoro, Rico, & Jiménez, 2004; Stanger & Gellman, 1998), and also a myristoyl N-terminal extension. Therefore, to discard the possibility that the conformational change triggered by DPC micelles in peptides LytA_{239–252}, LytA_{259–272}, and LytA_{239–272} is a consequence of a general nonspecific helix-inducer effect of DPC micelles, other unrelated linear peptides known to form stable β -hairpins in aqueous solution were examined to check whether they maintain their β -hairpin or become helical in DPC micelles. To this end, two previously reported β -hairpin-forming peptides, SESYV11 and SESYW11 (Berry & Paton, 2000; Eldholm, Johnsborg, Haugen, Ohnstad, & Håvarstein, 2009; Santiveri et al., 2000; Santiveri et al., 2001; Santiveri et al., 2002) were selected to be studied by NMR in 30 mM [D₃₈]-DPC. After the complete assignment of the spectra (see “Appendices”, Tables A4 and A5), the profiles of $\Delta\delta H_{\alpha}$, (Figure 2.48), $\Delta\delta C_{\alpha}$, and $\Delta\delta C_{\beta}$ (see “Appendices”, Figures A4 and A5) values vs. sequences were plotted. They were quite similar to those obtained in aqueous solution, indicating that the two peptides maintain their β -hairpin structures in the DPC micellar media, hence, confirming that the DPC micelles do not have a nonspecific helix-inducer effect.

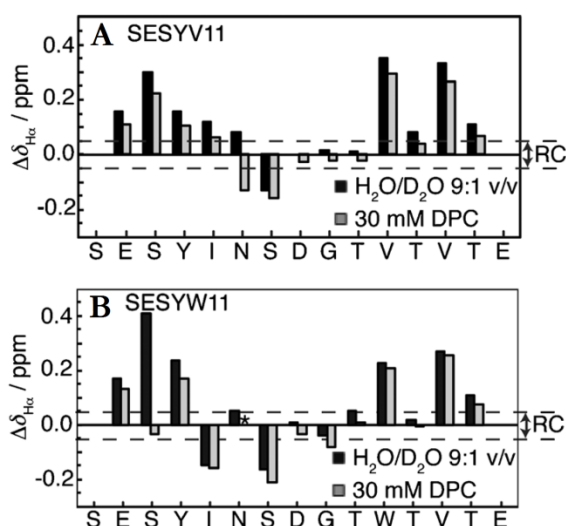


Figure 2.48. Bar plots of $\Delta\delta_{H\alpha}$ ($\Delta\delta_{H\alpha} = \delta_{H\alpha}^{\text{observed}} - \delta_{H\alpha}^{\text{RC}}$, ppm) as a function of sequence for the β -hairpin peptides SESYV11 (A) and SESYW11 (B) in H₂O/D₂O 9:1 v/v (in black; data taken from (Fesinmeyer et al., 2005; Santiveri, Rico, & Jiménez, 2000; Santiveri, Rico, & Jiménez, 2001; Santiveri, Santoro, Rico, & Jiménez, 2002)) and in 30 mM [D₃₈]-DPC (in grey). $\delta_{H\alpha}^{\text{RC}}$ values are taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range, and the asterisk indicates that the corresponding $\Delta\delta_{H\alpha}$ value was not determined.

5. STUDY OF LYTA₂₃₉₋₂₅₂ VARIANTS DESIGNED TO INVESTIGATE THE β -TO- α TRANSITION

5.1. DESIGN OF LYTA₂₃₉₋₂₅₂ VARIANTS

The physicochemical causes leading to the β -hairpin-to- α -helix transition observed in the studied peptides were further investigated. Based on the differences of hydrophobic and electrostatic surfaces shown by LytA₂₃₉₋₂₅₂ β -hairpin and α -helix (Figure 2.49, first line, detailed later in the “Discussion” section), a hypothesis was proposed: if a particular peptide sequence could form an amphipathic α -helix with a large hydrophobic patch, such α -helix would be formed in the presence of DPC and SDS micelles. However, amphipathic β -hairpins with large hydrophobic patches would remain unchanged in micelles. That is, detergent micelles stabilise amphipathic structures having an extensive hydrophobic patch on their surfaces. On the other hand, the LytA₂₃₉₋₂₅₂ sequence contains two Trp residues, which are both located at the hydrophobic face of the α -helix, and oriented towards the micelle core. Hence, taking into account that Trp residues are also known to be important for peptide-membrane and protein-membrane interactions (de Planque et al., 1999; Kachel, Asuncion-Punzalan, & London, 1995; Wimley & White, 1993; Yau, Wimley, Gawrisch, & White, 1998), these two residues might be also essential for micelle-triggered α -helix formation in peptide LytA₂₃₉₋₂₅₂. In this context, a series of LytA₂₃₉₋₂₅₂ variants were designed and synthesized, pursuing the obtaining of a less amphipathic α -helix, maintaining the non-amphipathic character of the β -hairpin and its conformation in aqueous solution.

The simplest way to achieve this is by exchanging the charged residue K5 with either W3 (peptide K3W5-LytA₂₃₉₋₂₅₂) or W10 (peptide W5K10-LytA₂₃₉₋₂₅₂). As can be seen in the helical wheels shown in Figure 2.49, these exchanges disrupt α -helix amphipathicity by introducing a positively charged side chain in the middle of the hydrophobic face. Considering that the side chains of residues W3, K5 and W10 are at the same β -sheet side

(downwards in the β -hairpin shown in Figure 2.13A), β -hairpin formation is expected to be unaffected or hardly affected. A third variant (peptide S3S10-LytA₂₃₉₋₂₅₂) introduces two polar side chains in the hydrophobic α -helix face by replacing the two Trp residues (W3 and W10) by Ser residues. In this case, β -hairpin formation might be somehow affected since the β -sheet propensity is lower for Ser than for Trp. (Fujiwara, Toda, & Ikeguchi, 2012)

To test whether the indole aromatic ring from the Trp residues also contribute to stabilise the α -helix in micelles, two variants were designed with different criteria: a peptide leading to an amphipathic α -helix but without Trp residues and maintaining the non-amphipathicity of the β -hairpin (peptide I3V10-LytA₂₃₉₋₂₅₂), and a peptide sequence leading to a non-amphipathic α -helix, maintaining the two Trp residues, and to an amphipathic β -hairpin (peptide I5Y6T11T13-LytA₂₃₉₋₂₅₂) (Figure 2.49). The first criterion was achieved in peptide I3V10-LytA₂₃₉₋₂₅₂ by the substitution of the two Trp residues by two β -branched hydrophobic residues, i.e., an Ile and a Val. β -branched residues, which are good β -sheet-formers were selected, to maintain β -hairpin formation in aqueous solution. The two hydrophobic residues were different to facilitate NMR spectral assignment. To achieve the second criterion it was needed to substitute four residues of the wild type LytA₂₃₉₋₂₅₂ sequence. Considering the β -hairpin structure, the Trp-containing β -sheet face became more hydrophobic by changing the charged residue Lys at position 5 to a hydrophobic β -branched residue, an Ile. At the same time, the other β -sheet face turned out more hydrophilic by changing the aromatic residues at positions 11 and 13 to the more polar Thr residues, and the Ile at position 6 to a Tyr. Thr residues at positions 11 and 13 were selected instead of other more polar or charged residues such as Ser or Asp because of their better β -sheet propensities. (Fujiwara et al., 2012)

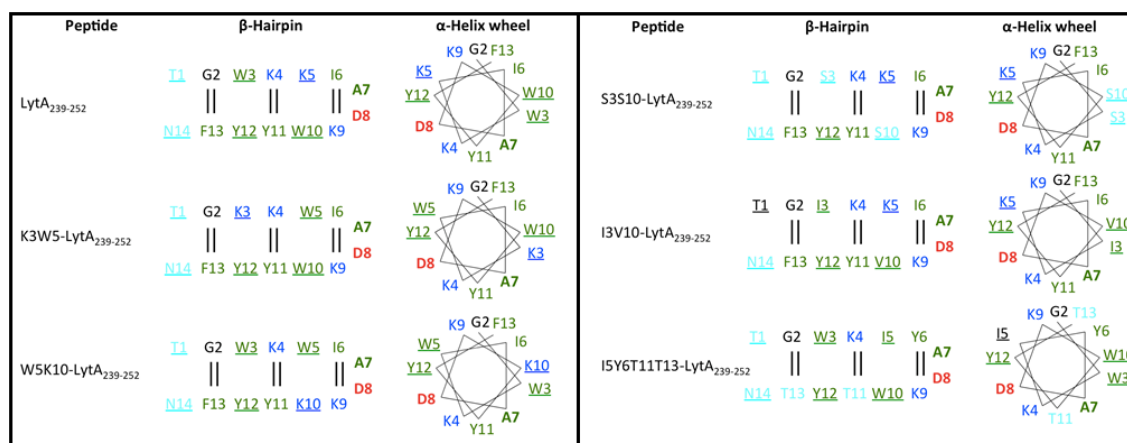


Figure 2.49. β -hairpin schemes and α -helical wheel representations for the structures formed by peptide LytA₂₃₉₋₂₅₂ in aqueous solution and in the presence of detergent micelles, respectively (Zamora-Carreras et al., 2015), and putative for the designed LytA₂₃₉₋₂₅₂ variants. In the β -hairpin schemes, hydrogen bonds are indicated by vertical lines, residues at non-hydrogen binding sites are underlined, and turn residues are shown in bold. Positively charged residues (K) are in blue, negatively charged (D) in red, polar residues (S, T, N) in cyan, and aromatic (F, Y, W), and aliphatic (I, V, A) in green. Numbering of the peptide positions 1 to 14 will be used in the next pages.

5.2. CD AND NMR STUDY IN AQUEOUS SOLUTION

To check whether the designed peptide variants in aqueous solution behave like the wild type sequence and form β -hairpin structures (Figure 2.50), far-UV CD experiments were performed.

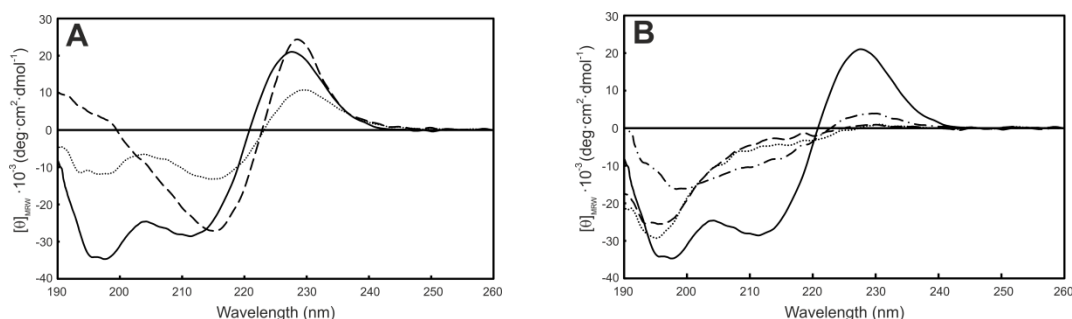


Figure 2.50. Far-UV CD spectra of LytA₂₃₉₋₂₅₂ variants in aqueous solution (20 mM Gly-HCl buffer, pH 3.0, 30 °C). A) Far-UV CD spectra superposition for LytA₂₃₉₋₂₅₂ wild type (solid line), K3W5-LytA₂₃₉₋₂₅₂ (dashed line), and W5K10-LytA₂₃₉₋₂₅₂ (dotted line). B) Far-UV CD spectra superposition for LytA₂₃₉₋₂₅₂ wild type (solid line), S3S10-LytA₂₃₉₋₂₅₂ (dashed line), I3V10-LytA₂₃₉₋₂₅₂ (dotted line), and I5Y6T11T13-LytA₂₃₉₋₂₅₂ (dashed-dotted line).

As observed in the far-UV CD spectra (Figure 2.50A), K3W5-LytA₂₃₉₋₂₅₂ and W5K10-LytA₂₃₉₋₂₅₂ variants show the strong positive bands around 227 nm, arising from the aromatic rings in rigid conformations. The slight wavelength shift and variations in the intensity reveal the changes occurred in the environment of the aromatic rings because of the sequence modification. Spectrum from K3W5-LytA₂₃₉₋₂₅₂ (Figure 2.50A, dashed line) displays more pronounced changes in the 195–205 nm regions, where the minima observed for the parent peptide are not present. This reveals that certain conformational distortion is occurring. In the case of W5K10-LytA₂₃₉₋₂₅₂, far-UV CD spectrum is characterised by a shape rather similar to that observed for the parent peptide, but with a decreased intensity. This could mean that the original β -hairpin conformation is essentially maintained, but its population is lower.

The far-UV CD spectra of S3S10-LytA₂₃₉₋₂₅₂, I3V10-LytA₂₃₉₋₂₅₂, and I5Y6T11T13-LytA₂₃₉₋₂₅₂ variants (Figure 2.50B), expose a large loss of the original β -hairpin structure. In the case of S3S10-LytA₂₃₉₋₂₅₂ and I3V10-LytA₂₃₉₋₂₅₂, spectra are those typical of random coil peptides (Figure 2.50B, dashed line and dotted line, respectively), with the characteristic minimum around 195 nm and no positive ellipticity values. The lower β propensity of the introduced amino acids, and the loss of important stabilizing interactions between aromatic side chains (since the two Trp are missing) are probably the main causes of the peptide destructure. On the other hand, the spectrum of peptide I5Y6T11T13-LytA₂₃₉₋₂₅₂ (Figure 2.50B, dashed-dotted line) points to the presence of some residual structure, despite the evident dramatic changes undergone.

LytA₂₃₉₋₂₅₂ variants in aqueous solution were also examined by NMR. A series of NMR spectra acquired for every peptide in H₂O/D₂O 9:1 v/v and in D₂O at pH 3.0 (see

Materials and methods) was acquired. The ^1H and ^{13}C chemical shifts were fully assigned for all the peptide variants by the standard strategy (see “Materials and methods” section, and Tables A6, A7, A8, A9 and A10 in “Appendices”). Once assigned the chemical shifts, the conformational shifts for H_α protons, C_α carbons, and C_β carbons were calculated and plotted as a function of the sequence (Figure 2.51 and Figure A6 in “Appendices”).

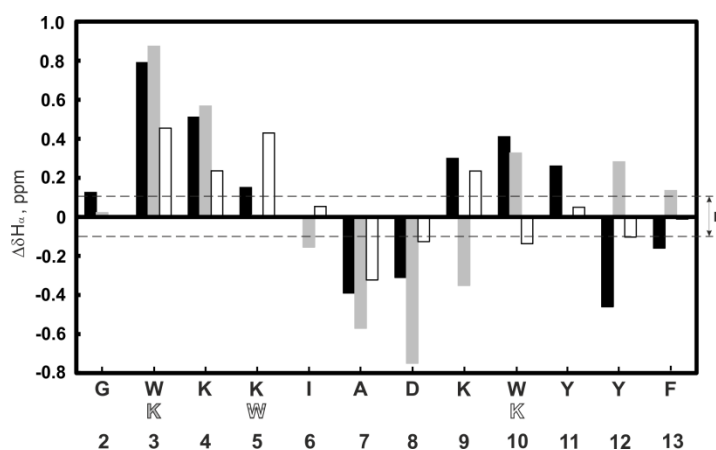


Figure 2.51. Bar plot of $\Delta\delta\text{H}_\alpha$ ($\Delta\delta\text{H}_\alpha = \delta\text{H}_{\alpha\text{observed}} - \delta\text{H}_{\alpha\text{RC}}$) as a function of sequence for peptide LytA₂₃₉₋₂₅₂ (black bars), K3W5-LytA₂₃₉₋₂₅₂ (grey bars), and W5K10-LytA₂₃₉₋₂₅₂ (white bars) in D₂O at pH 3.0 and 25 °C. $\delta\text{H}_{\alpha\text{RC}}$ values were taken from (Wishart, Bigam, Holm, Hodges, & Sykes, 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

As explained before, the characteristic $\Delta\delta\text{H}_\alpha$ profile of β -hairpins displays two stretches of positive values for the β -strands, separated by a short stretch of negative values at the turn region, as can be observed for LytA₂₃₉₋₂₅₂ (Figure 2.51, black bars) (Zamora-Carreras et al., 2015). The conformational shifts shown by the variant K3W5-LytA₂₃₉₋₂₅₂ conform to this pattern (Figure 2.51, grey bars). In the case of W5K10-LytA₂₃₉₋₂₅₂, the β -hairpin pattern is slightly distorted, as conformational shifts show small values in the region corresponding to the second β -strand. This fact suggests a diminished structural organization in this part of the peptide. The β -hairpin percentages estimated from the $\Delta\delta\text{H}_\alpha$ averaged for the β -strand residues are 79 % for peptide K3W5-LytA₂₃₉₋₂₅₂, 70 % for W5K10-LytA₂₃₉₋₂₅₂, and 100 % for the parent peptide (Table 2.11). These two variants formed β -hairpin structure, but slightly less populated than LytA₂₃₉₋₂₅₂.

Table 2.11. Averaged $\Delta\delta\text{H}_\alpha$, $\Delta\delta\text{C}_\alpha$ and $\Delta\delta\text{C}_\beta$, values in aqueous solution and in DPC micelles, and structure populations estimated from $\Delta\delta\text{H}_\alpha$ values measured at pH 3.0 and 25 °C. In aqueous solution, averaged $\Delta\delta\text{H}_\alpha$ and $\Delta\delta\text{C}_\beta$ are the mean of all positive values, and averaged $\Delta\delta\text{C}_\alpha$ the mean of all negative values. In DPC micelles, averaged $\Delta\delta\text{H}_\alpha$ is the mean of all negative values, and $\Delta\delta\text{C}_\alpha$ of all positive values. In all the cases, N- and C-terminal residues, as well as Gly 2 are excluded.

Peptide	Aqueous solution				DPC micelles		
	$\Delta\delta\text{H}_\alpha$, ppm	$\Delta\delta\text{C}_\alpha$, ppm	$\Delta\delta\text{C}_\beta$, ppm	% β -hairpin	$\Delta\delta\text{H}_\alpha$, ppm	$\Delta\delta\text{C}_\alpha$, ppm	% α -helix
LytA ₂₃₉₋₂₅₂	0.43	-0.52	2.12	100	- 0.26	1.5	67
K3W5-LytA ₂₃₉₋₂₅₂	0.31	-0.53	1.88	79	- 0.23	1.4	63
W5K10-LytA ₂₃₉₋₂₅₂	0.28	-0.39	1.35	70	- 0.22	1.6	57
S3S10-LytA ₂₃₉₋₂₅₂	< 0.01	-0.09	0.26	0	- 0.08	0.5	21
I3V10-LytA ₂₃₉₋₂₅₂	0.07 ^a	-0.13	0.52	0	- 0.21	2.0	58
I5Y6T11T13-LytA ₂₃₉₋₂₅₂	0.06	-0.14	0.67	16	- 0.16	1.2	43

^aThis value corresponds only to residues 3–5.

The acquired NOESY spectra were thoroughly analysed to identify structurally relevant NOE signals. β -hairpin characteristic NOE signals were observed for the variant W5K10-LytA₂₃₉₋₂₅₂, in particular those corresponding to the correlation between H_α from residues facing each other, and some NOE signals from side chains compatible with β -hairpin structure (Figure 2.52 and Table 2.12). In the case of K3W5-LytA₂₃₉₋₂₅₂, the distinctive H_α - H_α NOE signals were not distinguishable in the spectra (their chemical shifts were too close to the water signal), but many side chain NOE signals were consistent with the presence of β -hairpin structure (Table 2.10). In addition, the edge-to-face W5W10 interaction was confirmed for K3W5-LytA₂₃₉₋₂₅₂, regarding the corresponding chemical shift values.

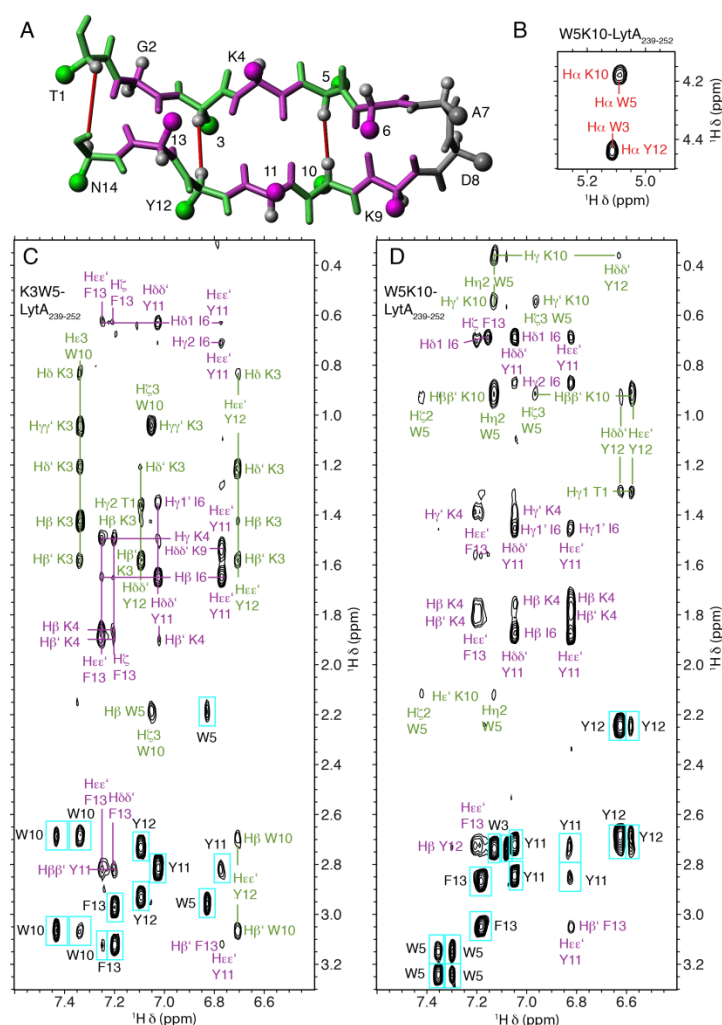


Figure 2.52. **A.** Backbone structure of the LytA₂₃₉₋₂₅₂ β -hairpin. Non-hydrogen sites are coloured in green, hydrogen-bonded sites in magenta, and turn residues in grey. H_α protons are displayed as small white spheres, and those pointing inwards connected by a red line. The side chain C_β carbons are shown as large spheres coloured according to their location at non-hydrogen-bonded sites (green), hydrogen-bonded sites (magenta), and turn residues (dark grey). Labels indicate the residue number (1 \equiv 239; 14 \equiv 252), and also the type of residue if maintained in all the designed variants. **B–D.** Selected NOESY spectral regions of K3W5-LytA₂₃₉₋₂₅₂ and W5K10-LytA₂₃₉₋₂₅₂ in D₂O at pH 3.0 and 25 °C. Panels **B** and **D** show non-sequential NOE signals involving side chain protons. Those from H-bonded residues are labelled in magenta, and those from non-H-bonded residues (the Trp-containing face) are in green. Intra-residual NOE cross-peaks are labelled in cyan. Panel **C** shows the H_α - H_α NOEs observed for W5K10-LytA₂₃₉₋₂₅₂.

Table 2.12. Summary of non-sequential cross-strand NOEs involving side chain protons observed for the peptides LytA₂₃₉₋₂₅₂ (Zamora-Carreras et al., 2015), K3W5–LytA₂₃₉₋₂₅₂, K5W10–LytA₂₃₉₋₂₅₂, and I5Y6T11T13–LytA₂₃₉₋₂₅₂ in aqueous solution. No cross-strand NOEs were observed for peptide S3S10–LytA₂₃₉₋₂₅₂, and only one, a diagonal between residues K5 and Y12, was found for peptide I3V10–LytA₂₃₉₋₂₅₂.

Side chains from non-HB residues	Peptide			
	LytA ₂₃₉₋₂₅₂	K3W5- LytA ₂₃₉₋₂₅₂	K5W10- LytA ₂₃₉₋₂₅₂	I5Y6T11T13- LytA ₂₃₉₋₂₅₂
Facing residues 5/10	K5/W10	W5/W10	W5/K10	I5/W10
No. NOEs	15	1	10	1
Facing residues 3/12	W3/Y12	K3/Y12	W3/Y12	W3/Y12
No. NOEs	8	7	0	0
Facing residues 1/14	T1/N14	T1/N14	T1/N14	T1/N14
No. NOEs	3	0	0	0
Diagonal residues 5/12	K5/Y12	W5/Y12	W5/Y12	I5/Y12
No. NOEs	0	0	0	1
Diagonal residues 3/10	W3/W10	K3/W10	W3/K10	W3/W10
No. NOEs	11	6	0	2
Diagonal residues 1/12	T1/Y12	T1/Y12	T1/Y12	T1/Y12
No. NOEs	1	1	2	1
Side chains from HB residues				
Facing residues 6/9	I6/K9	I6/K9	I6/K9	Y6/K9
No. NOEs	2	0	0	4
Facing residues 4/11	K4/Y11	K4/Y11	K4/Y11	K4/T11
No. NOEs	0	3	4	0
Diagonal residues 6/11	I6/Y11	I6/Y11	I6/Y11	Y6/T11
No. NOEs	11	9	10	1
Diagonal residues 4/13	K4/F13	K4/F13	K4/F13	K4/T13
No. NOEs	12	6	3	0

*Numbering 1 to 14 is used to simplify the data, correspondence to protein numbering: 1≡239, 14≡252.

In the case of the variant S3S10–LytA₂₃₉₋₂₅₂, the magnitudes of the conformational shifts are very low, i.e. $|\Delta\delta H_\alpha| \leq 0.05$ ppm, $|\Delta\delta C_\alpha| \leq 0.5$ ppm, and $|\Delta\delta C_\beta| \leq 0.5$ ppm (Figure 2.53, grey bars, and Figure A8 in “Appendices”). Only exceptions are H_α of Ser10 Tyr11, and Tyr12, but small (in the range -0.09 to -0.11 ppm) and negative, and C_α (-0.7 ppm) and C_β (1.8 ppm) of Asp8, which are suggestive of a certain turn tendency. This peptide is mainly random coil in aqueous solution, though it might maintain certain β–turn formation around Asp8. Non-sequential NOE signals were not observed, which is in agreement with the peptide being mainly random coil (Figure 2.54A).

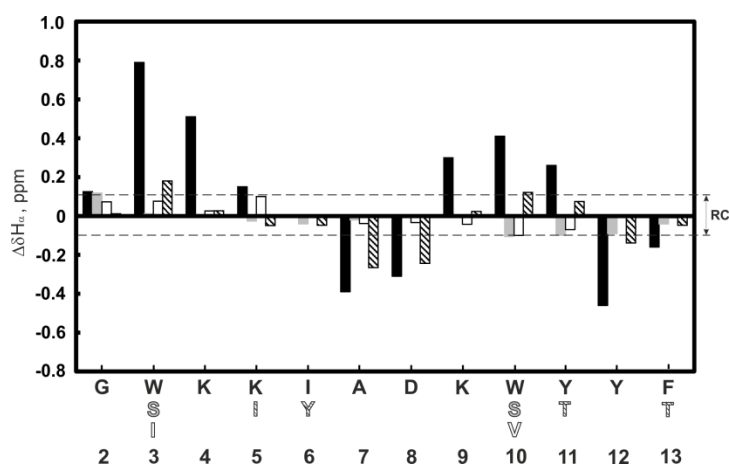


Figure 2.53. Bar plot of $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_{\alpha}^{\text{observed}} - \delta H_{\alpha}^{\text{RC}}$) as a function of sequence for peptide LytA₂₃₉₋₂₅₂ (black bars), S3S10–LytA₂₃₉₋₂₅₂ (grey bars), I3V10–LytA₂₃₉₋₂₅₂ (white bars), and I5Y6T11T13–LytA₂₃₉₋₂₅₂ (striped bars) in D₂O at pH 3.0 and 25 °C. $\delta H_{\alpha}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

Peptide I3V10–LytA₂₃₉₋₂₅₂ shows positive H_α values only at residues 3–5 (0.08, 0.03 and 0.1 ppm). The rest of residues display very low negative values ($|\Delta\delta H_\alpha| \leq 0.05$ ppm), except for Val10 and Tyr11 (–0.1 and –0.07 ppm) (Figure 2.53, white bars). The values of C_α (–0.6 ppm) and C_β (1.8 ppm) in Asp8 suggest a certain turn tendency (Figure A8 in “Appendices”). A few, weak non-sequential NOE signals observed for this peptide (Figure 2.54B and Table 2.12) suggest the presence of some low population of β –hairpin structure. These results indicate that I3V10–LytA₂₃₉₋₂₅₂ is a mainly random coil peptide, but certain residual β –hairpin population is present in the sample. NMR results for peptides S3S10–LytA₂₃₉₋₂₅₂ and I3V10–LytA₂₃₉₋₂₅₂ are in agreement with the previously described CD results.

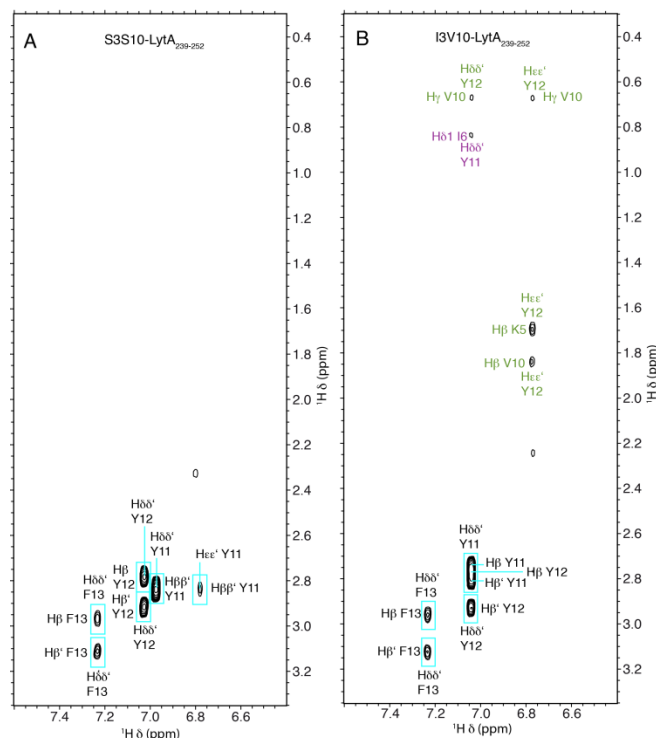


Figure 2.54. 2D ^1H – ^1H –NOESY spectra of peptides S3S10–LytA₂₃₉₋₂₅₂ (A) and I3V10–LytA₂₃₉₋₂₅₂ (B) in D₂O at pH 3.0 and 25 °C. Region showing NOE signals between aromatic protons (from Y11, Y12 and F13) and aliphatic protons.

Peptide I5Y6T11T13–LytA₂₃₉₋₂₅₂ forms a low populated β –hairpin structure. The profile of $\Delta\delta H_\alpha$ is similar to that from the parent peptide, but the magnitudes are rather small. The largest values are observed for the turn residues Ala7 and Asp8 (–0.27 and –0.24 ppm, respectively), and for the strand residues Trp3, Trp10 and Tyr11. The β –hairpin population estimated from the averaged $\Delta\delta H_\alpha$ for the β –strand residues (0.08 ppm) is 16 % (Table 2.11). β –hairpin formation is confirmed by the non-sequential NOE signals observed in the NOESY spectrum of this peptide (Figure 2.55, Table 2.12), which are compatible with the β –hairpin structure. The small number of these NOE signals is also in concordance with a low populated β –hairpin structure in aqueous solution.

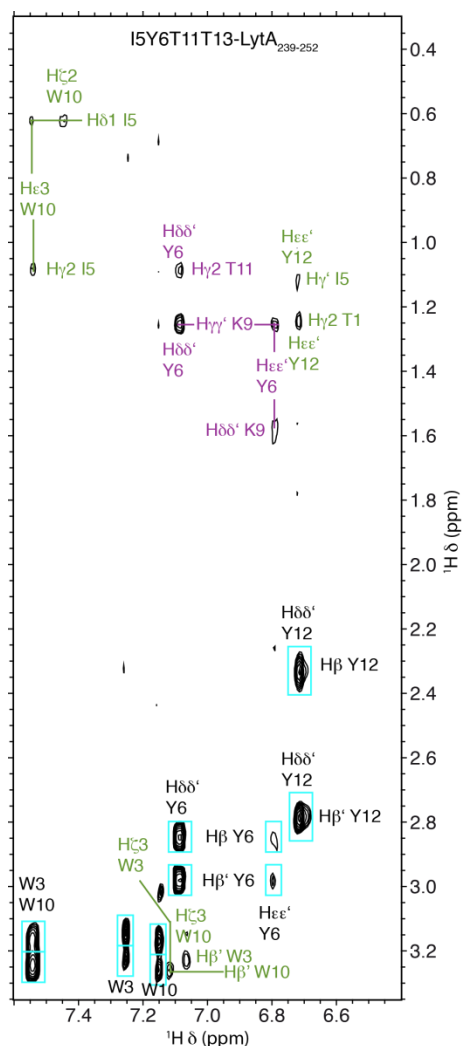


Figure 2.55. Selected 2D ^1H - ^1H -NOESY spectral region of peptide I5Y6T11T13-LytA₂₃₉₋₂₅₂ in D₂O at pH 3.0 and 25 °C showing non-sequential NOE signals involving side chain protons. Those from H-bonded residues are labelled in magenta, and those from non-H-bonded residues (the Trp-containing face) are in green. Intra-residual NOE cross-peaks are labelled in cyan.

5.3. CD AND NMR STUDY IN DPC MICELLES

To explore the effect of the sequence modifications executed on LytA₂₃₉₋₂₅₂ on the structural transition in the presence of a membrane-like milieu, the peptide variants were studied in the presence of DPC micelles.

The far-UV CD spectra of the three variants which form β -hairpin structure in aqueous solution (K3W5-LytA₂₃₉₋₂₅₂, W5K10-LytA₂₃₉₋₂₅₂, and I5Y6T11T13-LytA₂₃₉₋₂₅₂) are completely different in the absence and in the presence of DPC micelles (Figure 2.56A). Similarly to LytA₂₃₉₋₂₅₂, the three peptides in aqueous solution show a positive band at around 230 nm, which can arise from aromatic rings in rigid environments. The low intensity of this band in peptide I5Y6T11T13-LytA₂₃₉₋₂₅₂ is in agreement with the low β -hairpin population deduced from the NMR data (Table 2.11). The far-UV CD spectra of the three peptides in 30 mM DPC show a positive band at 197 nm and negative bands at 208 and 222 nm, which are typical of helices. Thus, these three peptides also act as conformational switches, as previously found for the parent peptide LytA₂₃₉₋₂₅₂.

Regarding the two peptides disordered in aqueous solution, the far-UV CD spectra of peptide S3S10-LytA₂₃₉₋₂₅₂ in the absence and in the presence of DPC micelles are almost identical (Figure 2.56B), suggesting that this peptide remains mainly disordered in DPC micelles. In contrast, the far-UV CD spectra of peptide I3V10-LytA₂₃₉₋₂₅₂ are different in aqueous solution and in 30 mM DPC (Figure 2.56B). Indeed, the far-UV CD spectrum of peptide I3V10-LytA₂₃₉₋₂₅₂ in DPC micelles displays the bands characteristic of helices, i.e. a positive band at 197 nm and negative bands at 208 and 222 nm.

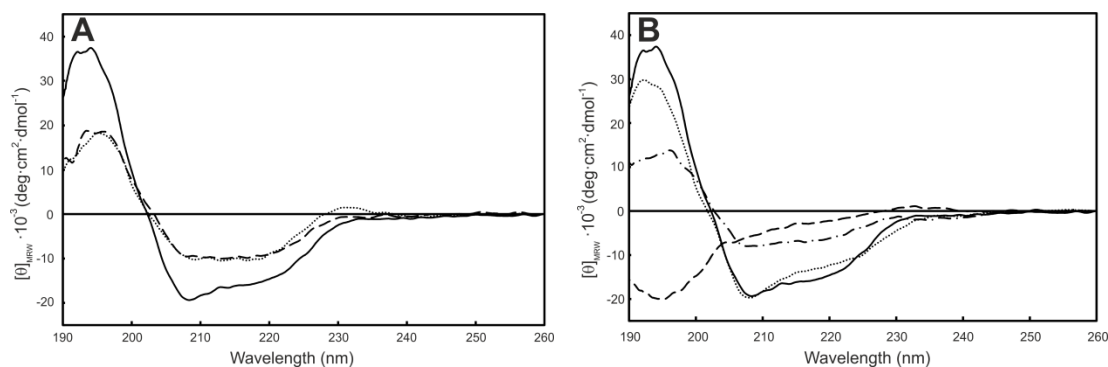


Figure 2.56. Far-UV CD spectra of LytA₂₃₉₋₂₅₂ variants in DPC micelles (30 mM DPC in 20 mM Gly-HCl buffer, pH 3.0, 30 °C). A) Far-UV CD spectra superposition for LytA₂₃₉₋₂₅₂ wild type (solid line), K3W5-LytA₂₃₉₋₂₅₂ (dashed line), and W5K10-LytA₂₃₉₋₂₅₂ (dotted line). B) Far-UV CD spectra superposition for LytA₂₃₉₋₂₅₂ wild type (solid line), S3S10-LytA₂₃₉₋₂₅₂ (dashed line), I3V10-LytA₂₃₉₋₂₅₂ (dotted line), and I5Y6T11T13-LytA₂₃₉₋₂₅₂ (dashed-dotted line).

To better characterise the structural behaviour of the designed peptides in the presence of DPC micelles, they were studied in 30 mM DPC by NMR. As in aqueous solution, once the ¹H and ¹³C chemical shifts were fully assigned by standard strategies (see “Materials and methods” section, and Tables A6, A7, A8, A9 and A10 in “Appendices”), the conformational shifts ($\Delta\delta H_\alpha$, $\Delta\delta C_\alpha$ and $\Delta\delta C_\beta$) and the set of NOEs were analysed. In agreement with CD data, peptides K3W5-LytA₂₃₉₋₂₅₂, W5K10-LytA₂₃₉₋₂₅₂, I3V10-LytA₂₃₉₋₂₅₂, and I5Y6T11T12-LytA₂₃₉₋₂₅₂ in DPC micelles displayed negative $\Delta\delta H_\alpha$ values for residues 3–13, and positive $\Delta\delta C_\alpha$ values for residues 2–13 (Figure 2.57, and Figures A8 in “Appendices”), which corroborates that the four peptides form a helix spanning residues 3–13. Further evidence about helix formation comes from the sets of NOEs, which include strong sequential NN(*i*, *i*+1) NOEs, and medium-range NOEs $\alpha\beta$ (*i*, *i*+3), αN (*i*, *i*+3), αN (*i*, *i*+2), αN (*i*, *i*+4), and NN(*i*, *i*+2) (Figure 2.58), all of them representative of helices. Based on the averaged $\Delta\delta H_\alpha$ for residues 3–13, the estimated α -helix populations were 63 % for peptide K3W5-LytA₂₃₉₋₂₅₂, 57 % for W5K10-LytA₂₃₉₋₂₅₂, 58 % for I3V10-LytA₂₃₉₋₂₅₂, and 43 % for I5Y6T11T13-LytA₂₃₉₋₂₅₂ (Table 2.11). In brief, these peptides become helical in the presence of DPC micelles, but the helices are slightly less populated than in the parent peptide LytA₂₃₉₋₂₅₂.

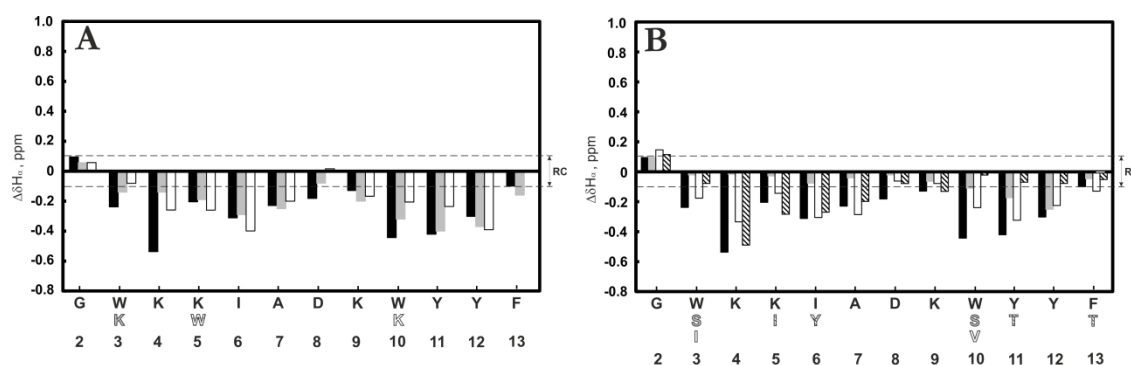


Figure 2.57. Bar plots of $\Delta\delta_{H\alpha}$ ($\Delta\delta_{H\alpha} = \delta_{H\alpha}^{\text{observed}} - \delta_{H\alpha}^{\text{RC}}$) as a function of peptide sequence. **A.** Plot for peptides LytA₂₃₉₋₂₅₂ (black bars), K3W5-LytA₂₃₉₋₂₅₂ (grey bars), and W5K10-LytA₂₃₉₋₂₅₂ (white bars). **B.** Plot for peptides LytA₂₃₉₋₂₅₂ (black bars), S3S10-LytA₂₃₉₋₂₅₂ (grey bars), I3V10-LytA₂₃₉₋₂₅₂ (white bars), and I5Y6T11T13-LytA₂₃₉₋₂₅₂ (striped bars). All samples were in a 30 mM DPC solution in D₂O at pH 3.0, and spectra were measured at 25 °C. $\delta_{H\alpha}^{\text{RC}}$ values were taken from (Wishart et al., 1995) The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

In the case of the peptide S3S10-LytA₂₃₉₋₂₅₂, as occurs with its CD spectra (Figure 2.56B), the profiles of $\Delta\delta_{H\alpha}$ and $\Delta\delta_{C\alpha}$ values observed in DPC micelles change very little relative to those in aqueous solution (Figure 2.57B). Although their magnitudes somewhat increase, most of the $\Delta\delta_{H\alpha}$ and $\Delta\delta_{C\alpha}$ values remain within the range typical of random coil peptides, except for those of residues 9–12. Also, the only two detected non-sequential NOEs involved residue K9, i.e. $\alpha\text{N}(i, i+3)$ between K9 and Y12, and that between the H α of I6 and H $\gamma\gamma'$ of K9, and the observed sequential NHi–NH_{i+1} are 6–7, 9–10, 11–12, and 13–14. All together, these data suggest that peptide S3S10-LytA₂₃₉₋₂₅₂ in DPC micelles is a mainly random coil peptide with a low populated short α -helix, probably spanning residues 6-13. Based on the averaged $\Delta\delta_{H\alpha}$ for residues 3–13, this peptide has about a 20 % α -helix population. In brief, peptide S3S10-LytA₂₃₉₋₂₅₂, which is disordered in aqueous solution, becomes only slightly helical in the presence of DPC micelles.

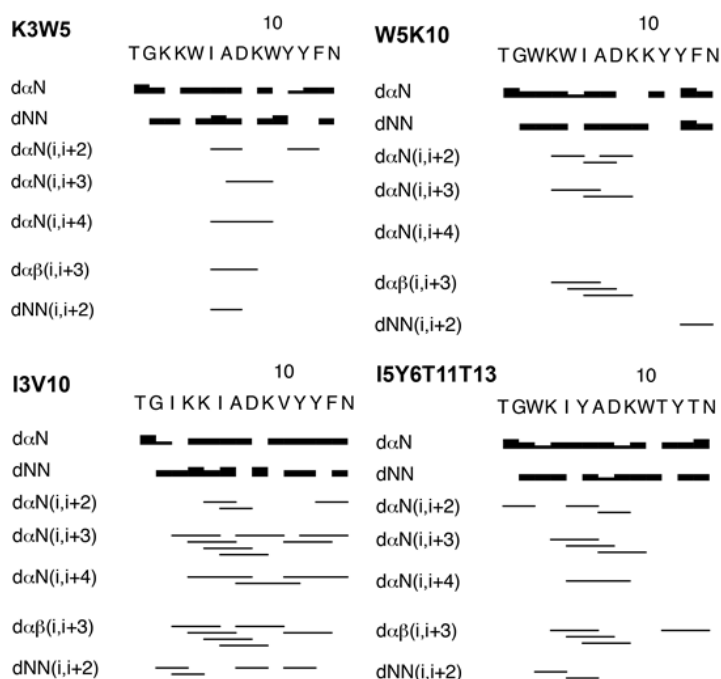


Figure 2.58. NOE summaries for peptides K3W5-LytA₂₃₉₋₂₅₂, W5K10-LytA₂₃₉₋₂₅₂, I3V10-LytA₂₃₉₋₂₅₂, and I5Y6T11T13-LytA₂₃₉₋₂₅₂, in 30 mM [D₃₈]-DPC, pH 3.0 (thickness of the lines is proportional to NOE signal intensities).

To confirm that helix formation was due to DPC micelles and not to its monomers, NMR spectra of the peptides at a DPC concentration below the cmc were acquired. The NMR spectra recorded in these conditions were essentially identical to those acquired in aqueous solution in the complete absence of DPC (Figure 2.59). Hence, as occurs in the parent peptide $\text{LytA}_{239-252}$, α -helix formation is induced by the presence of DPC micelles and not by DPC in its monomeric form.

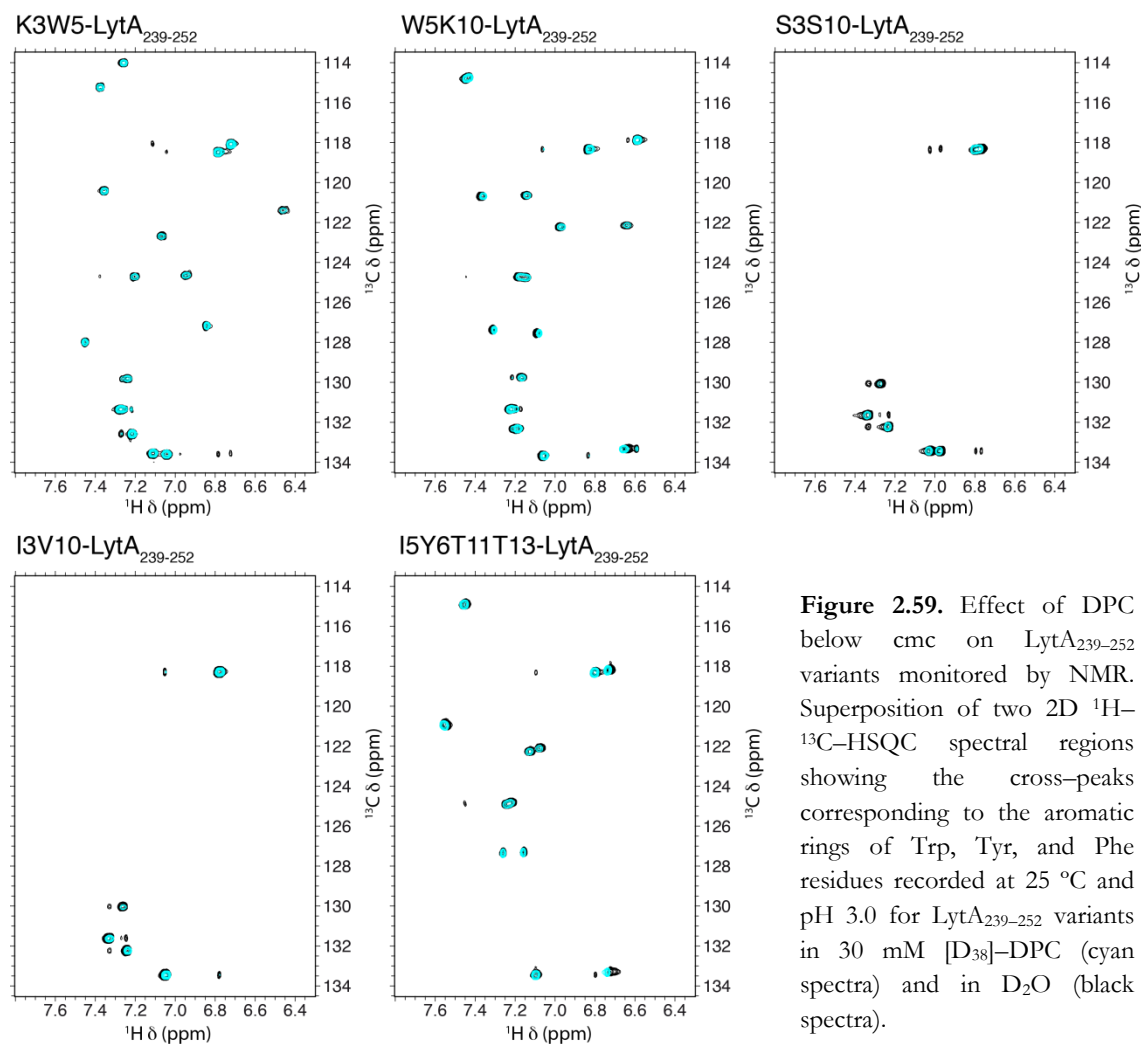


Figure 2.59. Effect of DPC below cmc on $\text{LytA}_{239-252}$ variants monitored by NMR. Superposition of two 2D ^1H – ^{13}C -HSQC spectral regions showing the cross-peaks corresponding to the aromatic rings of Trp, Tyr, and Phe residues recorded at 25 °C and pH 3.0 for $\text{LytA}_{239-252}$ variants in 30 mM $[\text{D}_{38}]$ -DPC (cyan spectra) and in D_2O (black spectra).

5.4. FLUORESCENCE MEASUREMENTS ON $\text{K3W5-LytA}_{239-252}$ AND $\text{W5K10-LytA}_{239-252}$

The fluorescence spectra of $\text{K3W5-LytA}_{239-252}$ and $\text{W5K10-LytA}_{239-252}$ variants in aqueous solution and in the presence of DPC micelles were recorded and compared to the results observed for the parent peptide (Figure 2.60)

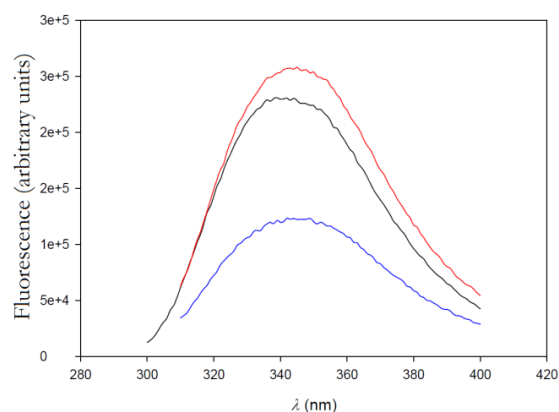
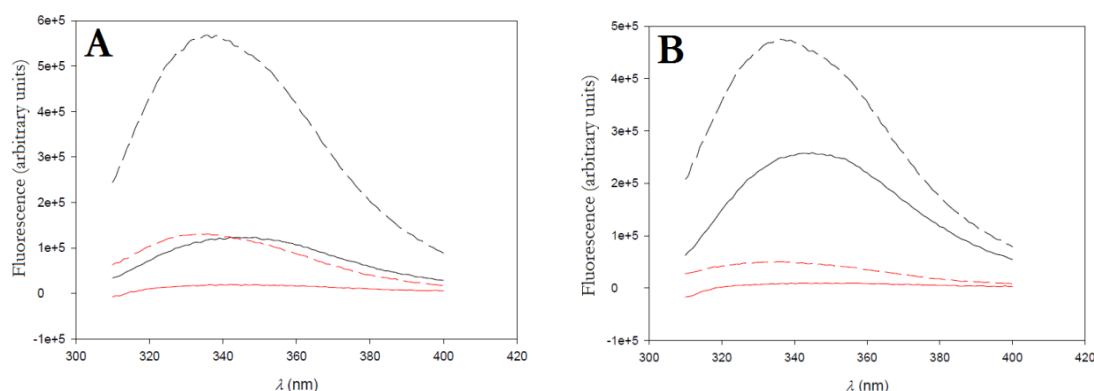


Figure 2.60. Intrinsic fluorescence of LytA₂₃₉₋₂₅₂ (black), K3W5-LytA₂₃₉₋₂₅₂ (blue), and W5K10-LytA₂₃₉₋₂₅₂ (red). All samples were in 20 mM HCl-Gly buffer, pH 3.0 and 25 °C, and the excitation wavelength was 280 nm.

As seen in Figure 2.60, K3W5-LytA₂₃₉₋₂₅₂ and W5K10-LytA₂₃₉₋₂₅₂ variants show a redshift of their maxima, compared to the parent peptide. These observations point to a higher exposition of the tryptophans to the solvent, derived from the fact that the structure is less compact in these variants as seen by NMR. Fluorescence quenching experiments were performed in the presence and the absence of DPC micelles for K3W5-LytA₂₃₉₋₂₅₂ and W5K10-LytA₂₃₉₋₂₅₂ variants (Figure 2.61).



Fluorescence quenching experiments for K3W5-LytA₂₃₉₋₂₅₂ (**A**), and W5K10-LytA₂₃₉₋₂₅₂ (**B**). Experiments were carried out in 20 mM HCl-Gly buffer, pH 3.0 and 25 °C, in the absence (solid lines) and the presence of 30 mM DPC (dashed lines), and in the absence (black) and the presence (red) of the quencher (500 mM acrylamide). The excitation wavelength was 280 nm.

A blueshift is observed upon the addition of DPC micelles, as a consequence of the insertion of one or both tryptophan rings in the micelles. As Trp aromatic rings come from a more exposed situation, the blueshift of the peptide variants is larger than the observed for the parent peptide (Figure 2.21A). Fluorescence intensities are lower for the variants and for those reasons the experiments were executed with different excitation and emission slits respect to the parent peptide. After the normalization of the spectra (Figure 2.62), it can be confirmed that the variants display larger blueshifts than the parent peptide, which is in agreement with the idea of a more superficial insertion of the Trp (probably, only one Trp is inserted and the other one remains exposed to the solvent).

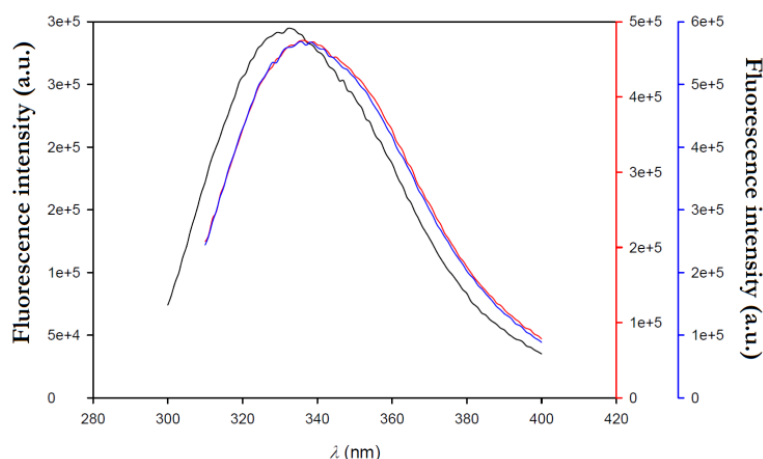


Figure 2.62. Normalized fluorescence spectra for LytA_{239–252} (black), K3W5–LytA_{239–252} (blue), and W5K10–LytA_{239–252} (red), in the presence of 30 mM DPC. Taken from figures 2.21A and 2.59. Colour code of the axes corresponds to colour code of the curves.

6. DISCUSSION

6.1. LYTA_{239–252}, LYTA_{259–272} AND LYTA_{239–272} HAVE NATIVE-LIKE β -HAIRPIN STRUCTURE AND MAINTAIN ABILITY TO BIND CHOLINE WITH LOW AFFINITY

In the work presented here, it has been demonstrated that two peptides derived from choline-binding repeats CBR3 and CBR4 from the CBD of pneumococcal autolysin LytA (LytA_{239–252} and LytA_{259–272}) form very stable native-like β -hairpins in aqueous solution, as previously reported for a peptide corresponding to CBR1, LytA_{197–210} (Beatriz Maestro et al., 2011). In addition, a longer peptide containing the sequences of LytA_{239–252} and LytA_{259–272} and the linker between them, LytA_{239–272} (see Figures 2.2 and 2.3), is also able to adopt very stable native-like β -hairpins in the 239–252 and 259–272 segments, but the relative position of the hairpins is not fixed since free rotation around the linker is allowed.

Thermodynamic analyses have proved that thermal denaturation of the three peptides takes place between two states with no intermediate states detected, as in the case of LytA_{197–210} (Beatriz Maestro et al., 2011). Calculated thermodynamic parameters indicate that LytA_{239–252} possesses the highest stability when compared with peptides LytA_{239–252} and LytA_{259–272}. Stability of the longer LytA_{239–272} seems to be intermediate to those observed for LytA_{239–252} and LytA_{259–272}, but the values of the thermodynamic parameters are closer to the latter. Thus, the overall stability is averaged between the two structured regions of LytA_{239–272}. Since the long LytA_{239–272} includes the sequences of the other two shorter peptides, and the corresponding segments adopt the native-like β -hairpin structures, it is reasonable to think that each segment reproduces the thermodynamic stability observed for the free peptides. The differences in stability between the two β -hairpins probably do not suffice to distinguish between the two transitions, so that a three state transition is not detectable in the long peptide. In addition, the β -hairpin that is more stable in the short peptide, formed by residues 239–259, is somehow less stable when in the context of the long peptide (100 % vs. 49 % β -hairpin population). Hence, stability differences between the two β -hairpins can be

even smaller in the long peptide than in the short peptides. The side chain interactions, determinants of β -hairpin stability, will be discussed in the next section.

Moreover, the ability of binding choline was tested for the three peptides. In all the cases, peptide–choline interaction was confirmed, and the affinity was calculated. $\text{LytA}_{239-252}$ showed the highest affinity, with a K_d of 80 ± 10 mM. $\text{LytA}_{239-252}$ and $\text{LytA}_{239-272}$ displayed a lower affinity, with K_d of 294 ± 28 mM and 184 ± 15 mM, respectively. As explained in the “Introduction” section, CBD from pneumococcal LytA possesses four choline-binding sites that are located between two consecutive CBRs and require the participation of aromatic residues from the β -hairpins present in those CBRs. Bearing this in mind, it was reasonable to expect very low or nil interaction in the short peptides, and a higher affinity in the long one. However, the three peptides were able to bind choline and, surprisingly, the highest affinity was observed for $\text{LytA}_{239-252}$, and not for $\text{LytA}_{239-272}$, a sequence containing the whole choline-binding site.

A plausible explanation for the lower affinity for choline of the long $\text{LytA}_{239-272}$ compared to the short $\text{LytA}_{239-252}$ may lie in the following two facts. First, the β -hairpin formed by the region 259–272 within the long peptide is less stable than the β -hairpin formed by the short peptide $\text{LytA}_{239-252}$. Second, the full canonical choline-binding site composed of aromatic residues from the two β -hairpins cores is not fixed, as it is in the native structure of the protein. The free rotation around the linker may difficult the optimal orientation of the β -hairpins to form the canonical choline-binding site.

It is remarkable the ability of binding choline of peptides $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, since they are significantly shorter than the canonical choline-binding sites. In the literature, a non-canonical choline-binding site comprising only two aromatic residues has been reported. In this short choline-binding site present in protein Pce (PDB code 2BIB), a Trp from a β -strand and a Tyr from the loop stabilise the choline molecule. (Galán-Bartual et al., 2015) Still, in the case of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$, only aromatics from the β -strands could be responsible for the stabilisation of choline.

6.2. DETERMINANTS OF β -HAIRPIN STABILITY IN AQUEOUS SOLUTION

$\text{LytA}_{239-252}$ and its variants showed a large variability in their ability to form β -hairpin structures, even though the variants were designed to affect as little as possible β -hairpin formation. This peptide system is probably very sensitive to changes at the strands because the sequence at the β -turn, which is conserved in the $\text{LytA}_{239-252}$ variants, is not optimal. It is well known that the characteristics of the β -turn are essential for β -hairpin formation (de Alba, Jiménez, & Rico, 1997; Hughes & Waters, 2006), and hence if the β -turn is not the most suitable, the stabilising or destabilising contributions from the strands are more noticeable than in peptides with optimal β -turns. In any case, it is interesting to analyse the origin of the observed differences in β -hairpin stability.

In the case of K3W5–LytA₂₃₉₋₂₅₂ and W5K10–LytA₂₃₉₋₂₅₂ variants, since two residues were exchanged (Figure 2.49), the β -sheet propensities are maintained compared to the parent peptide. On the other hand, as these changes take place on the same face of the β -hairpin, they do not have any effect on the side chain interactions of the non-Trp-containing face. According to this, the observed differences in β -hairpin populations should be explained by alterations occurring in side chain interactions of the Trp-containing face. In this face of the β -hairpin structure (Figure 2.13A; Zamora-Carreras et al. 2015), two stabilizing cross-strand side chain interactions are present in the parent peptide, a face-to-edge interaction between W3 and Y12, and a cation- π interaction between K5 and W10 (see also Figure 2.49). The presence of these two interactions can also be deduced from the chemical shifts of the involved residues. One residue in the two cross-strand pairs shows very large deviations from random coil values. In particular, the side chain protons of Y12 show chemical shifts characteristics of the edge aromatic ring of an edge-to-face interaction (see “Appendices”, Table A1; (Santiveri & Jiménez, 2010)), and the ^1H chemical shifts of the K5 side chain are up-field-shifted due to the anisotropy effects from the Trp indole ring (see “Appendices”, Table A1).

These two cross-strand pair interactions are also present in the W5K10–LytA₂₃₉₋₂₅₂ variant, but the “directionality” of the K/W interaction reverses. In the parent peptide, K and W are at positions 5 and 10, respectively, so that K belongs to the N-strand and W to the C-strand, whereas, in the variant, the W occupies position 5 at the N-strand, and the K position 10 at the C-strand (Figure 2.49). The existence of the two interactions is confirmed by the ^1H chemical shift deviations displayed by the side chains of Y12, the face residue of the face-to-edge W3/Y12 interaction, and of K10, from the cross-strand W5/K10 pair (see “Appendices”, Table A7). The fact that magnitudes of these chemical shift deviations in the W5K10–LytA₂₃₉₋₂₅₂ variant are slightly smaller than in the parent peptide (see “Appendices”, Table A1) agrees with the β -hairpin being less populated in the variant than in the parent peptide (Table 2.11). The directionality of the cross-strand K/W interaction must be the main responsible for the observed differences in β -hairpin populations, the K5/W10 pair being more stabilising than the W5/K10 pair. The existence of a “directionality effect” has been observed previously for cross-strand pair interactions. (Ramírez-Alvarado, Blanco, & Serrano, 2001; Russell, Blandl, Skelton, & Cochran, 2003; Russell & Cochran, 2000)

None of these two interactions is maintained in the case of K3W5–LytA₂₃₉₋₂₅₂ variant. The two cross-strand pairs in the Trp-containing face of this variant are K3/Y12 and W5/W10. The ^1H chemical shifts of the side chains of K3 and W5 deviate strongly from random coil values (see “Appendices”, Table A6). This confirms the presence of a cation- π interaction between K3 and Y12, and of an edge-to-face interaction between W5 and W10. Thus, the K3W5–LytA₂₃₉₋₂₅₂ variant and the parent peptide contain a cation- π and an aromatic-aromatic interaction in the Trp-containing face, but the residues involved in the two interactions differ. The cation- π pair in this variant is K/Y, instead of K/W in the parent peptide. This difference in the type of aromatic ring entails a diminution in terms of stabilizing energy, as it is generally considered that Trp is a better π -donor than Tyr. (Cochran et al., 2001; Russell & Cochran, 2000) As concerns the aromatic-aromatic interaction, in other β -hairpin systems (Santiveri & Jiménez, 2010), the edge-to-face W/W,

present in the K3W5–LytA₂₃₉₋₂₅₂ variant, has been proved to be more stabilizing than the face-to-edge W/Y interaction, present in the parent peptide. Hence, the slight β -hairpin destabilisation observed in this variant compared to the parent peptide (Table 2.11) comes from a fine balance of the energetic contributions of these two interactions.

The S3S10–LytA₂₃₉₋₂₅₂, and I3V10–LytA₂₃₉₋₂₅₂ variants are mainly disordered in aqueous solution (Table 2.11). In these variants, the substitution of the two W leads to differences in β -sheet propensities, and in side chain interactions in the β -sheet face where the W is located in the parent peptide (Figure 2.49). In the case of the S3S10–LytA₂₃₉₋₂₅₂ variant, the lower β -sheet propensity of Ser (0.86) in comparison with Trp (1.90) (Fujiwara, Toda et al., 2012), and the unfavourable cross-strand Ser/Tyr and Ser/Lys interactions (Wouters & Curmi, Proteins 1995) are undoubtedly contributing to the destabilization of the β -hairpin structure in aqueous solution. The complete loss of β -hairpin formation in the case of I3V10–LytA₂₃₉₋₂₅₂ variant is somehow unexpected, since β -sheet propensities for Ile and Val are good (2.02 and 2.31, respectively), even higher than that for Trp (1.90) (Fujiwara, Toda et al., 2012), and cross-strand I3/Y12 is a favourable hydrophobic interaction, though the Lys/Val is not favourable (Wouters & Curmi, Proteins 1995). This result indicates that strongly favourable cross-strands interactions are required to β -hairpin formation in the absence of optimal β -turn sequences.

The multiple sequence differences between I5Y6T11T13–LytA₂₃₉₋₂₅₂ variant and the parent peptide (Table 2.11) impede to explain the remarkable loss of β -hairpin stability in terms of particular contributing cross-strand interactions. It seems clear, however, that β -sheet propensities are not responsible for the β -hairpin destabilisation, since the overall β -sheet propensity is somehow even higher than the parent peptide. (Fujiwara, Toda et al., 2012). The ¹H-chemical shift deviations of side chains (see “Appendices”, Table A10) evidence the presence of two favourable interactions in this variant, the cross-strand face-to-edge W3/Y12 interaction, preserved from the parent peptide, and the hydrophobic cross-strand I5/W10 interaction (up-field shifts shown by the side chain protons of I5; see “Appendices”, Table A10), instead of the K5/W10 cation- π interaction observed in the parent peptide. Interactions at the non-Trp containing face, which are also completely different in the I5Y6T11T13–LytA₂₃₉₋₂₅₂ variant and in the parent peptide (Table 2.11), also account for the differences in β -hairpin stability.

All residue exchanges (peptides W3K5–LytA₂₃₉₋₂₅₂ and K5W10–LytA₂₃₉₋₂₅₂) and substitutions (peptides S3S10–LytA₂₃₉₋₂₅₂, I3V10–LytA₂₃₉₋₂₅₂, and I5Y6T11T13–LytA₂₃₉₋₂₅₂) executed on the β -strands of LytA₂₃₉₋₂₅₂ (Figure 2.49) destabilised the β -hairpin, even though a criterion in mind for the design was to maintain β -hairpin stability. This high sensitivity of the LytA₂₃₉₋₂₅₂ to any change in the strands probably arises from the fact that the β -turn sequence is not optimised for a 2:2 β -hairpin. In the absence of an optimal β -turn, favourable contributions from the strands, both β -sheet propensities and side chain interactions, become the key for β -hairpin stability. Trp residues highly contribute to β -hairpin stability, likely by aromatic-aromatic interactions (W/Y in LytA₂₃₉₋₂₅₂ and W5K10–LytA₂₃₉₋₂₅₂, and W/W in K3W5–LytA₂₃₉₋₂₅₂), and cation- π interactions (K/W in LytA₂₃₉₋₂₅₂, W/K in W5K10–LytA₂₃₉₋₂₅₂, and K/Y in K3W5–LytA₂₃₉₋₂₅₂).

LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ have very similar sequences (five residues are different; Figure 2.63), and display a significant difference in thermodynamic stability, as occurs with LytA₂₃₉₋₂₅₂ and its variants. The differences in sequence are localised at the turn region and at residues at non-hydrogen-bonding sites (the non-Trp containing face), whereas the stabilising cross-strand pairs at the Trp-containing face (W/W and K/W) are identical in the two peptides. Therefore, the difference in stability between these two peptides have to come from differences in turn propensities, even though both have a D residue at the position i+1, which is very favourable for turn formation, and from the differences in stabilising effect of cross-strand pair interactions at non-Trp containing face. It is not possible to identify particular residues or pair interactions as the main responsible for the observed difference in stability.

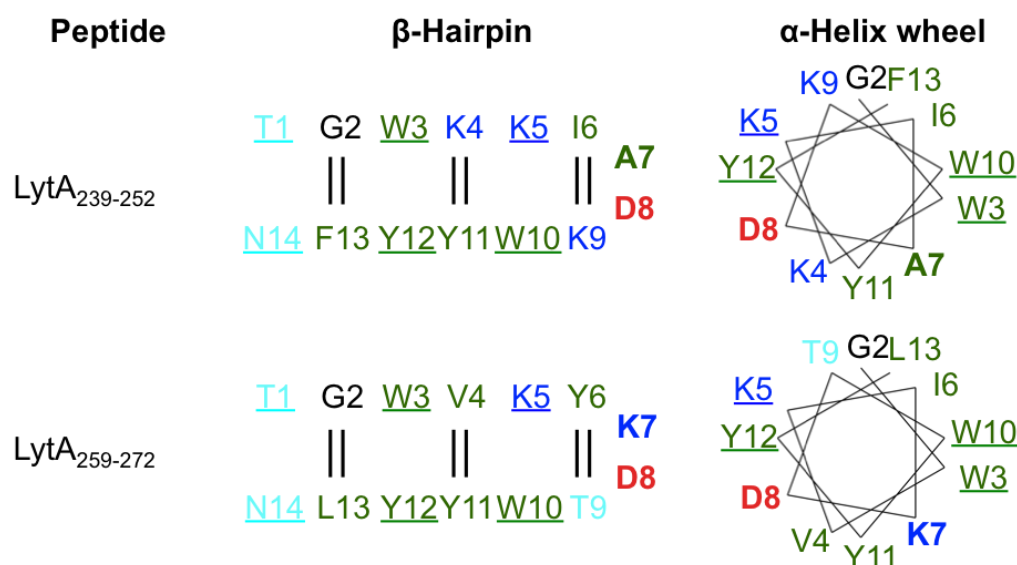


Figure 2.63. β-hairpin schemes and α-helical wheel representations for the structures formed by peptides LytA₂₃₉₋₂₅₂ and LytA₂₅₉₋₂₇₂ in aqueous solution and in the presence of detergent micelles, respectively (Zamora-Carreras et al., 2015). Hydrogen bonds are indicated by vertical lines, residues at non-hydrogen binding sites are underlined, and turn residues are shown in bold. Positively charged residues (K) are in blue, negatively charged (D) in red, polar residues (I) in cyan, and aromatic (F, Y, W), and aliphatic (I, V, A) in green. Numbering 1 to 14 is used to simplify the data, correspondence to protein numbering: 1=239, 14=252 for LytA₂₃₉₋₂₅₂; 1=259, 14=272 for LytA₂₅₉₋₂₇₂.

6.3. LYTA₂₃₉₋₂₅₂, A NOVEL SWITCH PEPTIDE

Some peptide sequences are able to adopt different conformations depending on different controllable external factors, such as pH (Schneider et al., 2002), metal binding (Anzini et al., 2013; Signarvic & DeGrado, 2009; Smith, Du, Radford, & Tezcan, 2013), redox conditions (Wang et al., 2012), and light irradiation (Aemissegger, Kräutler, van Gunsteren, & Hilvert, 2005). These isolated sequences –peptides– are called “conformational switches”. Reported conformational transitions in peptides include random coil to ordered secondary structures (Signarvic & DeGrado, 2009; Smith et al., 2013), or to self-assembled hydrogel β-sheets (Schneider et al., 2002); soluble monomeric α-helix to self-associated

oligomeric β -sheets (Wang et al., 2012); different registers in coil-coiled helices (Anzini et al., 2013), and dimer of coil-coiled helices to helical-hairpin (Pandya et al., 2004).

In this work, it has been demonstrated that a 14-residue peptide derived from the core of the third choline-binding repeat (CBR3) of the pneumococcal LytA autolysin, LytA_{239–252}, forms a very stable native-like β -hairpin and maintains this structure in the presence of TFE but, unexpectedly, it converts into a stable α -helix in the presence of DPC or SDS micelles, as well as in DMPC:DMPG and POPE:POPG SUVs. This α -helix can fold back into the native-like β -hairpin by dilution to sub-micellar detergent concentration.

Random coil to α -helix transitions have been reported to be induced by methanol or fluorinated alcohols such as TFE and HFIP (Buck, 1998) or, in the case of antimicrobial peptides, by micelles (Díaz et al., 2011). However, to our knowledge, LytA_{239–252} is the first documented case of a peptide that forms two completely different ordered structures depending on the solvent conditions. Moreover, reported chameleon sequences are up to seven residues long in natural proteins (Mezei, 1998), and even 11 in a particular designed sequence (Minor & Kim, 1996), so LytA_{239–252} represents the longest sequence known so far of this kind. Another interesting difference is that the LytA_{239–252} sequence can be predicted to form a β -hairpin by the program Betahairpred (<http://triton.iqfr.csic.es/software/behairpredv1.0/behairpred.htm>), but it is not predicted to be helical by AGADIR (<http://agadir.crg.es>), and only residues 240–246 show some very low helical propensity by PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) (data not shown).

It is known that certain sequences, known as “chameleon sequences” (Minor & Kim, 1996), can be either helical or extended, depending on their molecular context within the protein in which they are located. (Araki & Tamura, 2007; Sanz, Jiménez, & Giménez-Gallego, 2002) To investigate a possible relationship between the sequence of LytA_{239–252} peptide and reported chameleon sequences in proteins, the ChSeq database (<http://prodata.swmed.edu/chseq>) (Wenlin Li et al., 2015) was employed to look for coincidences. The sequence corresponding to the six central residues of LytA_{239–252} (²⁴²KKIADK²⁴⁷) was found to be a reported chameleon sequence. This six-residue fragment adopts a β -strand/turn conformation in protein LytA from *S. pneumoniae* (PDB code: 1GVM), as explained before, but it has a helix/turn conformation in other proteins, such as SsoPox from *Sulfolobus solfataricus* (PDB code: 2VC7) (Figure 2.64).

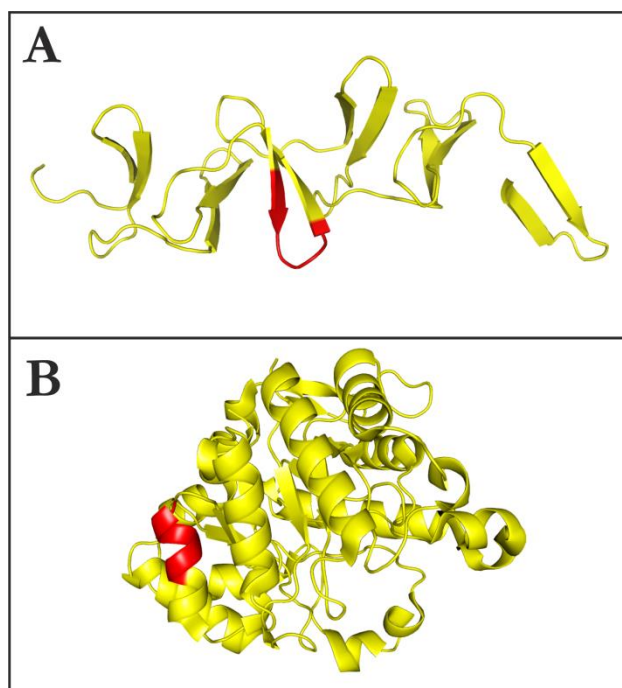


Figure 2.64. Chameleon sequence identified in $\text{LytA}_{239-252}$ ($^{242}\text{KKIADK}^{247}$). Chameleon sequence is highlighted in red in the crystallographic structures of pneumococcal LytA (monomer, PDB code: 1GVM) (**A**) and SsoPox from *S. solfataricus* (monomer, PDB code: 2VC7) (**B**), showing a β -strand/turn secondary structure, and a helical secondary structure, respectively.

A clear-cut difference between the two structures formed by $\text{LytA}_{239-252}$ is that whereas hydrophobic and polar side chains are evenly distributed between the two faces of the β -hairpin plane in aqueous solution (Figure 2.13A and 2.63), the α -helix is amphipathic, with hydrophobic residues clearly clustered in one face and polar/charged residues in the other (Figure 2.63). This amphipathic structure is very suitable to interact with a DPC micelle, as visualised in the model proposed for the peptide/micelle complex (Figure 2.26). In support, the side chains of the residues on the hydrophobic face are precisely those best defined in the NMR structure, probably as a consequence of their restricted mobility. In contrast, the non-amphipathic β -hairpin is not able to be inserted in the detergent micelles. β -hairpins formed by the control peptides SESYV11 and SESYW11 are more amphipathic than their putative helical structure, and hence more suitable to interact with micelles. This explanation is consistent with previous proposals concerning the importance of amphipathicity for the interaction with membranes of other peptides. (Hammen, Gorenstein, & Weiner, 1996)

6.4. OTHER LYTA SWITCH PEPTIDES

$\text{LytA}_{259-272}$, a peptide derived from the CBR4 from pneumococcal LytA choline-binding domain, which has a high sequence identity compared to $\text{LytA}_{239-252}$ (57 %; Figure 63), showed also a switching behaviour. CD and NMR data demonstrated that $\text{LytA}_{259-272}$ adopted a very stable native-like β -hairpin structure (though less stable than that of $\text{LytA}_{239-252}$, as already discussed), which was maintained in the presence of TFE. However, in the presence of DPC and SDS micelles, $\text{LytA}_{259-272}$ changed to a helical conformation, a change that was proved to be reversible. In brief, experimental data evidenced that $\text{LytA}_{259-272}$ is capable to act as a conformational switch, as found for the related peptide $\text{LytA}_{239-252}$.

The ChSeq database (<http://prodata.swmed.edu/chseq>) (Wenlin Li et al., 2015) was examined again to look for coincidences between $\text{LytA}_{259-272}$ sequence and chameleon sequences. In this case, the sequence corresponding to the five central residues of $\text{LytA}_{259-272}$ ($^{262}\text{VKYKD}^{266}$) was found to be a reported chameleon sequence. This five-residue fragment adopts a β -strand/turn conformation in protein LytA from *S. pneumoniae* (PDB code: 1GVM), as explained before, but it has a helix conformation in an endonuclease from *Fusarium graminearum* (PDB code: 4EFJ) (Figure 2.65).

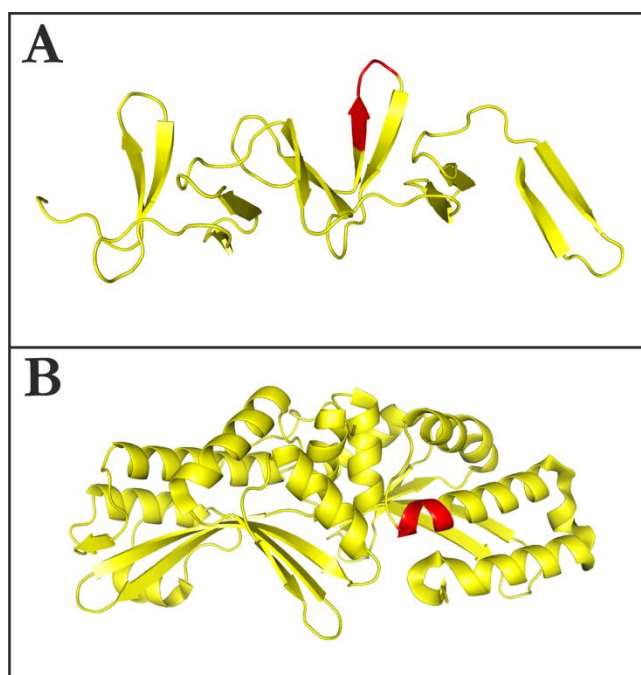


Figure 2.65. Chameleon sequence identified in $\text{LytA}_{259-272}$ ($^{262}\text{VKYKD}^{266}$). Chameleon sequence is highlighted in red in the crystallographic structures of pneumococcal LytA (monomer, PDB code: 1GVM) (**A**) and an endonuclease from *F. graminearum* (monomer, PDB code: 4EFJ) (**B**), showing a β -strand/turn secondary structure, and a helical secondary structure, respectively.

The possibility of a biological role of the conformational switching observed in these peptides in the context of the full-length protein lead to examine the conformational behaviour of peptide $\text{LytA}_{239-272}$, which contains the β -hairpin cores of CBR3 and CBR4, plus the linker between them (Figure 2.3B), and of the full-length LytA choline-binding domain in the presence of micelles. As previously explained, CD and NMR data showed that the regions corresponding to $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ adopted the native-like β -hairpin structure observed in aqueous solution for the isolated peptides, but the relative position of both hairpins was not fixed due to the flexibility of the linker. In the presence of DPC or SDS micelles, and even in SUVs, $\text{LytA}_{239-272}$ undergoes the β -to- α transition, and the observed helical structure is extended even through the linker. In the case of the full-length LytA choline-binding domain, CD data also evidences a β -to- α transition in the presence of DPC and SDS micelles.

It should be remarked that the conformational plasticity of peptides and proteins is at the basis of many relevant biological events. Well-known examples of these are the conformational helix to oligomer sheet transitions in the prion protein and the amyloid peptide, which cause important diseases. Thus, the conservation of the switching ability in the long peptide $\text{LytA}_{239-272}$, and in the full-length LytA choline-binding domain suggests that the switch could be biologically relevant for, at least, some of the biological processes in

which the full-length protein LytA is involved in the cell. In *S. pneumoniae*, access of the pneumococcal LytA amidase and other CBPs to the cell wall from the cytosol implies the interaction with and translocation across the cell membrane without the use of a signal peptide. Thus, CBRs both individually, in the context of a longer peptide, and also within the full-length CBD, seem to have the ability to undergo reversible disruption of their native structure and acquire an alternative helical conformation with the ability to recognise the lipid bilayer. This might constitute a general mechanism to complete the sorting of these proteins to the bacterial surface to carry out their physiological activity.

These results could be of big interest in the study of chameleon sequences, peptide switches, and in the emerging field of intrinsically disordered proteins. Additionally, these switching sequences may be useful for the design of new peptides and biosensors.

6.5. DETERMINANTS OF α -HELIX FORMATION IN DPC MICELLES

The fact that the K3W5–LytA₂₃₉₋₂₅₂ and W5K10–LytA₂₃₉₋₂₅₂ variants are able to form α -helices in the presence of DPC micelles demonstrate that a single Lys placed at the hydrophobic face of the helix does not impede its formation. The insertion of one positively charged residue in the hydrophobic patch (Figure 2.49) is not enough to cause a drastic change in the amphipathic nature of the helix. This situation is comparable to that of LytA₂₅₉₋₂₇₂, as its hydrophobic patch also contains a Lys that do not block the conformational transition.

Comparing α -helix formation in peptides S3S10–LytA₂₃₉₋₂₅₂ and I3V10–LytA₂₃₉₋₂₅₂ (Table 2.9), the proposal about that an amphipathic helix is stabilised in the presence of micelles was confirmed. This is the case of I3V10–LytA₂₃₉₋₂₅₂, which conserves the amphipathicity of the parent peptide (Figure 2.49). This variant does not form any ordered structure in aqueous solution, but it becomes helical in DPC micelles. Conversely, S3S10–LytA₂₃₉₋₂₅₂, which would form a non-amphipathic helix (Figure 2.49), remains mainly unstructured in the presence of DPC micelles.

To look for the importance of Trp side chains for the interaction with micelles, α -helix formation in the I3V10–LytA₂₃₉₋₂₅₂ variant and in the parent peptide can be compared (Table 2.11). In this variant, the helix amphipathicity is conserved, but the hydrophobic face does not contain any Trp residue. It was verified that, as observed in the parent peptide, the I3V10–LytA₂₃₉₋₂₅₂ variant turns into a α -helix in the presence of DPC micelles. However, helix population was lower in I3V10–LytA₂₃₉₋₂₅₂ than in the parent peptide (Table 2.11) indicating that the presence of Trp aromatic side chains are not essential, but it enhances the interaction with micelles, providing the helix with an additional stabilisation.

Further insights into the roles of amphipathicity and Trp side chains for α -helix formation in DPC micelles come from comparison of I5Y6T11T13–LytA₂₃₉₋₂₅₂ variant and the parent peptide. The I5Y6T11T13–LytA₂₃₉₋₂₅₂ variant maintains the two Trp residues, but in a hydrophobic patch less extended than the parent peptide (Figure 2.49). This variant becomes α -helical in DPC micelles, but the helix population is clearly lower than in the

parent peptide (Table 2.11). This result, taken together with those from the other variants, suggests that a decreased amphipathicity and a smaller hydrophobic patch destabilise α -helix formation in micelles, but the presence of the Trp residues favours interaction with micelles and partially counterbalances the destabilising contributions.

As mentioned before, in the case of $\text{LytA}_{259-272}$ the amphipathic character of the helix is maintained, and the presence of only one Lys residue in the hydrophobic patch is not enough to impede the structural transition. The conservation of most aromatic residues and their location in the sequence (i.e., W3, W10, Y11, Y12) when compared to $\text{LytA}_{239-252}$, suggests a very similar model of interaction with micelles for $\text{LytA}_{259-272}$, as proposed before. In brief, the similarities in amphipathicity and in the pattern of aromatic residues are the bases of the similar behaviour observed for $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$. In the case of the longer $\text{LytA}_{239-272}$, the presence of the sequences of the two short peptides provides the appropriate features to exhibit the same switching behaviour. The longer structure and the presence of a linker with a lower helical tendency suggest that the peptide/micelle interaction may be described by a more complex model, which will be studied more deeply in the future.

Regarding the DPC micelle triggered α -helix, the differential ability of the parent peptide and its variants to form an α -helix upon interaction with DPC micelles indicates that, as expected, amphipathicity plays a determining role for α -helix formation in micelles. Thus, two polar residues at the hydrophobic face of the helix led to an almost completely disordered peptide, only a residual α -helix is detected (peptide S3S10- $\text{LytA}_{239-252}$). However, a single positively charged Lys at the α -helix hydrophobic face (peptides W3K5- $\text{LytA}_{239-252}$, K5W10- $\text{LytA}_{239-252}$, and $\text{LytA}_{259-272}$) does not impede α -helix formation. It is possible to speculate that the flexibility of the long Lys side chain allows it to snorkel, making feasible for its aliphatic moiety contribute to the hydrophobic face, and/or for the positively charged amino to be positioned interacting with the negatively charged phosphate at the surface of the DPC micelle. The fact that the substitution of the Trp residues by hydrophobic residues (peptide I3V10- $\text{LytA}_{239-252}$) leads only to a slight destabilisation of the α -helix indicates that the Trp residues contribute to α -helix stability in micelles, but they are not essential.

CONCLUSIONS

1. The short peptides $\text{LytA}_{239-252}$, $\text{LytA}_{259-272}$, and the longer $\text{LytA}_{239-272}$, all derived from the CBD of LytA , undergo a two-state thermal unfolding, and they are able to bind choline in aqueous solution with different affinity and with a stabilizing effect.
2. $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ form stable, well-defined, native-like β -hairpins in aqueous solution, as seen by solution NMR and CD. The longer $\text{LytA}_{239-272}$ peptide also adopts stable, well-defined, native-like β -hairpin structure in the regions 239–252 and 259–272, corresponding to the sequences of the short peptides, but the relative position of the two β -hairpins is not fixed.
3. The aqueous β -hairpin structure of $\text{LytA}_{239-252}$, $\text{LytA}_{259-272}$, and $\text{LytA}_{239-272}$ is conserved in the presence of TFE, a known secondary-structure stabiliser. In the case of $\text{LytA}_{239-272}$, TFE favours some helical tendency in the linker region (residues 253–258).
4. $\text{LytA}_{239-252}$, $\text{LytA}_{259-272}$, $\text{LytA}_{239-272}$, and full-length C- LytA experience a β -to- α structural transition in the presence of either zwitterionic or negatively-charged micelles. For the three peptides, it was demonstrated that the transition is cooperative, takes place between two states, and it is reversible. It was also proved that the only presence of detergent monomers is not enough to trigger the structural transition, the presence of micelles is necessary. Transition is also observed in the presence of SUVs, except for peptide $\text{LytA}_{259-272}$.
5. $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ form well-defined, amphipathic α -helices in the presence of DPC and SDS micelles, as seen by solution NMR and CD. In the case of longer $\text{LytA}_{239-272}$, the helix structure spans the whole sequence, as the central linker also acquires some helical tendency.
6. A model for the interaction of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ with the micelles has been proposed, based on the experimental data (intermolecular NOEs, effect of relaxation agents, and fluorescence experiments). It is characterised by the immersion of the aromatic side chains in the hydrophobic core of the micelles, with a tilted orientation of the peptides. The hydrophobic face of the helices is oriented to the micelle, and the polar face is more exposed to the solvent. The N-terminus is inserted in the micelles, whereas the C-terminus is pointing outside the micelle, and the experimental data suggest a similar orientation for longer $\text{LytA}_{239-272}$. The interaction of $\text{LytA}_{239-252}$ and $\text{LytA}_{259-272}$ with the micelles is a dynamic process in which two equilibria are present: the detergent monomer-micelle equilibrium, and the peptide free-bound equilibrium.
7. DPC micelles are not general helix inducers, since they do not trigger any structural transition in other β -hairpin-forming peptides studied. The β -to- α structural transition is determined by the amphipathicity of the helix, and by the presence of aromatic and hydrophobic side chains able to interact with the micelles.

8. β -hairpin stability of the studied peptides strongly depends on the side chain interactions (i.e., aromatic–aromatic interactions, cation– π interactions). Any modification in the pattern of these interactions may significantly affect the hairpin stability, since the β -turn sequence is not optimal in these peptides.

BIBLIOGRAPHY

- Abel, S. p., Dupradeau, F. o.-Y., & Marchi, M. (2012). Molecular dynamics simulations of a characteristic dpc micelle in water. *Journal of Chemical Theory and Computation*, 8(11), 4610-4623.
- Aemissegger, A., Kräutler, V., van Gunsteren, W. F., & Hilvert, D. (2005). A photoinducible β -hairpin. *Journal of the American Chemical Society*, 127(9), 2929-2936.
- Albrecht, M., & Stortz, P. (2005). Metallacyclopeptides: Artificial analogues of naturally occurring peptides. *Chemical Society Reviews*, 34(6), 496-506.
- Ambroggio, X. I., & Kuhlman, B. (2006). Design of protein conformational switches. *Current Opinion in Structural Biology*, 16(4), 525-530.
- Anfinsen, C. B. (1973). Principles that Govern the Folding of Protein Chains. *Science*, 181(4096), 223-230. doi:10.1126/science.181.4096.223
- Anzini, P., Xu, C., Hughes, S., Magnotti, E., Jiang, T., Hemmingsen, L., . . . Conticello, V. P. (2013). Controlling self-assembly of a peptide-based material via metal-ion induced registry shift. *Journal of the American Chemical Society*, 135(28), 10278-10281.
- Araki, M., & Tamura, A. (2007). Transformation of an α -helix peptide into a β -hairpin induced by addition of a fragment results in creation of a coexisting state. *Proteins: Structure, Function, and Bioinformatics*, 66(4), 860-868.
- Berry, A. M., & Paton, J. C. (2000). Additive Attenuation of Virulence of *Streptococcus pneumoniae* by Mutation of the Genes Encoding Pneumolysin and Other Putative Pneumococcal Virulence Proteins. *Infection and Immunity*, 68(1), 133-140. doi:10.1128/iai.68.1.133-140.2000
- Blanco, F. J., Rivas, G., & Serrano, L. (1994). A short linear peptide that folds into a native stable β -hairpin in aqueous solution. *Nature Structural & Molecular Biology*, 1(9), 584-590.
- Bredenbeck, J., Helbing, J., Kumita, J. R., Woolley, G. A., & Hamm, P. (2005). α -Helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), 2379-2384.
- Bryan, P. N., & Orban, J. (2010). Proteins that switch folds. *Current Opinion in Structural Biology*, 20(4), 482-488. doi:<http://dx.doi.org/10.1016/j.sbi.2010.06.002>
- Buck, M. (1998). Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Quarterly reviews of biophysics*, 31(03), 297-355.
- Cerasoli, E., Sharpe, B. K., & Woolfson, D. N. (2005). ZiCo: A Peptide Designed to Switch Folded State upon Binding Zinc. *Journal of the American Chemical Society*, 127(43), 15008-15009. doi:10.1021/ja0543604
- Cerpa, R., Cohen, F. E., & Kuntz, I. D. (1996). Conformational switching in designed peptides: the helix/sheet transition. *Folding and Design*, 1(2), 91-101.
- Ciani, B., Hutchinson, E. G., Sessions, R. B., & Woolfson, D. N. (2002). A designed system for assessing how sequence affects α to β conformational transitions in proteins. *Journal of Biological Chemistry*, 277(12), 10150-10155.
- Cochran, A. G., Skelton, N. J., & Starovasnik, M. A. (2001). Tryptophan zippers: Stable, monomeric β -hairpins. *Proceedings of the National Academy of Sciences*, 98(10), 5578-5583.
- Chattopadhyay, A., & London, E. (1984). Fluorimetric determination of critical micelle concentration avoiding interference from detergent charge. *Analytical biochemistry*, 139(2), 408-412.
- Cheng, J., Goldstein, R., Gershenson, A., Stec, B., & Roberts, M. F. (2013). The cation- π box is a specific phosphatidylcholine membrane targeting motif. *Journal of Biological Chemistry*, 288(21), 14863-14873.

- Dado, G. P., & Gellman, S. H. (1993). Redox control of secondary structure in a designed peptide. *Journal of the American Chemical Society*, 115(26), 12609-12610.
- de Alba, E., Jiménez, M. A., & Rico, M. (1997). Turn Residue Sequence Determines β -Hairpin Conformation in Designed Peptides. *Journal of the American Chemical Society*, 119(1), 175-183. doi:10.1021/ja962325e
- de Planque, M. R., Kruijtz, J. A., Liskamp, R. M., Marsh, D., Greathouse, D. V., Koeppe, R. E., . . . Killian, J. A. (1999). Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane α -helical peptides. *Journal of Biological Chemistry*, 274(30), 20839-20846.
- Diaz, E., López, R., & Garcia, J. L. (1991). Chimeric pneumococcal cell wall lytic enzymes reveal important physiological and evolutionary traits. *Journal of Biological Chemistry*, 266(9), 5464-5471.
- Díaz, M. D., de la Torre, B. G., Fernández-Reyes, M., Rivas, L., Andreu, D., & Jiménez-Barbero, J. (2011). Structural Framework for the Modulation of the Activity of the Hybrid Antibiotic Peptide Cecropin A-Melittin [CA (1-7) M (2-9)] by N ϵ -Lysine Trimethylation. *Chembiochem*, 12(14), 2177-2183.
- Eldholm, V., Johnsberg, O., Haugen, K., Ohnstad, H. S., & Håvarstein, L. S. (2009). Fratricide in *Streptococcus pneumoniae*: contributions and role of the cell wall hydrolases CbpD, LytA and LytC. *Microbiology*, 155(7), 2223-2234. doi:10.1099/mic.0.026328-0
- Epstein, C. J., Goldberger, R. F., & Anfinsen, C. B. (1963). *The genetic control of tertiary protein structure: studies with model systems*. Paper presented at the Cold Spring Harbor symposia on quantitative biology.
- Fernandez-Tornero, C., Garcia, E., Lopez, R., Gimenez-Gallego, G., & Romero, A. (2002). Two new crystal forms of the choline-binding domain of the major pneumococcal autolysin: insights into the dynamics of the active homodimer. *J Mol Biol*, 321(1), 163-173.
- Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G., & Romero, A. (2001). A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. *Nature structural biology*, 8(12), 1020-1024.
- Fesinmeyer, R. M., Hudson, F. M., Olsen, K. A., White, G. W., Euser, A., & Andersen, N. H. (2005). Chemical shifts provide fold populations and register of β hairpins and β sheets. *Journal of Biomolecular NMR*, 33(4), 213-231.
- Fezoui, Y., Hartley, D. M., Walsh, D. M., Selkoe, D. J., Osterhout, J. J., & Teplow, D. B. (2000). A de novo designed helix-turn-helix peptide forms nontoxic amyloid fibrils. *Nature Structural & Molecular Biology*, 7(12), 1095-1099.
- Fujiwara, K., Toda, H., & Ikeguchi, M. (2012). Dependence of alpha-helical and beta-sheet amino acid propensities on the overall protein fold type. *BMC Structural Biology*, 12(18), 1-15. doi:Artn 18
- 10.1186/1472-6807-12-18
- Galán-Bartual, S., Pérez-Dorado, I., García, P., & Hermoso, J. A. (2015). Chapter 11 - Structure and Function of Choline-Binding Proteins A2 - Brown, Jeremy. In S. Hammerschmidt & C. Orihuela (Eds.), *Streptococcus Pneumoniae* (pp. 207-230). Amsterdam: Academic Press.
- García, E., García, J.-L., Ronda, C., García, P., & López, R. (1985). Cloning and expression of the pneumococcal autolysin gene in *Escherichia coli*. *Molecular and General Genetics MGG*, 201(2), 225-230.
- García, P., Moscoso, M., Rodríguez-Cerrato, V., Yuste, J., & García, E. (2010). *Streptococcus pneumoniae*: from molecular biology to host-pathogen interactions. *Journal of Applied Biomedicine*, 8(3), 131-140. doi:<http://dx.doi.org/10.2478/v10136-009-0016-6>

- Ghadiri, M. R., & Fernholz, A. K. (1990). Peptide architecture. Design of stable .alpha.-helical metallopeptides via a novel exchange-inert ruthenium(III) complex. *Journal of the American Chemical Society*, 112(26), 9633-9635. doi:10.1021/ja00182a030
- Ghozlane, A., Joseph, A., Bornot, A., & De Brevern, A. (2009). Analysis of protein chameleon sequence characteristics. *Bioinformation*, 3(9), 367-369.
- Gilbertson, S. R., Chen, G., & McLoughlin, M. (1994). Versatile building block for the synthesis of phosphine-containing peptides: The sulfide of diphenylphosphinoserine. *Journal of the American Chemical Society*, 116(10), 4481-4482.
- Goddard, T., & Kneller, D. (2004). SPARKY 3. *University of California, San Francisco*, 14, 15.
- Gosink, K. K., Mann, E. R., Guglielmo, C., Tuomanen, E. I., & Masure, H. R. (2000). Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infection and Immunity*, 68(10), 5690-5695.
- Güntert, P. (2004). Automated NMR structure calculation with CYANA. *Protein NMR Techniques*, 353-378.
- Guo, J.-T., Jaromczyk, J. W., & Xu, Y. (2007). Analysis of chameleon sequences and their implications in biological processes. *Proteins: Structure, Function, and Bioinformatics*, 67(3), 548-558. doi:10.1002/prot.21285
- Hakenbeck, R., Madhour, A., Denapate, D., & Brückner, R. (2009). Versatility of choline metabolism and choline-binding proteins in *Streptococcus pneumoniae* and commensal streptococci. *FEMS Microbiology Reviews*, 33(3), 572-586.
- Hammen, P., Gorenstein, D., & Weiner, H. (1996). Amphiphilicity determines binding properties of three mitochondrial presequences to lipid surfaces. *Biochemistry*, 35(12), 3772-3781.
- Haney, E. F., Petersen, A. P., Lau, C. K., Jing, W., Storey, D. G., & Vogel, H. J. (2013). Mechanism of action of puoroindoline derived tryptophan-rich antimicrobial peptides. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(8), 1802-1813.
- Hansen, M. B., Ruizendaal, L., Löwik, D. W. P. M., & van Hest, J. C. M. (2009). Switchable peptides. *Drug Discovery Today: Technologies*, 6(1-4), e33-e39. doi:<http://dx.doi.org/10.1016/j.ddtec.2009.03.002>
- Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., . . . Menéndez, M. (2003). Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure*, 11(10), 1239-1249.
- Hernández-Rocamora, V. M., Maestro, B., Mollá-Morales, A., & Sanz, J. M. (2008). Rational stabilization of the C-LytA affinity tag by protein engineering. *Protein Engineering Design and Selection*, 21(12), 709-720.
- Hernández-Rocamora, V. M., Maestro, B., de Waal, B., Morales, M., García, P., Meijer, E. e. W., . . . Sanz, J. M. (2009). Multivalent Choline Dendrimers as Potent Inhibitors of Pneumococcal Cell-Wall Hydrolysis. *Angewandte Chemie International Edition*, 48(5), 948-951.
- Honda, S., Kobayashi, N., & Munekata, E. (2000). Thermodynamics of a β -hairpin structure: evidence for cooperative formation of folding nucleus1. *J Mol Biol*, 295(2), 269-278. doi:<http://dx.doi.org/10.1006/jmbi.1999.3346>
- Hughes, R. M., & Waters, M. L. (2006). Model systems for β -hairpins and β -sheets. *Current Opinion in Structural Biology*, 16(4), 514-524. doi:<http://dx.doi.org/10.1016/j.sbi.2006.06.008>
- Jedrzejak, M. J. (2006). Unveiling molecular mechanisms of pneumococcal surface protein A interactions with antibodies and lactoferrin. *Clinica chimica acta*, 367(1), 1-10.
- Johnson Jr, W. C. (1988). Secondary structure of proteins through circular dichroism spectroscopy. *Annual review of biophysics and biophysical chemistry*, 17(1), 145-166.

- Kabsch, W., & Sander, C. (1984). On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. *Proceedings of the National Academy of Sciences*, 81(4), 1075-1078.
- Kachel, K., Asuncion-Punzalan, E., & London, E. (1995). Anchoring of Tryptophan and Tyrosine Analogs at the Hydrocarbon-Polar Boundary in Model Membrane Vesicles. *Biochemistry*, 34(47), 15475-15479. doi:10.1021/bi00047a012
- Kammerer, R. A., Kostrewa, D., Zurdo, J., Detken, A., García-Echeverría, C., Green, J. D., . . . Dobson, C. M. (2004). Exploring amyloid formation by a de novo design. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13), 4435-4440.
- Kammerer, R. A., & Steinmetz, M. O. (2006). De novo design of a two-stranded coiled-coil switch peptide. *Journal of structural biology*, 155(2), 146-153.
- Kobe, B., & Kajava, A. V. (2000). When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends in Biochemical Sciences*, 25(10), 509-515.
- Koradi, R., Billeter, M., & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics*, 14(1), 51-55.
- Li, W., Kinch, L. N., Karplus, P. A., & Grishin, N. V. (2015). ChSeq: A database of chameleon sequences. *Protein Science*, 24(7), 1075-1086.
- Li, W., Nicol, F., & Szoka Jr, F. C. (2004). GALA: a designed synthetic pH-responsive amphipathic peptide with applications in drug and gene delivery. *Advanced drug delivery reviews*, 56(7), 967-985. doi:<http://dx.doi.org/10.1016/j.addr.2003.10.041>
- López, R., García, E., García, P., & García, J. L. (1997). The Pneumococcal Cell Wall Degrading Enzymes: A Modular Design to Create New Lysins? *Microbial Drug Resistance*, 3(2), 199-211. doi:10.1089/mdr.1997.3.199
- Lungu, Oana I., Hallett, Ryan A., Choi, Eun J., Aiken, Mary J., Hahn, Klaus M., & Kuhlman, B. (2012). Designing Photoswitchable Peptides Using the AsLOV2 Domain. *Chemistry & Biology*, 19(4), 507-517. doi:<http://dx.doi.org/10.1016/j.chembiol.2012.02.006>
- Luo, P., & Baldwin, R. L. (1997). Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry*, 36(27), 8413-8421.
- Maestro, B., Santiveri, C. M., Jiménez, M. A., & Sanz, J. M. (2011). Structural autonomy of a β -hairpin peptide derived from the pneumococcal choline-binding protein LytA. *Protein Engineering Design and Selection*, 24(1-2), 113-122. doi:10.1093/protein/gzq087
- Maestro, B., & Sanz, J. M. (2005). Accumulation of partly folded states in the equilibrium unfolding of the pneumococcal choline-binding module C-LytA. *Biochem J*, 387(Pt 2), 479-488. doi:10.1042/bj20041194
- Maestro, B., & Sanz, J. M. (2007). Extensive unfolding of the C-LytA choline-binding module by submicellar concentrations of sodium dodecyl sulphate. *FEBS Letters*, 581(3), 375-381. doi:<http://dx.doi.org/10.1016/j.febslet.2006.12.042>
- Maestro, B., Velasco, I., Castillejo, I., Arévalo-Rodríguez, M., Cebolla, Á., & Sanz, J. M. (2008). Affinity partitioning of proteins tagged with choline-binding modules in aqueous two-phase systems. *Journal of Chromatography A*, 1208(1), 189-196.
- Mahajan, M., & Bhattacharjya, S. (2013). β -Hairpin Peptides: Heme Binding, Catalysis, and Structure in Detergent Micelles. *Angewandte Chemie*, 125(25), 6558-6562.
- Mäler, L. (2013). Solution NMR studies of cell-penetrating peptides in model membrane systems. *Advanced drug delivery reviews*, 65(8), 1002-1011.
- Marinelli, P., Castillo, V., & Ventura, S. (2013). Trifluoroethanol Modulates Amyloid Formation by the All α -Helical URN1 FF Domain. *International journal of molecular sciences*, 14(9), 17830-17844.

- Markley, J. L., Bax, A., Arata, Y., Hilbers, C., Kaptein, R., Sykes, B. D., . . . Wüthrich, K. (1998). Recommendations for the presentation of NMR structures of proteins and nucleic acids (IUPAC Recommendations 1998). *Pure and Applied Chemistry*, 70(1), 117-142.
- Medrano, F. J., Gasset, M., López-Zúmel, C., Usobiaga, P., García, J. L., & Menéndez, M. (1996). Structural Characterization of the Unligated and Choline-bound Forms of the Major Pneumococcal Autolysin LytA Amidase: CONFORMATIONAL TRANSITIONS INDUCED BY TEMPERATURE. *Journal of Biological Chemistry*, 271(46), 29152-29161. doi:10.1074/jbc.271.46.29152
- Mellroth, P., Daniels, R., Eberhardt, A., Rönnlund, D., Blom, H., Widengren, J., . . . Henriques-Normark, B. (2012). LytA, major autolysin of *Streptococcus pneumoniae*, requires access to nascent peptidoglycan. *Journal of Biological Chemistry*, 287(14), 11018-11029.
- Mellroth, P., Sandalova, T., Kikhney, A., Vilaplana, F., Heseck, D., Lee, M., . . . Achour, A. (2014). Structural and Functional Insights into Peptidoglycan Access for the Lytic Amidase LytA of *Streptococcus pneumoniae*. *mBio*, 5(1). doi:10.1128/mBio.01120-13
- Mezei, M. (1998). Chameleon sequences in the PDB. *Protein engineering*, 11(6), 411-414.
- Minor, D. L., & Kim, P. S. (1996). Context-dependent secondary structure formation of a designed protein sequence. *Nature*, 380(6576), 730-734.
- Mutter, M., Chandravarkar, A., Boyat, C., Lopez, J., Dos Santos, S., Mandal, B., . . . Saucède, L. (2004). Switch peptides in statu nascendi: induction of conformational transitions relevant to degenerative diseases. *Angewandte Chemie International Edition*, 43(32), 4172-4178.
- Nguyen, P. H., Gorbunov, R. D., & Stock, G. (2006). Photoinduced Conformational Dynamics of a Photoswitchable Peptide: A Nonequilibrium Molecular Dynamics Simulation Study. *Biophysical journal*, 91(4), 1224-1234. doi:<http://dx.doi.org/10.1529/biophysj.106.084996>
- Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci*, 4(11), 2411-2423. doi:10.1002/pro.5560041120
- Pagel, K., & Koksck, B. (2008). Following polypeptide folding and assembly with conformational switches. *Current Opinion in Chemical Biology*, 12(6), 730-739. doi:<http://dx.doi.org/10.1016/j.cbpa.2008.09.005>
- Pandya, M. J., Cerasoli, E., Joseph, A., Stoneman, R. G., Waite, E., & Woolfson, D. N. (2004). Sequence and structural duality: designing peptides to adopt two stable conformations. *Journal of the American Chemical Society*, 126(51), 17016-17024.
- Park, C., & Raines, R. T. (2001). Adjacent cysteine residues as a redox switch. *Protein engineering*, 14(11), 939-942. doi:10.1093/protein/14.11.939
- Pérez-Dorado, I., Galan-Bartual, S., & Hermoso, J. A. (2012). Pneumococcal surface proteins: when the whole is greater than the sum of its parts. *Molecular Oral Microbiology*, 27(4), 221-245. doi:10.1111/j.2041-1014.2012.00655.x
- Ramírez-Alvarado, M., Blanco, F. J., & Serrano, L. (2001). Elongation of the BH8 β -hairpin peptide: Electrostatic interactions in β -hairpin formation and stability. *Protein Science*, 10(7), 1381-1392.
- Rossi, P., Tecilla, P., Baltzer, L., & Scrimin, P. (2004). De novo metallonucleases based on helix-loop-helix motifs. *Chemistry—A European Journal*, 10(17), 4163-4170.
- Rozek, A., Friedrich, C. L., & Hancock, R. E. (2000). Structure of the bovine antimicrobial peptide indolicidin bound to dodecylphosphocholine and sodium dodecyl sulfate micelles. *Biochemistry*, 39(51), 15765-15774.

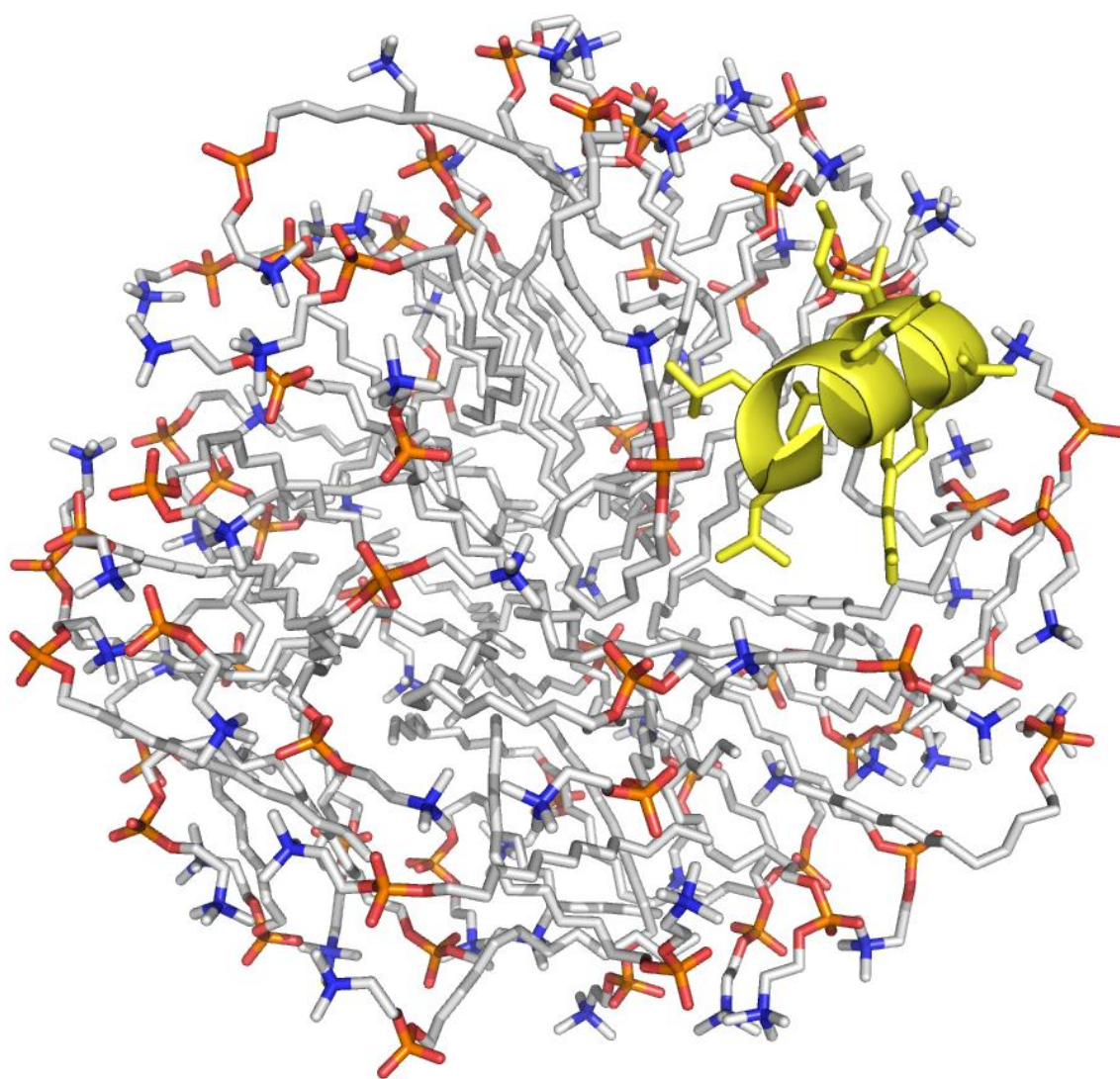
- Russell, S. J., Blandl, T., Skelton, N. J., & Cochran, A. G. (2003). Stability of cyclic β -hairpins: Asymmetric contributions from side chains of a hydrogen-bonded cross-strand residue pair. *Journal of the American Chemical Society*, 125(2), 388-395.
- Russell, S. J., & Cochran, A. G. (2000). Designing stable β -hairpins: Energetic contributions from cross-strand residues. *Journal of the American Chemical Society*, 122(50), 12600-12601.
- Sánchez-Puelles, J., Sanz, J. M., Garcia, J., & Garcia, E. (1990). Cloning and expression of gene fragments encoding the choline-binding domain of pneumococcal murein hydrolases. *Gene*, 89(1), 69-75. doi:[http://dx.doi.org/10.1016/0378-1119\(90\)90207-8](http://dx.doi.org/10.1016/0378-1119(90)90207-8)
- Sanchez-Puelles, J. M., Sanz, J. M., Garcia, J. L., & Garcia, E. (1992). Immobilization and single-step purification of fusion proteins using DEAE-cellulose. *European Journal of Biochemistry*, 203(1-2), 153-159. doi:10.1111/j.1432-1033.1992.tb19840.x
- Sanders, C. R., & Sönnichsen, F. (2006). Solution NMR of membrane proteins: practice and challenges. *Magnetic Resonance in Chemistry*, 44(S1), S24-S40.
- Santiveri, C. M., & Jiménez, M. A. (2010). Tryptophan residues: Scarce in proteins but strong stabilizers of β -hairpin peptides. *Peptide Science*, 94(6), 779-790. doi:10.1002/bip.21436
- Santiveri, C. M., Pantoja-Uceda, D., Rico, M., & Jimenez, M. (2005). β -Hairpin formation in aqueous solution and in the presence of trifluoroethanol: A ^1H and ^{13}C nuclear magnetic resonance conformational study of designed peptides. *Biopolymers*, 79(3), 150-162.
- Santiveri, C. M., Rico, M., & Jiménez, M. (2000). Position effect of cross-strand side-chain interactions on [beta]-hairpin formation. *Protein Science*, 9(11), 2151-2160.
- Santiveri, C. M., Rico, M., & Jiménez, M. A. (2001). $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ chemical shifts as a tool to delineate β -hairpin structures in peptides. *Journal of Biomolecular NMR*, 19(4), 331-345.
- Santiveri, C. M., Santoro, J., Rico, M., & Jiménez, M. (2004). Factors involved in the stability of isolated β -sheets: Turn sequence, β -sheet twisting, and hydrophobic surface burial. *Protein Science*, 13(4), 1134-1147.
- Santiveri, C. M., Santoro, J., Rico, M., & Jiménez, M. A. (2002). Thermodynamic Analysis of β -Hairpin-Forming Peptides from the Thermal Dependence of ^1H NMR Chemical Shifts. *Journal of the American Chemical Society*, 124(50), 14903-14909. doi:10.1021/ja0278537
- Sanz, J. M., Jiménez, M. A., & Giménez-Gallego, G. (2002). Hints of nonhierarchical folding of acidic fibroblast growth factor. *Biochemistry*, 41(6), 1923-1933.
- Saucède, L., Santos, S. D., Chandravarkar, A., Mandal, B., Mimna, R., Murat, K., . . . Adrian, M. (2006). Switch-peptides: From conformational studies to Alzheimer's disease. *CHIMIA International Journal for Chemistry*, 60(4), 199-202.
- Schneider, J. P., Pochan, D. J., Ozbas, B., Rajagopal, K., Pakstis, L., & Kretsinger, J. (2002). Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide. *Journal of the American Chemical Society*, 124(50), 15030-15037.
- Schrodinger, LLC. (2015). *The PyMOL Molecular Graphics System, Version 1.8*.
- Shen, Y., Delaglio, F., Cornilescu, G., & Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR*, 44(4), 213-223.
- Shenkarev, Z. O., Balandin, S. V., Trunov, K. I., Paramonov, A. S., Sukhanov, S. V., Barsukov, L. I., . . . Ovchinnikova, T. V. (2011). Molecular mechanism of action of β -hairpin antimicrobial peptide arenicin: oligomeric structure in dodecylphosphocholine micelles and pore formation in planar lipid bilayers. *Biochemistry*, 50(28), 6255-6265.

- Signarvic, R. S., & DeGrado, W. F. (2009). Metal-binding dependent disruption of membranes by designed helices. *Journal of the American Chemical Society*, 131(9), 3377-3384.
- Smith, S. J., Du, K., Radford, R. J., & Tezcan, F. A. (2013). Functional, metal-based crosslinkers for α -helix induction in short peptides. *Chemical Science*, 4(9), 3740-3747.
- Sohma, Y., Hayashi, Y., Kimura, M., Chiyomori, Y., Taniguchi, A., Sasaki, M., . . . Kiso, Y. (2005). The 'O-acyl isopeptide method' for the synthesis of difficult sequence-containing peptides: application to the synthesis of Alzheimer's disease-related amyloid β peptide (A β) 1–42. *Journal of Peptide Science*, 11(8), 441-451.
- Stanger, H. E., & Gellman, S. H. (1998). Rules for antiparallel beta-sheet design: D-Pro-Gly is superior to L-Asn-Gly for beta-hairpin nucleation. *Journal of the American Chemical Society*, 120(17), 4236-4237. doi:10.1021/ja973704q
- Stern, O., & Volmer, M. (1919). On the quenching-time of fluorescence. *Phys. Z*, 20, 183-188.
- Takano, K., Katagiri, Y., Mukaiyama, A., Chon, H., Matsumura, H., Koga, Y., & Kanaya, S. (2007). Conformational contagion in a protein: Structural properties of a chameleon sequence. *Proteins: Structure, Function, and Bioinformatics*, 68(3), 617-625. doi:10.1002/prot.21451
- Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., . . . Dodson, R. J. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science*, 293(5529), 498-506.
- Usachev, K., Efimov, S., Kolosova, O., Filippov, A., & Klochkov, V. (2014). High-resolution NMR structure of the antimicrobial peptide protegrin-2 in the presence of DPC micelles. *Journal of Biomolecular NMR*, 1-8.
- Vallée-Bélisle, A., & Plaxco, K. W. (2010). Structure-switching biosensors: inspired by Nature. *Current Opinion in Structural Biology*, 20(4), 518-526. doi:<http://dx.doi.org/10.1016/j.sbi.2010.05.001>
- Vila, R., Ponte, I., Suau, P., & Rico, M. (2000). A helix-turn motif in the C-terminal domain of histone H1. *Protein Science*, 9(4), 627-636.
- Wang, X., Bergenfeld, I., Arora, P. S., & Canary, J. W. (2012). Reversible Redox Reconfiguration of Secondary Structures in a Designed Peptide. *Angewandte Chemie*, 124(48), 12265-12267.
- Wimley, W. C., & White, S. H. (1993). Membrane partitioning: distinguishing bilayer effects from the hydrophobic effect. *Biochemistry*, 32(25), 6307-6312.
- Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., & Sykes, B. D. (1995). ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR*, 5(1), 67-81.
- Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol*, 222(2), 311-333.
- Wuthrich, K. (1986). *NMR of proteins and nucleic acids*: Wiley.
- Wüthrich, K., Billeter, M., & Braun, W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J Mol Biol*, 180(3), 715-740.
- Xu, Y., Oyola, R., & Gai, F. (2003). Infrared Study of the Stability and Folding Kinetics of a 15-Residue β -Hairpin. *Journal of the American Chemical Society*, 125(50), 15388-15394. doi:10.1021/ja037053b
- Yau, W.-M., Wimley, W. C., Gawrisch, K., & White, S. H. (1998). The Preference of Tryptophan for Membrane Interfaces. *Biochemistry*, 37(42), 14713-14718. doi:10.1021/bi980809c

- Zamora-Carreras, H., Maestro, B., Strandberg, E., Ulrich, A. S., Sanz, J. M., & Jiménez, M. Á. (2015). Micelle-Triggered β -Hairpin to α -Helix Transition in a 14-Residue Peptide from a Choline-Binding Repeat of the Pneumococcal Autolysin LytA. *Chemistry – A European Journal*, n/a-n/a. doi:10.1002/chem.201500447
- Zhang, S., & Rich, A. (1997). Direct conversion of an oligopeptide from a β -sheet to an α -helix: a model for amyloid formation. *Proceedings of the National Academy of Sciences*, 94(1), 23-28.

CHAPTER 3

MEMBRANE-ACTIVE PEPTIDES (MAPs)



This chapter is divided into two different parts. The first part, corresponding to the study of BP100 peptide is the result of my work during a short stay in the group of Prof. A. S. Ulrich, under the kind supervision of Dr. E. Strandberg, from the Institute for Biological Interfaces (Karlsruhe Institute for Technology, Karlsruhe, Germany). In the second part of this chapter, I have used solution NMR to assess the structure–function relationship of crotalictidin and its fragments, in order to complete the extensive biological investigation developed by the group of Prof. D. Andreu (Universitat Pompeu Fabra, Barcelona, Spain) on this peptide.

GENERAL INTRODUCTION

1. MEMBRANE-ACTIVE PEPTIDES (MAPs)

Membrane-active peptides (MAPs) constitute a major class amongst bioactive peptides. MAPs show a vast diversity in their structure and function but they can be defined as natural or synthetic peptides, of typically 5–60 amino acids in length, which possess the ability to interact with the lipid membrane or disrupt it upon the interaction with its components and, thereby, execute their membrane-related activity. (Ponnappan, Budagavi, Yadav, & Chugh, 2015) They are usually unstructured in solution but undergo a disorder-to-order transition upon membrane association which promotes either α -helical or β -strand secondary structural states which are generally amphipathic. (Last, Schlamadinger, & Miranker, 2013)

Natural occurring MAPs can be found in virtually every organism. Besides, lots of synthetic MAPs have been developed in the last decades, most of them starting from natural sequences that are modified in order to achieve better properties. (Manuel N Melo, Ferre, & Castanho, 2009) According to their biological functions, MAPs can be classified as antimicrobial (AMPs), antiviral (AVPs), antifungal, antiparasitic (APPs), anticancer (ACPs), immunomodulatory, cell-penetrating (CPPs), fusogenic (FPs), and amyloid peptides. Due to the similarity of the mechanisms that are behind these different biological functions, some MAPs can show more than one type of activity. (Wadhwani, Reichert, Burck, & Ulrich, 2012) Considering the potential biological functions which MAPs can perform, the clinical relevance of their study, comprehension and development is evident.

2. ANTIMICROBIAL PEPTIDES (AMPs)

Antimicrobial peptides (AMPs) are MAPs capable to interact with microbial membranes. In spite of the fact that they share some common characteristics (such as a short length, the presence of a high number of positive and hydrophobic residues, and a significant amphipathicity), their sequences vary enormously from one to another. Most AMPs, in common with general MAPs, undergo a transition from disorder to order upon encountering a membrane. (Last et al., 2013; Li, Xiang, Zhang, Huang, & Su, 2012; Manuel N Melo et al., 2009)

Natural AMPs are an essential part of the innate host defence system against infections. These peptides constitute an evolutionarily ancient defensive weapon, but remain very effective, as they are far less prone to trigger bacterial resistance than conventional antibiotics. AMPs are often toxic to a broad spectrum of bacteria, but relatively innocuous toward host eukaryotic cells at bactericidal concentrations. (Last et al., 2013; Sato & Feix, 2006; Seo, Won, Kim, Mishig-Ochir, & Lee, 2012; Zasloff, 2002)

2.1 CLASSIFICATION AND ORIGIN OF THE AMPs

AMPs show a huge diversity and different ways of sorting them have been proposed. Generally, they are classified by their amino acid composition and structure: (Figure 3.1) (Brogden, 2005; Ponnappan et al., 2015)

- Anionic peptides (Class I)

Small peptides (~700–800 Da) that require zinc as a cofactor for antimicrobial activity. They are active against both Gram-positive and Gram-negative bacteria. Examples: maximin H5 (amphibians), dermcidin (humans).

- Linear cationic α -helical peptides (Class II)

Short peptides (< 40 amino acids) with spatially segregated hydrophobic and hydrophilic regions in a linear structure. Some of them occur in the form of homodimer or heterodimer. They lack cysteine and sometimes present a hinge or kink in the middle. They are usually unstructured in aqueous solution and become helical in TFE, micelles, vesicles or liposomes. Active against both Gram-positive and Gram-negative bacteria. Activity is normally correlated with the α -helix content. Examples: cecropins (A) (insects), LL37 (humans), BP100.

- Cationic peptides enriched with specific amino acids (Class III)

Their sequences have over-representation of one type of amino acid (usually Pro, His, Trp, Arg or Gly), which seems to be essential for their antimicrobial activity. They usually adopt random coil conformations, but sometimes form extended coils. Examples: prophenin (pigs), indolicidin (cattle).

- Anionic and cationic peptides that contain cysteine and form disulphide bonds (Class IV)

Cysteine-rich peptides which contain β -sheet structure and disulphide bonds. They can contain one disulphide bond (e.g., brevinins), two (e.g., tachyplesins), three (e.g., human α - and β -defensins) or more (e.g., drosomycin).

- Anionic and cationic peptides that are fragments of larger proteins (Class V)

Similar in composition and structure to those described above, but their role in innate immunity is not yet clear. Examples: lactoferricin, casocidin I (humans).

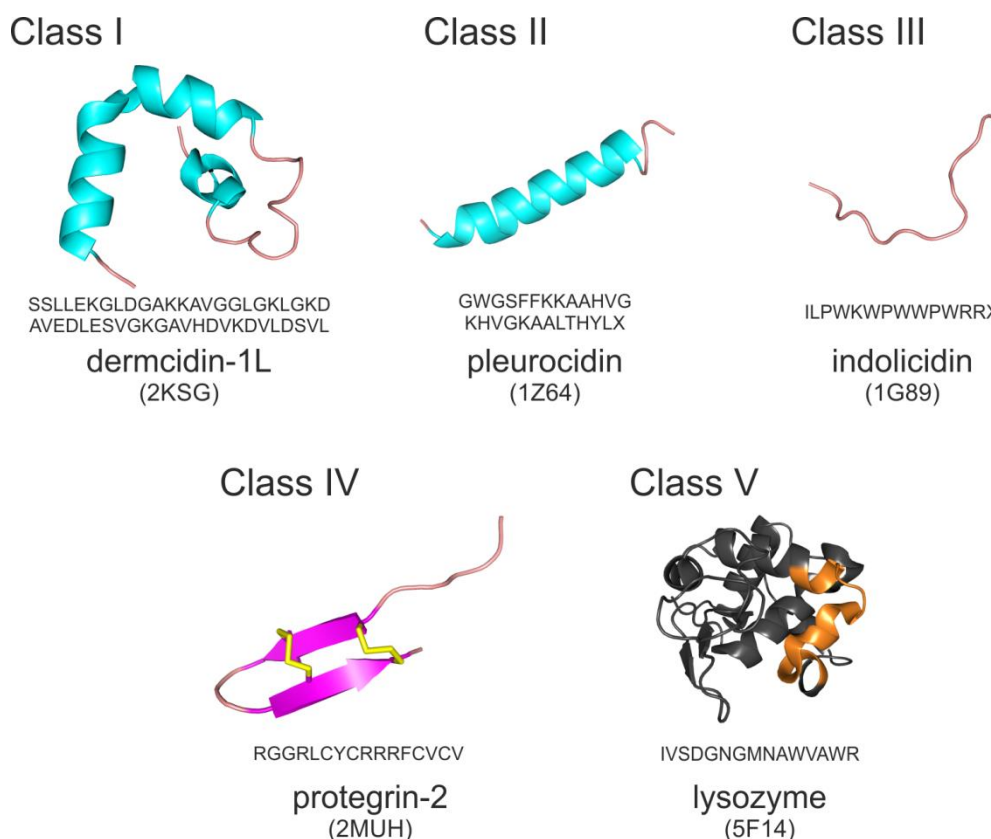


Figure 3.1. Examples of the AMPs classes by sequence and structure. Sequence, name and PDB code are indicated for each structure. Secondary structure elements are coloured cyan for α -helices, magenta for β -strands and salmon for loops, except in class V example in which the antimicrobial region of the whole protein is highlighted in orange (and the sequence indicated corresponds to this highlighted region). In the class IV example, disulphide bonds are indicated in yellow.

AMPs can be also classified depending on their origin in two main groups: natural and synthetic AMPs. Natural AMPs has been described in all kinds of living organisms and even in viruses.

Arg-rich peptides with antimicrobial and cytolytic activities have been found in some lentivirus, such as HIV. AMPs are also secreted by both Gram-positive and Gram-negative bacteria. These bacterial AMPs are grouped as bacteriocins and their function is to kill bacterial competitors without causing significant harm to the host cell, thanks to posttranscriptional modifications or specific immunity mechanisms. It is thought these bacterial AMPs can also induce the permeabilization of the target cell membranes. (R. E. W. Hancock & Chapple, 1999)

In plants, AMPs with additional antifungal or cytolytic activities can be found. In fact, defensive peptides in plants are mostly oriented to kill fungal organisms, as they are their main pathogens. In insects, AMPs play an important role in humoral defence reactions and they are synthesized during systemic response against pathogens. AMPs are also present in insect venoms and, in this case, they usually show cytotoxic activity. The organism of some insects, e.g. *Drosophila*, is able to discriminate between different types of

pathogens and produce the appropriate AMP to fight the infection. (R. E. W. Hancock & Chapple, 1999; Li et al., 2012)

AMPs are not exclusive of microorganisms and invertebrates, as they have been also described in vertebrates. AMPs from vertebrates show diverse sequences, structures and target specificity. AMPs are present in many tissues and cell types of fish, amphibians and mammals, such as neutrophils, mucosal or skin secretions from epithelial cells, platelets, Paneth cells, and leukocytes. A well-known example of animal AMPs is the cathelicidins family, and one of its members, crotalicidin, has been studied in the second part of the present chapter. (Li et al., 2012)

AMPs can be of synthetic origin. Synthetic AMPs are designed by systematic variation of naturally occurring peptides in order to improve their antimicrobial activity and reduce the cytotoxic effects. The goal of these modifications is to obtain new AMPs with good therapeutic properties to find new alternatives to the classic antibiotics which have developed bacterial resistance. One example of this kind of peptides is BP100. (Badosa et al., 2007; R. E. W. Hancock & Chapple, 1999)

2.2 AMPs CELL-TARGETING AND MECHANISMS OF ACTION

Bacterial killing by AMPs is a process which must imply several specific steps. Firstly, AMPs must be attracted to the bacterial surface and this event takes place mediated by electrostatic interactions between the charged peptide and components present in the cell surface. Then, AMPs must traverse the bacterial polysaccharide layer to interact with the cell membrane. Once the peptide is in close contact with the outer (in Gram-negative bacteria) or the cytoplasmic (in Gram-positive bacteria) membrane (Figure 3.2), the process of insertion can happen. It is proposed that the initial binding of AMPs to the membrane displaces divalent cations, destabilizing the cell envelope and resulting in the subsequent uptake of the peptide. (Brogden, 2005; Sato & Feix, 2006)

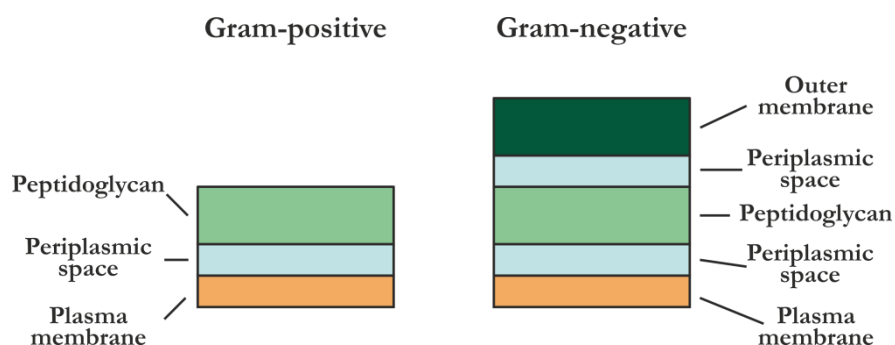


Figure 3.2. Schematic representation of the cell membrane structure of Gram-positive and Gram-negative bacteria.

Most AMPs act through a non-receptor-mediated mechanism which implies the unspecific electrostatic interactions between the peptide and the membrane. A very

common property of the AMPs is their net positive charge, which allows a preferential binding to negatively charged components present in the bacterial cell surface, such as lipopolysaccharides (LPS) and teichoic acids contained in the envelope of Gram-negative and Gram-positive bacteria, respectively, or the negatively charged phospholipids of the cytoplasmic membrane. However, as explained before, there are AMPs with net negative charge that may use a different binding mechanism. Some AMPs expressed by bacteria contain receptor-binding domains and are able to bind specifically to concrete components of the cell membrane. This is the case of nisin Z, an AMP which uses the membrane-anchored cell wall precursor Lipid II as receptor. (Shai, 2002)

Once an AMP is interacting with the target cell membrane, cell destruction can be achieved either by membrane disruption or by intracellular killing processes. The latter comprises a variety of modes of action that includes flocculation of intracellular contents, alteration of cytoplasmic membrane septum formation, inhibition of cell-wall synthesis, binding of nucleic acids, inhibition of nucleic-acid synthesis, inhibition of protein synthesis and inhibition of enzymatic activity. (Brogden, 2005)

The antimicrobial activity of AMPs based on the capacity of membrane disruption has been studied from a mechanistic point of view and three mechanism models have been proposed so far: (Ponnappan et al., 2015)

- Carpet model

Peptides accumulate on the bilayer with a parallel orientation respect to the bilayer plane forming a carpet-like structure on the membrane. As commented before, electrostatic interaction between the cationic peptide and the negatively charged membrane phospholipids acts as the initial driving force for membrane perturbation. Once a threshold peptide concentration is reached, the membrane becomes permeable, breaks and cell lysis takes place in a detergent-like manner with the formation of micelles. (Figure 3.3)

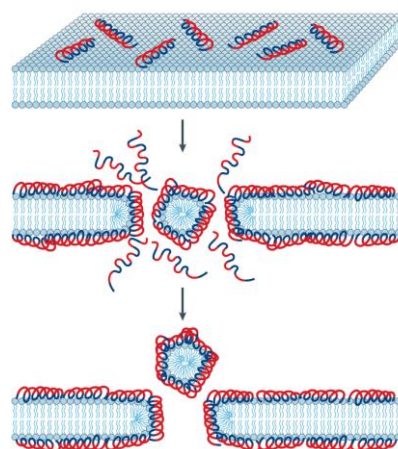


Figure 3.3. Carpet model of AMPs. Peptides interact with the membrane in a parallel orientation forming a carpet. Hydrophobic region of the peptides (blue) orient to the membrane core and the hydrophilic face (red) is exposed to the solvent. Membrane is destabilized and disrupted forming micelles. (Brogden, 2005)

- Barrel-stave model

Helical AMPs usually have an amphipathic structure with separated hydrophilic and hydrophobic regions. This amphipathic structure permits the insertion of the peptide into the cell membrane. According to the barrel-stave model, peptides bind to the membrane, recognize each other and oligomerize. The oligomer inserts into the hydrophobic core of the membrane, forming a transmembrane pore. Upon oligomerization, AMPs orient themselves, allowing the hydrophobic surface to interact with the hydrophobic core of the membrane and hydrophilic surface to point inward to create a hydrophilic pore. (Figure 3.4) Oligomerization of these peptides can occur on the membrane surface or in hydrophobic core of the membrane. To follow this model mechanism, MAPs are required to have net charge close to neutral and amphipathic structure.

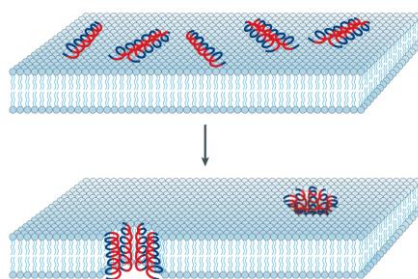


Figure 3.4. Barrel-stave model of AMPs. The attached peptides aggregate and insert into the membrane bilayer. The hydrophobic peptide regions (blue) align with the lipid core region and the hydrophilic regions (red) form the inner coating of the pore. (Brogden, 2005)

- Toroidal-pore model

In this model, peptides are always associated with the lipid head groups even when they are perpendicularly inserted in the lipid bilayer. Moreover, the phospholipid monolayer bends continuously from top to bottom of the bilayer and the toroidal pore is lined by both peptides and phospholipid head groups. (Figure 3.5)

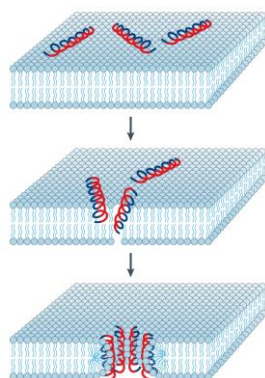


Figure 3.5. Toroidal-pore model of AMPs. The attached peptides aggregate and induce the lipid monolayers to bend continuously through the pore. The water core is lined by both the inserted peptides and the lipid head groups. Hydrophobic peptide regions are colored in blue and hydrophilic regions in red. (Brogden, 2005)

Some AMPs show mechanisms that have been categorized in more than one model. Differences in experimental parameters, such as peptide concentration, peptide-to-lipid ratio (P:L), and membrane composition may account for the variety of behaviours observed for a single peptide. (Last et al., 2013)

2.3 STRUCTURE–FUNCTION RELATIONSHIPS (SAR) OF AMPs

Understanding the structure–activity relationships of AMPs is essential for the design and development of novel antimicrobial agents with improved properties. In spite of the fact that the mechanisms of action of some AMPs are not precisely defined, the main factors that contribute to an increase in the activity and selectivity towards bacteria have been identified. (Manuel N Melo et al., 2009; Seo et al., 2012) The factors relative to the peptide are:

- Size: AMPs described so far vary from very short sequences (6 amino acids) to peptides longer than 59 residues. AMPs forming dimers and trimers have been also reported.
- Sequence: AMPs are often enriched with basic (such as Lys or Arg) and hydrophobic amino acids (such as Ala, Leu, Trp, Phe, Tyr, Ile and Val). Percentage of hydrophobic residues in the sequence usually ranges between 50–67 %.
- Charge: anionic AMPs contain a high number of Asp and Glu residues, whereas cationic AMPs are rich in Lys and Arg. Anionic peptides complexed with zinc or highly cationic peptides are often more active than neutral peptides or those with lower net charge.
- Conformation and structure: AMPs can assume a variety of secondary structures (α -helices, relaxed coils and antiparallel β -sheets). Peptides with amphipathic α -helical structure are often more active than those with less-defined secondary structures. Peptides with a γ -core motif (two antiparallel β -sheets with an interposed short turn in defensin-like molecules) are usually very active.
- Hydrophobicity: the high presence of hydrophobic residues allows AMPs to partition into the membrane bilayer, and it is correlated with increased levels of haemolysis.
- Amphipathicity: this property arises from the segregation of hydrophilic and hydrophobic amino acids in opposite regions of a helical peptide. In helices, amphipathicity is often expressed as the hydrophobic moment, which is the vector sum of hydrophobic indices, treated as vectors normal to the helical axis. Amphipathicity leads to higher levels of peptide internalization and membrane perturbation.

Some factors relative to the bacterial target cell that contribute to increase the activity and specificity of AMPs are: (Manuel N Melo et al., 2009)

- Hydrophobicity and charge: cationic AMPs tend to bind the bacterial membranes because they contain a large proportion of negatively-charged phospholipids, whereas their activity against other types of cells, such as mammalian cells, is lower

because their membranes are composed by essentially neutral phospholipids and sterols.

- Transmembrane potential: AMP interaction is promoted by the usually higher, negative –inside transmembrane potential found in bacteria.
- AMP ligands: some AMPs show specific structural affinity for bacterial membrane components, such as phospholipid head groups.

2.4 ANTIBIOTIC RESISTANCE AND CLINICAL IMPORTANCE OF AMPs

Bacteria are exceptionally adaptable organisms that possess a big capacity for surviving under adverse conditions. Antimicrobial resistance is one of the bacterial survival mechanisms and it is an inevitable evolutionary response to antibiotics use. However, the abuse of antibacterial drugs has sped up the development of new resistances. ((WHO), 2014; Brusselaers, Vogelaers, & Blot, 2011; French, 2005)

Consequences of antibiotic resistance are matter of concern since infections caused by antibiotic–resistant microorganisms do not respond to the treatments, provoking the prolongation of the illness and increasing the risk of death. Furthermore, the failure in the treatment entails long periods of drug ineffectiveness that augments the probability of spread the contagion with resistant strains. When infections become first–line–antibiotic–resistant, treatment must be modified using other kind of antimicrobial drugs which are usually more expensive and more toxic.

According to WHO data, in 2012 at least 7.2 % of annual world deaths (this is over 4 million people) were due to diseases potentially treatable with antibiotics. But if we focus on the African region the percentage rises at least to 15.5 % (over 1.4 million people only in Africa). ((WHO), 2012) Regarding these data, it is evident the importance of the emergence of new resistances and the loss of effectiveness of antibiotics, and thus, the development of new alternative antibacterial drugs.

Therefore, AMPs have been interesting targets as novel antibiotics in the last decades due to their broad–spectrum activity, which include drug–resistance bacterial strains. However, natural AMPs have some limitations for drug development, since they are labile and their stability depends greatly on the surrounding conditions (pH, presence of proteases...). In addition, natural AMPs usually show low stability for oral administration and the production costs are high. For these reasons, different approaches have been followed for AMPs development, such as: (Seo et al., 2012)

- Utilization of non–natural amino acids (such as D–amino acids) or modifications in the terminal regions of the peptide in order to make them more resistant to degradation.

- Modification of the AMPs sequences to improve their therapeutic index (higher antimicrobial activities and lower cytotoxic/haemolytic effects).
- Use of efficient delivery systems, such as liposome encapsulation, to reduce degradation and toxic effects.
- Search for minimal antimicrobial sequences to reduce production costs.

AMPs show a relatively low possibility of resistance emergence, but all the same they are not exempt from it. Amongst the resistance strategies developed by bacteria to circumvent antibiotic activity, some of them are directly oriented to AMPs, such as those that affect the attachment, insertion or permeabilization ability of the AMPs. The mechanisms of resistance against AMPs include: (Seo et al., 2012; Zasloff, 2002)

- Alteration of net surface charges: some bacteria transport basic or acid molecules to their surfaces in order to make the net surface charge less negative or positive. This way the electrostatic interaction between AMPs and the cell surface is weakened.
- Modification of the membrane fluidity: *Salmonella* species reduce the fluidity of their outer membrane by increasing hydrophobic interactions with the addition of acyl tails to the membrane component lipid A.
- Changes in membrane proteins: a correlation between the alteration in the production of outer membrane proteins and resistance to killing by AMPs has been observed in some Gram-negative bacteria.
- Role of transporters: some membrane transporters have been associated with AMP resistance, as they are sometimes involved in AMP import/export processes.
- Proteolytic enzymes: some of the proteolytic enzymes produced by bacteria are able to degrade AMPs.

Despite being promising therapeutic agents, no AMPs have been approved as drugs so far. Some AMPs have reached the clinical trial phase, such as Pexiganan, Plectasin and Brilacidin, and it is probable that in the near future some of them will be commercially available. (Zasloff, 2016)

Amongst the AMPs that are under development for clinical uses, some of them are oriented to the direct treatment of bacterial infections. However, this is not the only use proposed for AMPs. In this regard, immobilization of AMPs onto solid supports has been proposed to improve biomedical devices, drug delivery systems, or bio-sensors, and avoid the formation of biofilms. AMPs attached to nanoparticles have been studied as site-specific targeting and delivering drugs to be used in the treatment of a variety of diseases, including cancer. AMPs can also serve as indicators and diagnostic agents as they can be immobilized in chips to detect bacterial pathogens. (Fox, 2013; Wang et al., 2015)

3. ANTICANCER PEPTIDES (ACPs)

Broadly speaking, ACPs are peptides with a specific cytotoxic activity against cancerous cells. However, a variety of specificities can be found in these peptides: some of them show a general toxicity to eukaryotic cells, killing cancer and non–cancer alike, whilst others exhibit toxicity to cancer cells alone, and even some ACPs are selective for specific types of cancerous cells. (Dennison, Harris, Bhatt, Singh, & Phoenix, 2009; Gaspar, Veiga, & Castanho, 2014; Li et al., 2012)

The factors contributing to the diverse efficiency and specificity of ACPs are not well understood yet, being this a crucial question to develop therapeutic ACPs in a rational manner. All the same, it is known that the characteristics of the target cell membrane and some physicochemical properties of the ACPs affect their activity. (Dennison et al., 2009)

3.1 FACTORS CONTRIBUTING TO ACPs SPECIFICITY

As mentioned before, some factors effecting ACPs specificity are related to the target cell membrane properties, such as: (Dennison et al., 2009; Gaspar et al., 2014; Hoskin & Ramamoorthy, 2008)

- Membrane components: this is thought to be a major factor in the selectivity of ACPs. Cancerous cells typically possess an increased expression of anionic molecules (such as phosphatidylserine (PS), O–glycosylated mucins, sialylated gangliosides or heparin sulphate) that provides a net negative charge to the membrane. This net negative charge favours electrostatic interactions between ACPs and membrane, what does not occur in healthy cells because of the overall neutral charge conferred by their zwitterionic major membrane components.
- Membrane fluidity: cancerous cells usually have decreased levels of cholesterol and thus membrane fluidity is greater than that of healthy cells, which facilitates membrane destabilization. Cholesterol is a major component of non–cancer eukaryotic cell membranes and it may protect them from the cytotoxic effect of ACPs by making more difficult the peptide insertion.
- Cell surface area: ACPs binding is favoured in cancerous cells due to the augmentation of the cell surface area, as a consequence of the presence of higher numbers of microvilli (tiny projections of the cell membrane).
- Membrane glycosylation pattern: it has been suggested that changes in the membrane glycosylation pattern suffered by healthy cells when undergo cancerous transformation could affect the interaction and activity of ACPs on the cells.

On the other hand, the peptide properties that facilitate the ACPs interaction with membranes are analogous to those described for AMPs:

- Net charge: net positive charge of the peptide favours the electrostatic interaction with the net negative charged membrane of cancerous cells.
- Hydrophobicity and amphipathicity: the presence of hydrophobic amino acids in AMPs sequences contributes to the generation of an amphipathic structure which is able to interact with and insert more efficiently in membranes.

3.3 STRUCTURE AND MODES OF ACTION OF ACPs

The adoption of a defined structure has been described as a crucial feature for the activity of many ACPs. In this regard, α -helical, β -sheet and linear ACPs have been reported so far. (Hoskin & Ramamoorthy, 2008)

The modes of action of ACPs can be divided in two main groups: membranolytic and non-membranolytic modes of action. Membranolytic modes of action are based in plasma and/or mitochondrial membrane disruption. The overexpression of anionic molecules in cancerous membranes triggers the interaction between ACPs and cell membranes. After engagement of ACPs in the cell membrane, the peptides can cause membrane destabilization accompanied by pore formation and/or changes on the cell membrane charge. The insertion of bulky hydrophobic amino acids on the cell membrane hydrophobic core with the acquisition of a stable structure can drive the pore formation. ACPs can follow the classic pore formation models described for AMPs: carpet model, toroidal pore model or barrel-stave model. (Gaspar et al., 2014; Hoskin & Ramamoorthy, 2008; Mulder, Lima, Miranda, Dias, & Franco, 2013) (Figures 3.3, 3.4, and 3.5)

Non-membranolytic modes of action include: (Gaspar et al., 2014; Mulder et al., 2013)(Figure 3.6)

- Activation of immune modulatory pathway by induction of natural killer (NK) cells and interferon (IFN).
- Alteration of the lysosome membrane leading to an acidification of the intracellular environment and cell death.
- Amplification of the proteasome activity.
- Induction of mitochondrial pathway of apoptosis by either the cytochrome c release into the cytoplasm or activation of the caspase cascade.
- Increase of the influx of Ca^{2+} .
- Inhibition of genes involved in DNA replication.

- Arrest of cell cycle G0, G1, or S phases.

Some ACPs combine more than one of these mechanisms.

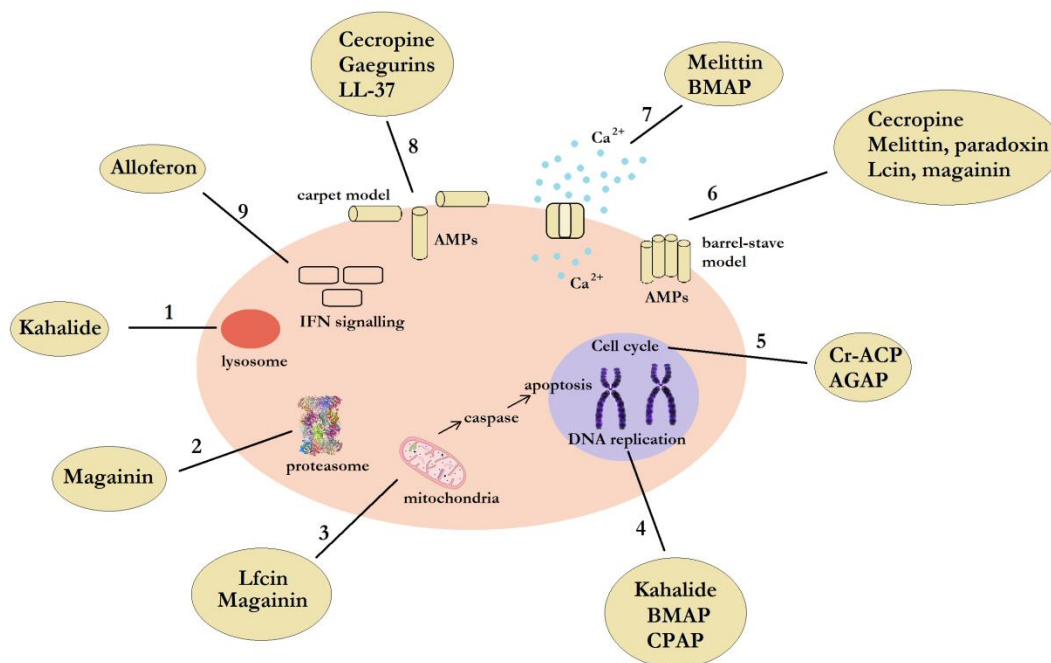


Figure 3.6. Mechanisms of action of ACPs. **1.** bModification of the lysosome membrane, and subsequent acidification of intracellular environment. **2.** Amplification of the proteasome activity. **3.** Induction of the mitochondrial pathway of apoptosis. **4.** Inhibition of genes involved in DNA replication. **5.** Arrest of cell cycle, G0, G1, or S phases. **6.** Formation of pores by either toroidal or barrel-stave models. **7.** Increase of the influx of Ca^{2+} . **8.** Pore formation by carpet model. **9.** Activation of an immune modulatory pathway by induction of NK and IFN. Figure adapted from (Mulder et al., 2013)

3.4 DEVELOPMENT OF THERAPEUTIC ACPs

Cancer treatments imply harmful side effects, as the currently used antineoplastic drugs attack all rapidly dividing cells, rather than solely cancerous cells. In this regard, ACPs are promising molecules for cancer treatment since many of them seem to show specificity towards malignant cells. (Gaspar et al., 2014; Hoskin & Ramamoorthy, 2008)

The challenge in the design of therapeutic ACPs lies on achieving a very high specificity and killing ability with low cytotoxic activity to healthy mammalian cells, and improving the resistance to proteolytic cleavage. In order to get better therapeutic properties and overcome their weaknesses, ACPs sequences are manipulated altering their physicochemical and biochemical characteristics, such as net charge, secondary structure, oligomerization ability, hydrophobicity, or amphipathicity. For instance, higher serum stability has been obtained by introducing D-amino acids in ACP sequences. Other strategies have been proposed, such as the co-administration of immunosuppressive drugs to avoid anti-ACP immune responses, or encapsulation of ACPs to protect them against

proteolytic degradation. Finally, the current high costs of ACPs manufacturing could be mitigated by future investments from pharmaceutical companies, which will lead to economies of scale and innovative strategies for ACP isolation or synthesis. (Gaspar et al., 2014; Hoskin & Ramamoorthy, 2008)

4. CELL-PENETRATING PEPTIDES (CPPs)

CPPs are defined as short (at most 30–35 amino acids), water –soluble, polybasic or partly hydrophobic peptides with a net positive charge at physiological pH. They are able to penetrate the live cell membrane at low micromolar concentrations by a seemingly energy-independent pathway, without using any chiral receptors and without causing significant membrane damage. In addition, CPPs are capable of internalizing conjugated biological active cargoes, typically big hydrophilic molecules, with high efficiency and low toxicity. (Heitz, Morris, & Divita, 2009; Lindgren, Hällbrink, Prochiantz, & Langel, 2000; Madani et al., 2011; Patel, Zaro, & Shen, 2007)

Penetration is a seemingly energy-independent translocation mechanism by means of which peptides cross the cell membrane transporting conjugated cargoes into the cytoplasm and nucleus. Cell penetration is a non-receptor-dependent process and thus, not cell-specific. Traditionally, hydrophilic molecules enter the cell through endocytic processes. Endocytosis involves the absorption to the cell membrane or a membrane-bound receptor, followed by energy-dependent formation of a vesicle. (Lindgren et al., 2000)

4.1 CLASSIFICATION OF THE CPPs

CPPs can be classified according to their origin in three classes: (Lindgren et al., 2000; Madani et al., 2011; Zorko & Langel, 2005)

- Protein-derived CPPs: also known as protein transduction domains (PTDs) or membrane translocation sequences, they usually consist of the minimal effective partial sequence of the parent translocation protein. Examples: penetratin, Tat.
- Model or synthetic CPPs: they comprise sequences that have been designed with the aim of producing well defined amphipathic α -helical structures or mimicking the structures of known CPPs. Examples: MAP, (Arg)₇.
- Designed CPPs: they are usually chimeric peptides composed of a hydrophilic and a hydrophobic domain of different origin. Examples: MPG, transportan.

Based upon their sequence and structural properties, CPPs can also be divided into three groups: (Madani et al., 2011)

- Primary amphipathic CPPs: they contain typically more than 20 amino acids and possess two sequential hydrophilic and hydrophobic domains in their primary structure. Some of them are toxic to cells even at low concentrations, and they can interact with both natural and anionic lipid membranes. Some examples are transportan or TP10.
- Secondary amphipathic CPPs: they often contain a smaller number of amino acids compared with primary amphipathic CPPs. Amphipathicity appears when the peptides adopt their secondary structure (α -helix or β -sheet) upon interaction with a phospholipid membrane. They typically bind to model membranes with a certain fraction of anionic lipids. Some examples are penetratin or pVEC.
- Non-amphipathic CPPs: they are rather short peptides with a high content of cationic residues (typically, Arg). They bind to lipid membranes with a high amount of anionic lipids. Some examples are R9 or TAT(48–60).

4.2 MECHANISMS OF UPTAKE

The mechanism of CPPs cellular translocation has been cause of discussion during the last decades. At present, it is believed that there is not a unique mechanism of translocation, and it depends on the CPP type. Furthermore, most CPPs employ two or more uptake pathways depending on the conditions. Two main mechanisms of uptake have been described so far: non-endocytic (energy-independent) or endocytic pathways: (Madani et al., 2011)

- Non-endocytic pathways

Also known as direct penetration, they involve energy-independent mechanisms. The first step is always the interaction between the positive charged peptide and anionic components of the cell membrane. A transient destabilization of the membrane takes place, followed by the folding of the peptide on the membrane. Next steps depend on the specific mechanism (see section 2.2), which can follow a carpet-like model (Figure 3.3), pore formation (barrel-stave model or toroidal, Figures 3.4 and 3.5), a membrane thinning model or the formation of inverted micelle structure. (Madani et al., 2011; Patel et al., 2007)

- Endocytic pathways

Nowadays it is generally accepted that endocytosis is involved in the translocation mechanism of most CPPs. Earlier results pointing to the opposite were due to experimental artefacts. Two classes of endocytic pathways of peptide internalization are known: non-receptor-mediated endocytosis (macropinocytosis) and receptor-

mediated endocytosis. The latter can be divided in three types: clathrin-mediated, caveolin-mediated and clathrin-and-caveolin-mediated endocytosis. The mechanism followed for each peptide usually is related to the nature of the cargo attached to it, as different mechanisms imply distinct vesicle sizes. (Madani et al., 2011; Patel et al., 2007)

4.3 CONJUGATED CARGOES

As the main potential application of CPPs is the transport of molecules through cell membranes, they are required to be capable of forming a chemical linkage with the cargo. The link between the CPP and cargo is usually a covalent bond and, when the cargo is a peptide or protein, the whole CPP-cargo ensemble is often synthesized or expressed in tandem as a fusion protein. Suitable amino acid side-chains or bifunctional spacer molecules can also be used to this purpose. The covalent strategy of fusion is mainly reported for the delivery of DNA mimic molecules or steric block oligonucleotides. This strategy offers some advantages, including rationalization, reproducibility of the procedure, and the control of the stoichiometry of the CPP-cargo, but some limitations have been reported, such as the risk of altering the biological activity of the cargo. (Heitz et al., 2009; Zorko & Langel, 2005)

In short, amphipathic peptides with hydrophobic and hydrophilic differentiated regions (primary amphipathic CPPs), sometimes the cargo-CPP coupling is achieved by the formation of non-covalent complexes, for instance using the biotin-avidin interaction. This strategy has been used for gene delivery. (Heitz et al., 2009; Zorko & Langel, 2005)

4.4 CLINICAL RELEVANCE OF CPPS

Many molecules have been traditionally considered of limited therapeutic interest because of their low biomembrane permeability and relatively rapid degradation. In this regard, CPPs could be used to facilitate the transport of those molecules through the cell membrane and its releasing into the cytoplasm or nucleus where they can perform their therapeutic activity. Since CPPs show a low cytotoxicity and no limitations for the type of the cargo they can carry, a great variety of molecules have been studied as CPP cargoes (small molecules, oligonucleotides, plasmid DNA, peptides, proteins, nanoparticles, lipid-based formulations, viruses, quantum dots,...). However, most of the applications describe the delivery of oligopeptides/proteins and nucleic acids or analogues. (Heitz et al., 2009; Lindgren et al., 2000)

Plasma membrane of eukaryotic cells is poorly permeable to DNA, and nucleic acids or oligonucleotides reach their targets within cells with a low efficiency. These limitations are the two main obstacles for the development of gene delivery as a therapeutic strategy. To overcome these difficulties, some CPP-based approaches have been explored. However, only a few CPPs have been validated *in vivo* for gene delivery, and so far, PPTG1 peptide constitutes one of the few examples with a significant reported *in vivo* gene expression response following intravenous injection. (Heitz et al., 2009)

The poor nuclear translocation shown by non-viral gene delivery systems is one of the most important barriers to be surpassed for transfection of non-dividing cells and gene therapy. In this regard, synthetic CPPs containing the nuclear localization signal (NLS) have been developed to facilitate the nuclear translocation of these peptides and augmentate their therapeutic efficiency. (Heitz et al., 2009)

The delivery of steric block small neutral oligonucleotides, such as PNAs or PMO, constitutes a potent tool for either antisense application or mRNA splicing correction. To facilitate the internalization of uncharged PNAs and PMO, many CPPs have been designed but, up to now, only a few have been used *in vivo* and none of them were active at submicromolar concentrations, showing some limitations due to their endosomal sequestration. Improvements have been made introducing some chemical modifications, and an increase of stabilization of the complexes and a reduction of non-specific cytotoxic effects have been observed. (Heitz et al., 2009)

A powerful instrument to control protein activation and/or gene expression post-transcriptionally is the utilization of decoy oligonucleotides and short interfering RNAs (siRNA), but the poor cellular uptake associated to the low permeability of the cell membranes to nucleic acids constitutes an important limitation. The use of CPPs to get a more efficient internalization of these nucleic acids is challenging in the case of charged siRNAs, since their multiple anionic charges interact with the CPP moiety and inhibit the uptake by steric hindrance. To solve this situation, non-covalent strategies are more appropriate and yields significant associated biological response. (Heitz et al., 2009)

The utilization of CPP-based strategies to deliver peptides and proteins to target different diseases, including cancer, asthma, ischaemia and diabetes, has been successful. Most of these applications use CPPs covalently linked to peptides or as fusion proteins. Successful *in vivo* applications of bioactive peptides attached to CPPs has been reported through very different modes of administration including intravenous, intra-tumoural, intra-tracheal injections, transduction into oocytes, sprays for nasal delivery or direct penetration through the skin. The use of CPPs associated to peptides has been reported in anti-proliferation treatments in cancer, or anti-apoptotic treatments in ischaemic events. CPPs have been also employed to deliver small molecules through the blood brain barrier and for the treatment of asthma by inhibiting the airway inflammatory response by cytokine blockage. (Heitz et al., 2009)

At present, many CPP-based strategies are under evaluation in preclinical and clinical assays. Promising results obtained in numerous studies during the last years have led to an increasing interest in CPPs from pharmaceutical companies. (Heitz et al., 2009)

5. OTHER MEMBRANE–ACTIVE PEPTIDES

As stated before, apart from the previously explained activities, membrane–active peptides can play many other roles, this is the case of antiviral peptides (AVPs), antifungal peptides, antiparasitic peptides (APPs), immunomodulatory peptides, fusogenic (FPs), and amyloid peptides.

5.1 ANTIVIRAL PEPTIDES (AVPs)

AVPs share many common features with MAPs with activities against bacteria, fungi, and other microorganisms. AVPs must be capable of disabling the viral infection in such a way that the host cell would remain intact and operational. Sometimes AVPs are created through CPP–based approaches, attaching antiviral drugs to a CPP. Other strategies of AVP design include the modification of antiviral molecules to achieve a cell–penetrating ability, or the synthesis of rationally designed peptides. (Pärn, Eriste, & Langel, 2015)

5.2 ANTIFUNGAL PEPTIDES

Peptides with antifungal activity are expressed by numerous organisms, including bacteria, fungi, plants, insects, amphibians and mammals, and many others have been developed through modifications of the sequence of known natural antifungal peptides, oriented to improve their therapeutic properties. Two main modes of action have been described for antifungal peptides: cell lysis and interference of the cell wall synthesis. The different mechanisms of cell lysis include membrane disruption through surface interactions, internalization and specific interaction with target molecules, and aggregation and formation of aqueous membrane pores. (De Lucca & Walsh, 1999)

5.3 ANTIPARASITIC PEPTIDES (APPs)

Since the late 1980s, many AMPs have been reported to show an additional antiparasitic activity. Several studies on protozoan parasites, such as *Plasmodium* sp. (malaria), *Leishmania* sp. (leishmaniasis), *Trypanosoma* sp. (sleeping sickness, Chagas), *Babesia* sp. (babesiosis), *Toxoplasma gondii* (toxoplasmosis) have been carried out. Effects of APPs on other kind of parasites, such as flatworms or hookworms, have been also studied, for instance with *Schistosoma* sp. (schistosomiasis), *Ancylostoma duodenale* or *Necator americanus* (hookworm disease). At present, only a few peptides with antiparasitic activity have reached phase III clinical trials and there is still a long way to go before any APPs will be commercially available. (Pretzel, Mohring, Rahlfs, & Becker, 2013)

5.4 IMMUNOMODULATORY PEPTIDES

Immunomodulatory properties have been reported in many known peptides since the 1970s. Glycopeptides, hormones and peptide fragments of immunoglobulins are usually considered as immunomodulatory peptides that regulate cell –mediated and humoral immune functions. Many of these peptides have been tested as immunologic

response stimulators in bacterial infections or as cancer cell growth inhibitors by stimulating the activity of immune competent cells. (Caccavo et al., 2002; Mohanty, Mohapatra, Misra, & Sahu, 2015)

5.5 FUSOGENIC PEPTIDES (FPs)

The infection process performed by enveloped viruses requires fusion of the viral and the target cell membranes in order to internalize the viral nucleocapsid in the host cell and initiate viral replication. The envelope proteins responsible for membrane fusion contain fusion domains known as fusogenic peptides (FPs). FPs interact with the target cell membrane and trigger the fusion process. (Reichert et al., 2006)

5.6 AMYLOID PEPTIDES

Amyloid peptides are characterised by their ability to form highly stable β -sheet-rich structures known as amyloid fibres. Amyloid peptides play an important role in many biological processes and they are related to over twenty mortal human diseases, including Alzheimer's and Parkinson's diseases and amyotrophic lateral sclerosis (ALS). Surprisingly, many other amyloid peptides are essential for vital physiological processes, including RNA regulation and memory consolidation. Amyloid peptides can be membrane permeabilizing in synthetic systems, and the hypothesis that membrane disruption may contribute to their pathogenicity has been proposed. (Last et al., 2013; Mompeán García, 2015)

6. MAPs STUDIED IN THIS THESIS

In this work, two representative MAPs have been studied in detail. In the first part, we have investigated the mechanism of action of BP100, a synthetic, 11-residue peptide derived from natural sequences that displays good antimicrobial and cell-penetrating properties with low cytotoxicity and high stability. In the second part, the selected system is crotalictidin, a member of the cathelicidins family from a South American pit viper. Antimicrobial and antitumour activities have been reported for this kind of peptides, and thus we have explored its properties and structure–function relationship.

Part 1

**BP100, a short peptide with good
antimicrobial properties**

INTRODUCTION

1. ORIGIN OF BP100

1.1 MELITTIN

Melittin, a peptide present in the apitoxin (European honey bee venom), was first mentioned by Ernst Habermann in 1954, who named it after the Latin word for honey – *mel*–, which appears in the specific name of the European honey bee (*Apis mellifera*) and in a genus of African bees (*Melitta*). In that work, Habermann reported the haemolytic effects of the protein fraction of the bee venom –melittin–. (Habermann, 1954) In 1965, melittin was characterised as a 26–residue basic polypeptide with a molecular mass of about 2850 Da and a glycine residue in the N–terminus, pharmacologically very active and capable of haemolysing erythrocytes, even in high dilution. (Kreil, 1965) One year later, the exact sequence of melittin was published by Habermann: (Habermann & Jentsch, 1966)

Melittin amino acid sequence
G-I-G-A-V-L-K-V-L-T-T-G-L-P-A-L-I-S-W-I-K-R-K-R-Q-Q

In 1967, antimicrobial activity of melittin was reported for the very first time by a research group of the U.S. Naval Radiological Defence Laboratory. Melittin showed antibacterial activity against most of the Gram–positive and about half of the Gram–negative bacteria tested, being active even against some penicillin–resistant strains. (Jean F Fennell, Shipman, & Cole, 1967; Jean F. Fennell, Shipman, & Cole, 1968)

At present, it is known that melittin forms an amphipathic α –helix (Figure 3.7) and acts against a wide range of infectious agents, including Gram–positive and Gram–negative bacteria, and shows high haemolytic effects. It has been also proved that melittin is capable of inhibit replication of murine retroviruses, tobacco mosaic virus and herpes simplex virus, exposing its antiviral properties. (Wachinger et al., 1998)

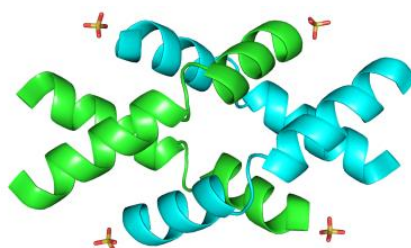


Figure 3.7. Structure of melittin forming a tetramer in the presence of phosphate. (PDB: 2MLT)

1.2 CECROPIN A

In 1980, two small basic proteins with potential antibacterial activity against some Gram-negative bacteria were purified and characterised from the immune system of cecropia moth (*Hyalophora cecropia*). Authors named them as “cecropins” and suggested that they represented a new class of antibacterial agents. (Hultmark, Steiner, Rasmuson, & Boman, 1980) One year after, this group published the sequence of the two cecropins, called cecropin A (CA) and cecropin B, and demonstrated that cecropin A activity was specific for bacteria, showing low haemolytic effects. (Steiner, Hultmark, Engstrom, Bennich, & Barman, 1981)

Cecropin A amino acid sequence
K-W-K-L-F-K-K-I-E-K-V-G-Q-N-I-R-D-G-I-I-K-A-G-P-A-V-A-V-V-G-Q-A-T-Q-I-A-K

1.3 CECROPIN A–MELITTIN HYBRIDS

Cecropin A–melittin hybrid peptides were designed in order to improve the properties of the isolated peptides as antimicrobials. CA(1–13)Mlt(1–13) and CA(1–8)Mlt(1–18) hybrids exhibited a wider antimicrobial spectrum, with an improved potency relative to cecropin A and without the undesirable cytotoxic effects of melittin. Although these designed hybrids show better properties, they are still too long (26 residues each) to be therapeutically appropriate. (Boman, Wade, Boman, Wählin, & Merrifield, 1989; Wade et al., 1990)

Cecropin A–melittin hybrids with shorter lengths (15 amino acids) were later developed using one of the previously mentioned longer hybrids, CA(1–8)Mlt(1–18), as a starting point, trying to retain the original antibacterial activity. Two new shorter peptides with antimicrobial activity comparable to that of the large one were found: CA(1–7)Mlt(2–9) and CA(1–7)Mlt(5–12). These peptides were the shorter cecropin A–based antibiotics retaining the original antimicrobial potency known in that moment, and they also showed antimalarial properties. (Andreu et al., 1992)

The 11–residue peptide Pep3 (WKLFKKILKVL–NH₂), derived from the CA(1–7)Mlt(2–9) hybrid was proved to be sufficient to maintain the antibacterial and antifungal activities. Variations of Pep3 sequence were performed and numerous analogues were synthesized and studied in order to find molecules with improved therapeutic characteristics. One of the analogues analysed, BP76 (KKLFKKILKFL–NH₂), resulted in a peptide with increased bactericidal activity and minimized cytotoxicity and susceptibility to protease degradation compared to the parent peptide Pep3. (Cavallarin, Andreu, & San Segundo, 1998; R. Ferre et al., 2006)

Based on peptide B76, a 125–member library of synthetic linear undecapeptides was prepared and studied trying to go further in the therapeutic improvement of this antimicrobial peptide. As a result of this approach, a new 11–residue peptide called BP100 (KKLFFKKILKYL–NH₂) was found to show higher bactericidal activity, being slightly lower than that of the potent streptomycin for plant bacterial infections. This peptide also possesses a low susceptibility to protease degradation and acceptable haemolysis effects. (Badosa et al., 2007)

BP100 amino acid sequence
K-K-L-F-K-K-I-L-K-Y-L-NH₂

2. PREVIOUS STUDIES ON BP100

AMPs and CPPs are very interesting tools from a therapeutic point of view, and thus BP100, due to its promising characteristics, has been studied in the last years in order to both unveil its mode of action and to improve its therapeutic index by increasing the antimicrobial activity and stability, and reducing the cytotoxic effects. The works mentioned in Table 3.1 were mainly oriented to describe the antimicrobial and cytotoxic activities of BP100, as well as to try to understand the mechanism behind its bactericidal activity. However, other works have been focused in the improvement of BP100 as an antimicrobial agent by introducing chemical modifications in its sequence. For instance, designed BP100–based peptidotriazoles (peptides containing a triazole ring in the backbone or onto the side chain of a selected residue) have been analysed as agents against bacterial and fungal phytopathogens. (Güell et al., 2012) Other example is the modification of BP100 by exchanging amino acids of the same kind, for example peptide R–BP100 in which Lys residues have been replaced by Arg, or peptide RW–BP100, in which Lys residues have been replaced by Arg and the Tyr has been replaced by Trp. These BP100 variants were designed to be more effective against Gram–positive bacteria. (Torcato et al., 2013)

Table 3.1. Chronological summary of the published works related to BP100.

Reference	Techniques utilized	Main results
Badosa, Ferre et al. (2007) <i>Peptides</i> 28 : 2276–2285 (Badosa et al., 2007)	MIC assays, haemolysis assays, bactericidal assays, protease degradation assays, infection assays	Low MIC values (2.5–7.5 μ M) against Gram –negative phytopathogenic bacteria. Low haemolysis at MIC values. Significant protease stability. Good bactericidal properties both <i>in vitro</i> and <i>in vivo</i> .
Ferre, Melo et al. (2009) <i>Biophys J</i> 96 : 1815–1827 (Rafael Ferre et al., 2009)	Fluorescence spectroscopy, turbidimetry, DLS, cytotoxicity assays	Preference for negatively charged membranes. Electroneutrality is reached at membrane saturation point. Importance of the saturation point in the mode of action is revealed.

(Table continues in the next page)

Alves, Melo et al. (2010) <i>J Biol Chem</i> 285 (36) : 27536-27544 (Alves et al., 2010)	MIC assays, LDME, AFM	Active against Gram-negative <i>E. coli</i> with low MIC values. Inactive against Gram-positive <i>S. aureus</i> (MIC > 200 μ M). Membrane neutralization achieved at MIC values. Changes in the bacterial cell observed upon treatment with BP100 (corrugation, membrane collapse and, finally, vesicle-like structure formation)
Eggenberger, Mink et al. (2011) <i>Chembiochem</i> 12 : 132-137 (Eggenberger, Mink, Wadhwani, Ulrich, & Nick, 2011)	Uptake fluorescence assays, fluorescence microscopy	BP100 is an efficient CPP and is capable of transporting functional cargoes across the membrane. Uptake of BP100 reaches high levels and occurs by a mechanism other than standard endocytic turnover, and actin filaments are essential to the mechanism. Uptake is enhanced in post-cycling cells.
Wadhwani, Epand et al. (2012) <i>Biophys J</i> 103 : 265-274 (P. Wadhwani, R. F. Epand, et al., 2012)	DSC, MIC assays, CD spectroscopy	Active against both Gram-negative (<i>E. coli</i> and <i>P. aeruginosa</i>) and Gram-positive strains (<i>S. aureus</i> and <i>S. epidermidis</i>) (MICs \leq 20 μ M). Random coil structure in aqueous solution and α -helix in the presence of lipid vesicles. Anionic lipid clustering effect not observed.
Wadhwani, Strandberg et al. (2014) <i>BBA-Biomembranes</i> 1838 : 940-949 (Wadhwani et al., 2014)	CD spectroscopy, OCD spectroscopy, ssNMR spectroscopy	Confirmed random coil structure in aqueous solution and α -helix in the presence of lipid vesicles. BP100 lies almost flat on the membrane ($\tau \approx 110^\circ$ and $\varrho \approx 160^\circ$), the uncharged C-terminus points deeper to the membrane than the N-terminus, and the Lys side chains point out of the membrane into the solvent. BP100 is rather dynamic.
Manzini, Perez et al. (2014) <i>BBA-Biomembranes</i> 1838 : 1985-1999 (Manzini et al., 2014)	CD spectroscopy, NMR spectroscopy, DLS, LDME, leakage assays, optical microscopy	BP100 is random coil in aqueous solution and become helical in the presence of TFE, or lipid vesicles. This was confirmed by ^1H -NMR. Electroneutrality of the membrane is reached at P/L = 1:6. Mechanism seems to be cooperative and hydrophobic residues may play an important role in peptide binding. Mechanism depends on P/L ratio and on the membrane composition.
Misiewicz, Afonin et al. (2015) <i>J Biomol NMR</i> 61 : 287-298 (Misiewicz et al., 2015)	ssNMR	BP100 forms a "carpet" on a variety of membranes (from model bilayers to native membranes), which leads to a membrane thinning.

OBJECTIVES

BP100, a synthetic peptide derived from a melittin–cecropin A hybrid has been studied in collaboration with the group of Prof. Anne S. Ulrich from the Karlsruhe Institute of Technology (Karlsruhe, Germany). This peptide is known to possess a potent bactericidal activity, low cytotoxicity, and a good resistance to protease degradation. Cell-penetrating properties have also been described for this peptide in previous works (Eggenberger et al., 2011).

This AMP acts on the bacterial membrane inducing a damage which is the base of its antibacterial effect. However, its particularly short length, only 11 residues, makes it unlikely to operate through a transmembrane–pore mechanism. In order to elucidate the peptide orientation in different lipid membranes and obtain information that makes possible a better understanding of the mode of action of BP100, an alanine scan approach was carried out combined with CD, microbiological and haemolysis assays, ssNMR and dynamical NMR data analysis.

The specific objectives proposed in this part are:

- To check the importance of each BP100 residue for structure, antimicrobial activity and cytotoxicity.
- To evaluate the antimicrobial potency and cytotoxic effects of some BP100 variants.
- To study the orientation of BP100 in distinct lipid membranes and propose a mechanism of action for this peptide.

MATERIALS AND METHODS

1. MATERIALS

All Fmoc protected amino acids were purchased from Iris Biotech GmbH (Marktredwitz, Germany) or Novabiochem (Merck Chemicals Ltd., Nottingham, UK), except for 2,2,2-²H₃-Ala (Ala-d₃), which was from Cambridge Isotope Laboratories (Andover, USA). Coupling reagents used for peptide synthesis were purchased from Iris Biotech GmbH or Novabiochem. Solvents for synthesis and purification were purchased from Biosolve (Valkenswaard, The Netherlands) or Acros Organics (Geel, Belgium). 1,2-dimyristoyl-*sn*-glycero-3-phosphatidylcholine (DMPC) was from NOF Corporation (Grobbendonk, Belgium), while the other lipids 1,2-dimyristoyl-*sn*-glycero-3-phosphatidylglycerol (DMPG), 1-myristoyl-2-hydroxy-*sn*-glycero-3-phosphatidylcholine (lyso-MPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC), and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylglycerol (POPG) were purchased from Avanti Polar Lipids (Alabaster, AL, USA).

2. PEPTIDE SYNTHESIS

The parent peptide BP100, with the sequence KKLFKKILKYL-NH₂ (Figure 3.8), was synthesized with a single Ala-d₃ substitution at each position in the sequence. Additionally, the parent peptide was synthesised with a ¹⁵N-labelled leucine (¹⁵N-Leu) at position Leu8. Sequences and abbreviations of the synthesized peptides used in this work are listed in Table 3.2. Standard Fmoc-solid phase synthesis protocols were followed (Fields & Noble, 1990) on an automated Syro II peptide synthesizer (MultiSynTech, Witten, Germany). The identity of the peptides was confirmed by an LC-MS system equipped with a 1100 Series LC-system from Agilent (Santa Clara, USA) coupled to an ESI micro-TOF mass spectrometer from Bruker Daltonics (Bremen, Germany). Peptides were purified on a C18 reverse phase HPLC columns using water/acetonitrile gradients, each containing 5 mM HCl, and were > 95 % pure.

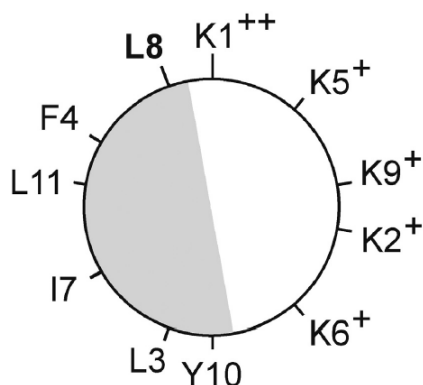


Figure 3.8. Helical wheel projection of BP100. The hydrophobic region is shown in grey and the polar region in white. Charged lysine residues and the N-terminus are marked with +, and the ¹⁵N-labelled Leu-8 is highlighted in bold.

Table 3.2 Sequences of BP100 analogues used in this study.

Peptide	Labelled position	Sequence	Mw	Charge
BP100-PP	None	KKLFKKILKYL-NH ₂	1421.8	+6
BP100K1A	Lys-1	Ala-d ₃ -KLFKKILKYL-NH ₂	1367.7	+5
BP100K2A	Lys-2	K-Ala-d ₃ -KLFKKILKYL-NH ₂	1367.7	+5
BP100L3A	Leu-3	KK-Ala-d ₃ -FKKILKYL-NH ₂	1382.8	+6
BP100F4A	Phe-4	KKL-Ala-d ₃ -KKILKYL-NH ₂	1348.7	+6
BP100K5A	Lys-5	KKLF-Ala-d ₃ -KILKYL-NH ₂	1367.7	+5
BP100K6A	Lys-6	KKLFK-Ala-d ₃ -ILKYL-NH ₂	1367.7	+5
BP100I7A	Ile-7	KKLFKK-Ala-d ₃ -LKYL-NH ₂	1382.8	+6
BP100L8A	Leu-8	KKLFKKI-Ala-d ₃ -KYL-NH ₂	1382.8	+6
BP100K9A	Lys-9	KKLFKKIL-Ala-d ₃ -YL-NH ₂	1367.7	+5
BP100Y10A	Tyr-10	KKLFKKILK-Ala-d ₃ -L-NH ₂	1332.7	+6
BP100L11A	Leu-11	KKLFKKILKY-Ala-d ₃ -NH ₂	1382.8	+6
BP100- ¹⁵ N-L8	Leu-8	KKLFKKI- ¹⁵ N-Leu-KYL-NH ₂	1422.8	+6

3. MINIMUM INHIBITORY CONCENTRATION ASSAYS

The antimicrobial activity of BP100 and its Ala-substituted variants were examined using a serial 2-fold dilution MIC assay, a standard procedure utilized in previous works of the group (P. Wadhwani, R. F. Epand, et al., 2012). Four bacterial strains were tested: two Gram-positive, *Staphylococcus aureus* (DSM 1104) and *Staphylococcus epidermidis* (DSM 1798), and two Gram-negative, *Escherichia coli* (DSM 1103) and *Pseudomonas aeruginosa* (DSM 1117).

4. HAEMOLYSIS ASSAYS

Haemolytic activities of the peptides were determined by a serial 2-fold dilution assay modified from previously published protocols (Strandberg et al., 2007). Citrate phosphate dextrose (CPD)-stabilized blood preservations with erythrocyte concentrate of healthy donors were provided from the local blood bank (Municipal Clinical Centre, Karlsruhe, Germany). The erythrocytes were washed twice with 9-fold excess Tris washing buffer (172 mM, pH 7.6 at 20°C), followed by centrifugation at 600 g for 10 min at 20°C. Afterwards, the erythrocytes were diluted from the sediment with Tris washing buffer to about 10 % (v/v) haematocrit. This stock erythrocyte suspension was further diluted to about 0.5 % with Tris reaction buffer (172 mM, pH 7.6 at 37 °C) for each peptide just prior to incubation. The peptides were dissolved in reaction buffer/DMSO (9/1 v/v). For each peptide a series of 2-fold dilutions of 200 µL in Tris reaction buffer were prepared in reaction tubes to get twice the desired end concentration. 200 µL of the 0.5 % erythrocyte suspension were transferred to each reaction tube to yield a final erythrocyte concentration of 0.25 %. The final peptide concentration was between 160 and 1.25 µM in 2-fold dilution steps. For each dilution series, a negative control (zero haemolysis) was obtained by adding the same amount of the erythrocytes to 200 µL Tris reaction buffer and measuring the background lysis in the absence of peptide. As a positive control (100 % haemolysis), 200 µL of the erythrocytes were added to 200 µL Tris reaction buffer with 0.2 % of Triton X-100 (Sigma, Germany), giving a final concentration of 0.1 % Triton X-100. Incubation was

performed at 37 °C for 30 min under mild agitation. Afterwards, the tubes were centrifuged at 20,000 g for 5 min to pellet the cells. The supernatant was transferred to cuvettes and the absorbance at 540 nm was measured against Tris reaction buffer. The lysis percentage was then calculated relative to 0 % lysis by reaction buffer and 100 % lysis by 0.1 % Triton X-100. The absorbance measurement was repeated three times and the average value was used. For each peptide, the haemolysis assay was performed at a series of concentrations, three times for each concentration, and for each data point the average of those three measurements and the standard deviation (SD) was calculated.

5. CIRCULAR DICHROISM (CD) SPECTROPOLARIMETRY

The lipid powders (DMPC and DMPG) were dissolved in chloroform/methanol 50/50 (v/v) to obtain stock solutions with a concentration of 5 mg·mL⁻¹ (7.4 mM for DMPC, and 7.3 mM for DMPG). Aliquots of the stock solutions were mixed in a glass vial and thoroughly vortexed to obtain a DMPC:DMPG 3:1 mixture (molar ratio). Subsequently, the organic solvents were removed under a gentle stream of nitrogen, followed by overnight vacuum. The lipid film formed at the bottom of the vial was dispersed by addition of 10 mM phosphate buffer (PB, pH 7.0) and strong vortexing during 10 min. The dispersion was homogenized by 10 freeze–thaw cycles followed by vigorously vortexing for 1 min after each cycle. Afterwards, small unilamellar vesicles (SUVs) were formed by sonication of the MLVs for 16 min in a strong ultrasonic bath at 35 °C (UTR 200, Hielscher, Germany).

To prepare the final CD samples, peptide stock solutions in water were prepared with a nominal concentration from weighed material of 1 mg·mL⁻¹. After UV measurements, the corrected concentrations of the peptide stock solutions were in the range 618–704 µM. An aliquot of the respective peptide stock solution in water was added to either pure 10 mM PB or to the DMPC:DMPG (3:1) vesicle dispersion in 10 mM PB. Final peptide concentration was in the 124–141 µM range for the 10 mM PB solutions, and in the liposome samples it was adjusted to ~29 µM, which resulted in a peptide–to–lipid molar ratio (P/L) of 1:50, given a lipid concentration of 1.47 mM.

CD spectra of these samples were recorded on a J-815 spectropolarimeter (JASCO, Groß-Umstadt, Germany). Measurements were performed in quartz glass cuvettes (Suprasil, Hellma, Müllheim, Germany) of 1 mm path length, between 260 and 180 nm at 0.1 nm intervals. Spectra were recorded at 25 °C for the pure PB buffer, and at 30°C for the vesicle suspensions (i.e. above the phase transition temperature of the lipids), using a water–thermostatted rectangular cell holder. Three repeat scans at a scan rate of 10 nm·min⁻¹, 8 s response time, and 1 nm bandwidth were averaged for each sample and for the baseline of the corresponding peptide–free sample. After subtracting the baseline spectra from the sample spectra, CD data were processed with an adaptive smoothing method, which is part of the Jasco Spectra Analysis software. Finally, the spectral data were converted to mean residue ellipticities ($[\theta]_{\text{MRE}}$), using the concentration of each peptide stock solution based on the 280 nm UV absorbance of the Tyr residue contained in the

sequence (Pace, Vajdos, Fee, Grimsley, & Gray, 1995). The absorption spectrum in the range of the Tyr aromatic bands was recorded from 340 to 240 nm in a quartz glass half-micro-cuvette, with 1 cm optical path length (Hellma, Müllheim), using milliQ-water as a blank. The corrected peptide concentration was calculated from the baseline-corrected absorbance using a molar extinction coefficient of $1490 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ for the Tyr absorption at 280 nm (Pace et al., 1995). For BP100Y10A, which has no tyrosine, the concentration used was the average value of the corrected concentrations for the other peptides, since all showed similar values.

Secondary structure was estimated from CD spectra using the CDSSTR program with the implemented singular value decomposition (SVD) algorithm (Johnson, 1999; Sreerama, Venyaminov, & Woody, 2000), by the CONTIN-LL program, which is based on the ridge regression algorithm (Provencher & Glockner, 1981; van Stokkum, Spoelder, Bloemendal, van Grondelle, & Groen, 1990), and by the SELCON-3 program, which incorporates the self-consistent method together with the SVD algorithm to assign protein secondary structure. (Sreerama, Venyaminov, & Woody, 1999; Sreerama & Woody, 1993) The three algorithms are provided by the DICHROWEB on-line server. (Lobley, Whitmore, & Wallace, 2002; Whitmore & Wallace, 2004) In each program, an appropriate reference protein dataset provided by DICHROWEB (in this case, reference set 7) was employed, which was the most suitable for the secondary structure expected and the wavelength range used. The lowest data point used in the analysis was 190 nm (below that value the signal acquired by the spectrometer is not reliable because of saturation of the detector). Input and output units were $[\theta]_{\text{MRE}}$. The quality of the fit between experimental and back-calculated spectra corresponding to the estimated secondary structure fractions was assessed from the normalized root mean square deviation (NRMSD), with a value <0.1 (for CONTIN-LL and CDSSTR) and <0.25 (for SELCON-3) considered as a good fit. (Whitmore & Wallace, 2004) Finally, the secondary structure element fractions of each sample were calculated as the average value of the individual data obtained with the three algorithms. Individual values were not considered for the average if the sum of all structural elements fractions was <0.98 or >1.02 , or when the NRMSD (normalized root mean square deviation) between the experimental and back-calculated CD spectrum was above the threshold value (0.1 for CONTIN-LL and CDSSTR, and 0.25 for SELCON-3).

6. NMR SAMPLE PREPARATION

Oriented NMR samples were prepared by dissolving lipids and peptides in a mixture of methanol, chloroform, and milliQ-water, and spreading onto thin glass plates of dimensions $9 \text{ mm} \times 7.5 \text{ mm} \times 0.08 \text{ mm}$ (Marienfeld Laboratory Glassware, Lauda-Königshofen, Germany). Then, the organic solvents were removed under overnight vacuum, and subsequently, samples were hydrated at 48°C and 96 % relative humidity, using deuterium-depleted water for the samples labelled with Ala- d_3 . Further experimental details have been published in previous works of the group (Strandberg, Wadhwani, Tremouilhac, Durr, & Ulrich, 2006; Parvesh Wadhwani et al., 2012). Usually, 0.5–1 mg of ^2H - or ^{15}N -labelled peptide were used, and appropriate amounts of lipids to obtain the desired P/L ratio.

7. SOLID STATE NMR

^{31}P -NMR measurements to evaluate the sample quality and the degree of lipid alignment were performed at 202.5 MHz on an Avance III Bruker NMR spectrometer (Bruker Biospin, Karlsruhe, Germany), with a wide bore 500 MHz magnet, using a Hahn echo sequence (Rance & Byrd, 1983) (90° pulse of 5 μs and a 30 μs echo time), with ^1H SPINAL64 decoupling (Fung, Khitritin, & Ermolaev, 2000) during acquisition. 256 scans were accumulated using an acquisition time of 10 ms and a relaxation delay of 1 s. ^2H -NMR was performed using a home-built goniometer probe head at a frequency of 76.77 MHz using a solid echo pulse sequence (Davis, Jeffrey, Bloom, Valic, & Higgs, 1976) with a 90° pulse of 5 μs , a 70 μs echo time, a sweep width of 500 kHz, and 4096 data points. Typically, between 100,000 and 700,000 scans were acquired, with a relaxation delay of 100 ms.

^1H - ^{15}N cross polarization experiments using a CP-MOIST pulse sequence (Levitt, Suter, & Ernst, 1986) were performed at 60.8 MHz using a home-built double-tuned probe with a low-E flat-coil resonator, employing a ^1H and ^{15}N radiofrequency field strength of 65 kHz during the cross polarization, and 36 kHz ^1H SPINAL16 decoupling during acquisition. A mixing time of 250 μs was used (Wadhvani et al., 2014), and up to 25,000 scans were accumulated. The acquisition time was 10 ms and the recycle time 3 s. The ^{15}N chemical shift was referenced using the signal of a dry powder of ^{15}N -ammonium sulphate, of which the chemical shift was set to 26.8 ppm. All NMR experiments were performed at 35 $^\circ\text{C}$, unless stated otherwise. The temperature of the sample inside the NMR probe was calibrated using a methanol sample. (Ammann, Meier, & Merbach, 1982)

8. NMR DATA ANALYSIS

The orientation of a helical peptide in a membrane is defined by two angles, the tilt angle, τ , defined as the angle between the long axis of the helix and the membrane normal, and the azimuthal rotation angle, ϱ , which defines the rotation of the peptide around its long axis (Figure 3.9) (Glaser, Sachse, Durr, Wadhvani, & Ulrich, 2004; Strandberg et al., 2008; Strandberg et al., 2004; van der Wel, Strandberg, Killian, & Koeppe, 2002). For calculating the peptide orientation and dynamics, the structure of BP100 is assumed to be an ideal α -helix, based on a poly-alanine model generated with SYBYL (Tripos, St. Louis, USA) (Glaser et al., 2004). Using ^2H -NMR data from the labelled positions, the helix orientation is determined from a grid search of the tilt angle (τ) and azimuthal rotation angle (ϱ) giving the lowest root mean square deviation (RMSD) between calculated and experimental quadrupolar splittings from different labelled positions. Supplementary information about this approach can be found in previous publications of the group. (Glaser et al., 2005; Strandberg et al., 2004)

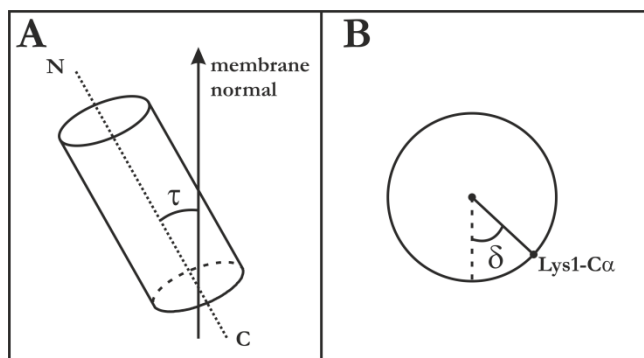


Figure 3.9. Definition of τ and ϱ angles. **A.** Tilt angle (τ) is defined between the peptide axis and the membrane normal (assumed to be along the magnetic field direction). **B.** Azimuthal rotation angle (ϱ) is the rotation of the helix compared to a standard orientation with C_α of Lys1 in the xz plane (dashed line).

For BP100, rotation azimuthal angle (ϱ) was defined using the C_α of residue-1. (Wadhvani et al., 2014) Peptide dynamics is described as Gaussian distributions of the τ and ϱ angles, with their respective widths, σ_τ and σ_ϱ . Larger widths correspond to more dynamic situations, in which the angles undergo fluctuations with bigger amplitudes. It is assumed that the whole-body helix fluctuations are fast on the NMR time scale, so that the measured splittings represent time-averages over these distributions. (Strandberg, Esteban-Martin, Salgado, & Ulrich, 2009) The quality of the fits was analysed by τ - ϱ RMSD plots and quadrupolar wave plots.

RESULTS AND DISCUSSION

1. ANTIMICROBIAL AND HAEMOLYTIC ACTIVITY OF BP100 ANALOGUES

Besides the original parent peptide BP100, eleven analogues were synthesized with a single Ala-d₃ at each position, as shown in Table 3.2. To investigate the influence of these Ala substitutions on the biological function, the antimicrobial activity of BP100 and the Ala analogues was tested against four representative Gram-positive and Gram-negative strains. The minimum inhibitory concentrations (MICs) are given in Table 3.3. A clear trend can be seen in all four bacterial strains: When any Lys is replaced by Ala, the activity is the same as in BP100 (within one dilution factor) or even higher (in *S. aureus* for BP100KA5 and BP100KA6). Thus, removing one of the charges from BP100 (5 Lys, free N-terminus) does not reduce its antimicrobial activity. On the other hand, when one of the hydrophobic residues (Leu3, Phe4, Ile7, Leu8, or Leu11) is replaced by Ala, the activity drops dramatically, giving MIC values 4–16 times higher than BP100. Replacement of Tyr10 leads to a slight reduction in activity.

Table 3.3. MICs (in μM) of BP100-PP peptide and Ala-substituted analogues in four different bacterial strains. The assay was done twice and when the values differed between the two experiments both values are given. The highest tested concentration was 160 μM .

Peptide	Gram-negative		Gram-positive	
	<i>E. coli</i> (DSM 1103)	<i>P. aeruginosa</i> (DSM 1117)	<i>S. aureus</i> (DSM 1104)	<i>S. epidermidis</i> (DSM 1798)
BP100-PP	2.5–5	5	40	10
BP100K1A	2.5–5	10	40	10
BP100K2A	2.5–5	10–20	40	10
BP100L3A	20–40	20	160/>160	40
BP100F4A	40	20	>160	80–160
BP100K5A	2.5	10	10–20	10
BP100K6A	2.5–5	10	10	10
BP100I7A	20	40	>160	160
BP100L8A	20–40	20–40	>160	80
BP100K9A	2.5–5	10–20	40	10
BP100Y10A	20	10–20	80	20
BP100L11A	40	40	>160	160/>160

The haemolytic activity of the peptides was determined as a function of peptide concentration, and results are given in Figure 3.10. All Ala analogues were less haemolytic than BP100, but the differences between them were large. The most haemolytic peptides were BP100K1A, BP100K5A, and BP100K6A, which behaved similar to the parent peptide. BP100K2A, BP100K9A, and BP100Y10A had intermediate haemolytic activity, while replacement of the hydrophobic sites (BP100L3A, BP100F4A, BP100I7A, BP100L8A and BP100L11A) reduced haemolysis dramatically, even at 160 μM peptide concentration. Lowering the hydrophobicity of the peptide was thus generally seen to lead to a lower activity against both bacterial and erythrocyte membranes.

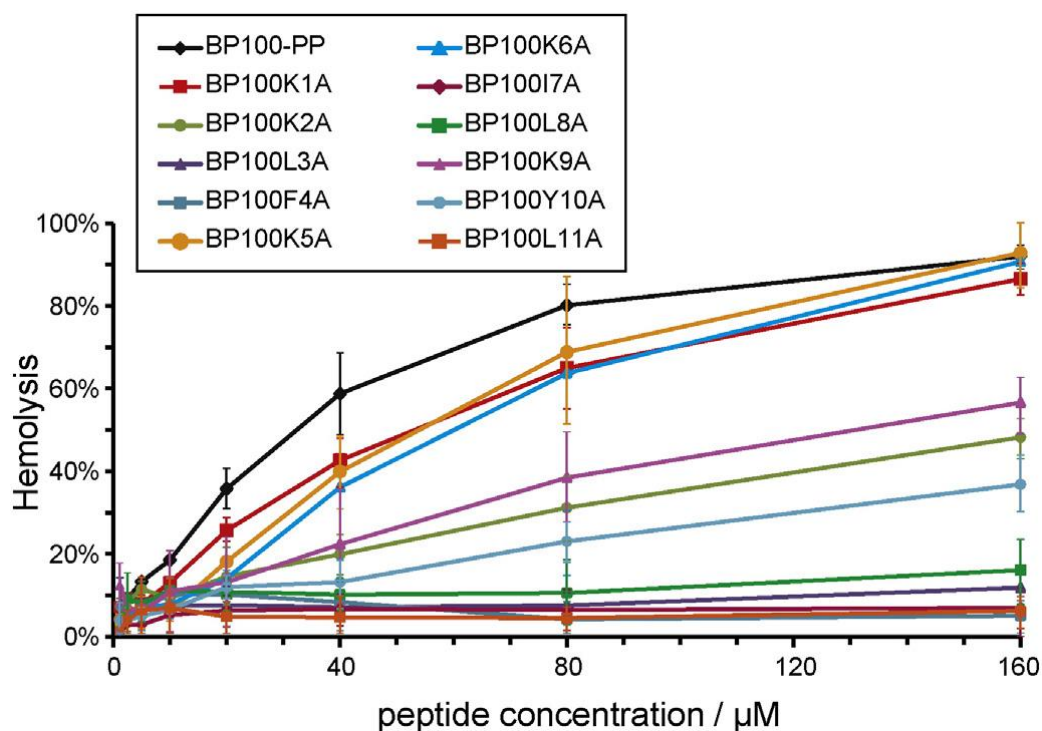


Figure 3.10. Haemolysis of BP100–PP and Ala–d₃ substituted analogues. 100 % haemolysis is the value after addition of Triton–X100. The error bars correspond to \pm the standard deviation for each data point. All Ala analogues show less haemolytic activity than BP100–PP.

A quantitative comparison of the haemolytic activities and the MIC values (Table 3.4) shows that about 5–13 % haemolysis occurs at the MIC value against *E. coli*, but no clear trend is seen. Amongst the K→A analogues, which have a comparably high antimicrobial activity similar to the parent peptide, BP100K2A shows the lowest haemolysis. As this analogue is only about half as haemolytic as the parent peptide, it may be a suitable candidate for developing improved agents with a better therapeutic index.

Table 3.4. Haemolysis of BP100 and Ala–d₃ substituted analogues in %, at different peptide concentrations. 100 % haemolysis is the value after addition of triton–X100. The haemolysis found at the MIC against *E. coli* for each peptide is marked in bold and underlined.

Conc. (μM)	BP100–PP	Position labelled with Ala–d ₃										
		K1A	K2A	L3A	F4A	K5A	K6A	I7A	L8A	K9A	Y10A	L11A
1.25	7	7	7	8	4	4	4	1	3	12	4	1
2.5	9	9	8	4	3	<u>6</u>	5	3	9	7	4	4
5	<u>13</u>	<u>8</u>	<u>11</u>	7	7	7	<u>7</u>	3	8	<u>6</u>	5	6
10	18	13	9	7	10	6	8	5	11	11	6	7
20	36	26	15	8	10	18	14	<u>6</u>	11	13	<u>12</u>	5
40	59	43	20	<u>7</u>	<u>8</u>	40	36	7	<u>10</u>	22	13	<u>5</u>
80	80	65	31	7	4	69	64	6	11	38	23	4
160	92	87	48	12	5	93	91	7	16	57	37	6

2. CD CHARACTERIZATION OF BP100 ANALOGUES

To determine the effect of the Ala replacements on the conformation of BP100, circular dichroism spectra were recorded in 10 mM PB, and in the presence of DMPC:DMPG (3:1) vesicles at a P/L ratio of 1:50. In phosphate buffer, BP100 and all Ala analogues are unstructured, as seen in the spectra with typical random coil features, such as the minimum around 198 nm, and mostly negative ellipticities over the full spectral range from 185 to 260 nm (Figure 3.11A). All peptides fold as α -helices in the presence of negatively charged DMPC:DMPG (3:1) vesicles, as deduced from the spectral line shapes in Figure 3.11B with a positive maximum around 192 nm, and two negative bands at 208 nm and 223 nm.

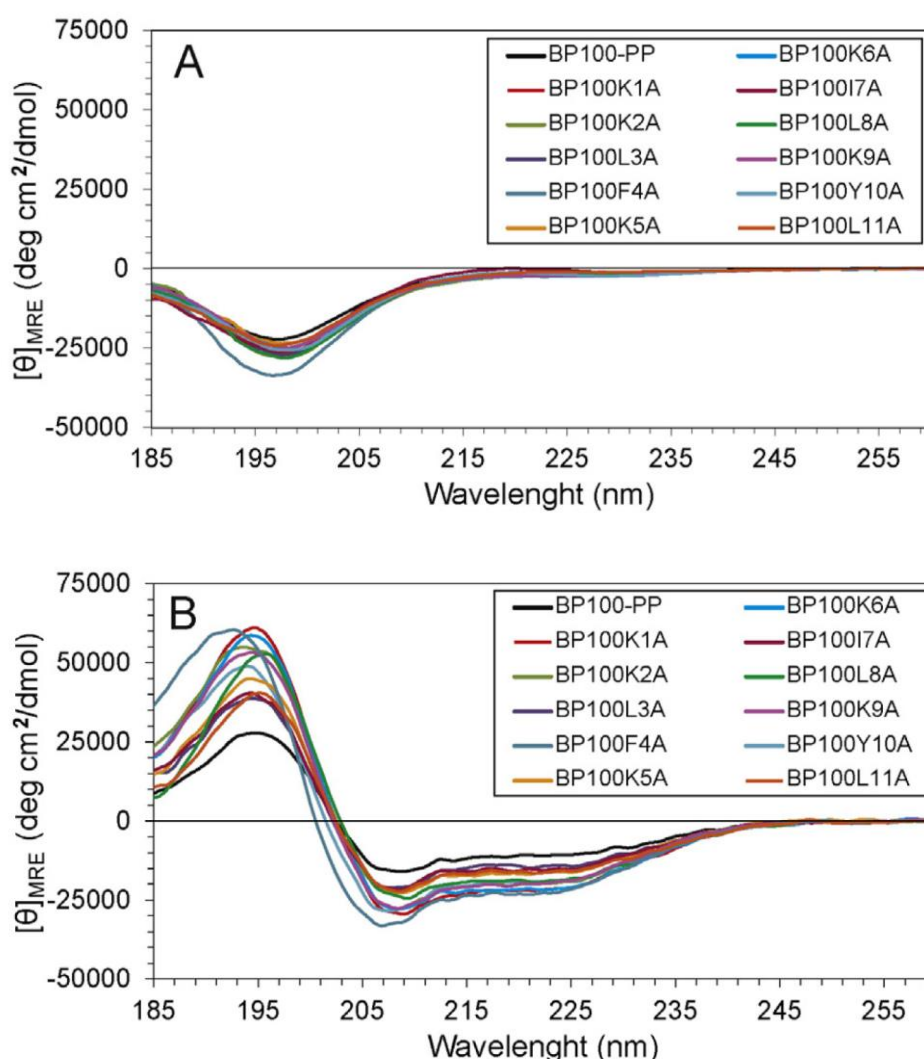


Figure 3.11. CD spectra of BP100 and Ala-d₃ substituted analogues in 10 mM PB (A), and in the presence of DMPC:DMPG (3:1) vesicles, P/L = 1:50, 10 mM PB (B).

The CD spectra in the presence of DMPC:DMPG (3:1) were deconvoluted to obtain the secondary structure elements of the different BP100 analogues, as given in Table 3.5. The short BP100 peptides are highly α -helical. The Ala substituted analogues showed between

73 % and 89 % α -helicity, compared to 58 % for the parent peptide. In a previous study, BP100 under the same conditions had given 61 % α -helix, and selectively CF₃-Bpg substituted peptides values from 71 % to 87 % α -helix. (Wadhwani et al., 2014) Thus, the CD conformational analyses together with the antimicrobial tests confirm that Ala substitutions of BP100 do not perturb the well folded and highly α -helical structure of the peptide, but rather seem to stabilize it.

Table 3.5. Secondary structure fractions of BP100-PP peptide and Ala-labelled analogues in DMPC:DMPG (3:1) vesicles, evaluated from the CD spectra using three different algorithms.

Peptide	Fraction of secondary structure element*				
	α -helix	β -sheet	turn	unordered	total
BP100-PP	0.58	0.09	0.09	0.23	1.00
BP100K1A	0.87	0.03	0.04	0.05	1.00
BP100K2A	0.88	0.00	0.04	0.08	1.00
BP100L3A	0.73	0.05	0.08	0.14	1.00
BP100F4A	0.83	0.01	0.05	0.11	1.00
BP100K5A	0.80	0.02	0.06	0.11	1.00
BP100K6A	0.89	0.02	0.04	0.05	1.00
BP100I7A	0.74	0.02	0.07	0.15	0.99
BP100L8A	0.84	0.00	0.06	0.09	1.00
BP100K9A	0.87	0.01	0.03	0.10	1.00
BP100Y10A	0.87	0.00	0.03	0.11	1.00
BP100L11A	0.75	0.05	0.07	0.13	1.00

*Data represent mean values of the results obtained with three different secondary structure estimation algorithms. Individual results of the different analyses were not considered when the sum of all structural element fractions was <0.98 or >1.02, or when the NRMSD (normalized root mean square deviation) between the experimental and back-calculated CD spectrum was above the threshold value (0.1 for CONTIN-LL and CDSSTR, and 0.25 for SELCON-3).

3. SOLID STATE NMR STUDY OF BP100 ANALOGUES IN LIPID SYSTEMS

²H-NMR experiments on the Ala-d₃ labelled peptides were performed in three different lipid systems. Previous ¹⁹F-NMR experiments had been done on CF₃-labelled BP100 in DMPC:DMPG (3:1), (Wadhwani et al., 2014), so this system was used here for comparison. For the case of several other peptides, the group of Prof. Ulrich has demonstrated that spontaneous curvature of the lipids can have an influence on peptide orientation. (Strandberg, Tiltak, Ehni, Wadhwani, & Ulrich, 2012; Strandberg & Ulrich, 2015; Strandberg, Zerweck, Wadhwani, & Ulrich, 2013) Therefore, the experiments have been extended here to include POPC:POPG (3:1), which has a negative spontaneous curvature, as well as DMPC:DMPG:lyso-MPC (1:1:1), which has a positive spontaneous curvature.

In the previous ¹⁹F-NMR study performed by the group of Prof. Anne S. Ulrich (Wadhwani et al., 2014), only minor changes in the dipolar splittings were found in DMPC:DMPG (3:1) when the P/L ratio was varied between 1:10 and 1:3000, showing that the peptide orientation did not change significantly as a function of concentration. In this work, the influence of peptide concentration was also monitored as a control by preparing

samples for BP100K5A (a peptide which gave large, well resolved splittings) at P/L values of 1:10, 1:20, 1:50 and 1:100 in each of the three lipid systems.

In Figure 3.12, the ^{31}P -NMR spectra of the oriented NMR BP100K5A samples are shown. The peak around 30 ppm stems from oriented phospholipids, and the small signal around -15 ppm originates from unoriented lipids, which would usually correspond to the highest intensity in a non-oriented sample. Integration of the two peaks showed that 70–95 % of the lipids were well oriented, depending on sample. It can be noted that at peptide concentrations between P/L = 1:100 and 1:20, there is only a single oriented peak in the POPC:POPG and DMPC:DMPG samples (Figures 3.12A and 3.12B), while at P/L = 1:10 the signals from oriented PC and PG head groups are resolved. This observation indicates that the cationic peptide binds to the membrane and preferentially interacts with the negatively charged PG lipid head groups, shifting their chemical shift slightly. In the DMPC:DMPG:lyso-MPC samples (Figure 3.12C), there is an additional peak around 15 ppm from oriented lyso-MPC. Also here, at P/L = 1:100 and 1:50, the PC and PG signals are overlapping, but at P/L = 1:20 and 1:10 there is a shift of the PG peak and three peaks can be distinguished in the spectra.

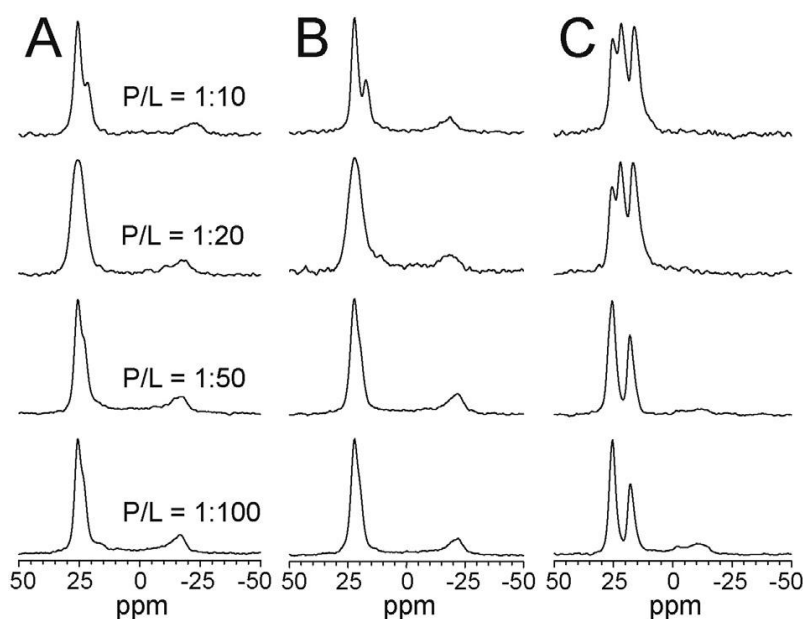


Figure 3.12. ^{31}P NMR spectra of oriented samples of lipids with BP100K5A, measured at 0° sample orientation, at different P/L ratios from 1:10 to 1:100, in (OPC:POPG (3:1) (A); DMPC:DMPG (3:1) (B); and DMPC:DMPG:lyso-MPC (1:1:1) (C).

Concentration dependent ^2H -NMR spectra of the Ala- d_3 labelled peptides were measured at both 0° inclination of the oriented samples (with the membrane normal parallel to the external magnetic field B_0) and 90° inclination (with the membrane normal perpendicular to B_0). The spectra are shown in Figure 3.13, and splittings are listed in Table 3.6. All spectra displayed a central peak, representing signals from unbound peptides and possibly from residual deuterium in the water. It can be noted that the amount of peptide is the same at different P/L ratios, meaning that the amount of lipid and water is higher at low P/L. This relationship can, at least partly, explain why the central peak is more prominent at low peptide concentration. This central signal is strongly reduced at higher P/L for POPC:POPG and DMPC:DMPG:lyso-MPC. There is also a large splitting around 40 kHz from peptides bound to the membranes, which is the interesting part of the

spectrum, as it gives information about the peptide orientation, and which is used in the analysis below. A smaller splitting of approximately 20 kHz is visible in some of the spectra, and is attributed to a small fraction of non-oriented material in the sample. It can be identified from the fact that the same splitting is found at a 90° inclination of the sample. Some POPC:POPG samples show a splitting of approximately 60 kHz with very low intensity, being probably due to residual deuterium in the oriented lipids, which produces intensity in this characteristic region. (Balla, Bowie, & Separovic, 2004; Misiewicz et al., 2015; Oldfield, Meadows, Rice, & Jacobs, 1978)

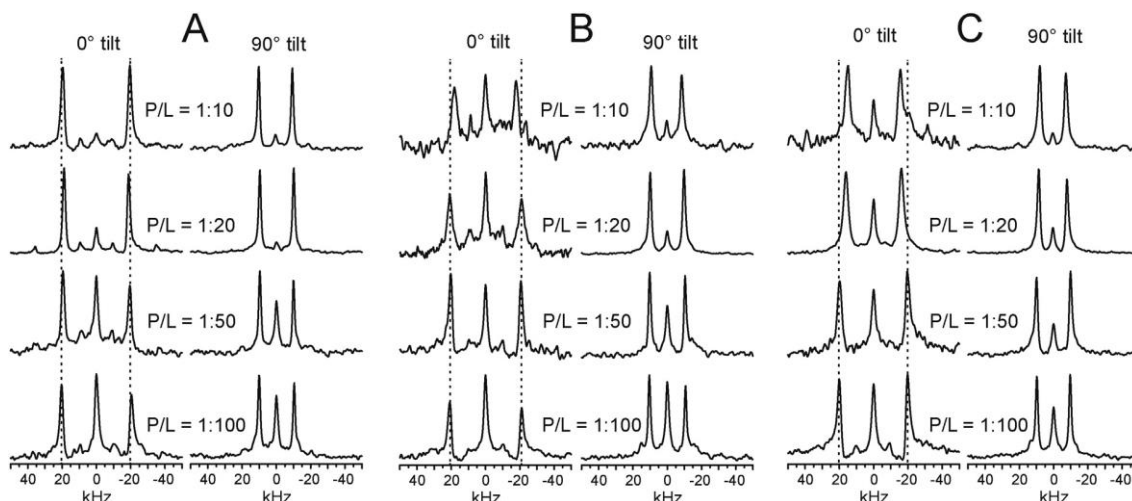


Figure 3.13. ^2H NMR spectra of oriented samples of BP100K5A, measured at 0° and 90° sample orientation, at different P/L ratios from 1:10 to 1:100 in POPC:POPG (3:1) bilayers (A); (DMPC:DMPG (3:1) (B); and DMPC:DMPG:lyso-MPC (1:1:1) (C). Dashed lines indicate the splittings at P/L = 1:100.

Table 3.6. ^2H quadrupolar splittings (in kHz) of BP100K5A in different oriented lipid bilayers, at P/L=1:20, measured with the membrane normal parallel (0°) and perpendicular (90°) to the magnetic field.

Lipid system						
P/L	POPC:POPG (3:1)		DMPC:DMPG (3:1)		DMPC:DMPG:lyso-MPC (1:1:1)	
	0°	90°	0°	90°	0°	90°
1:10	39.2	19.6	36.1	17.7	30.6	15.2
1:20	39.7	19.8	39.6	19.7	33.0	16.5
1:50	39.4	20.6	40.8	20.6	39.4	19.8
1:100	40.7	20.4	41.9	21.1	39.8	19.4
1:20 (repeated)	37.6	-	41.8	-	32.3	-

At a low peptide concentration of P/L = 1:100, the peptide splitting is almost the same in all three lipid systems, namely 41 ± 1 kHz at 0° sample inclination. In POPC:POPG, the splitting does not change much with concentration, and remains essentially the same up to 1:10. In DMPC:DMPG, splittings are the same between 1:100 and 1:20, but there is a change in 1:10 to 36 kHz. In DMPC:DMPG:lyso-MPC, the splittings are the same at 1:100 and 1:50, but at 1:20 a significantly smaller splitting of 33 kHz is observed, and at 1:10 this is further reduced to 30 kHz. Thus, it seems that

BP100 at low concentration always has the same helix orientation in the bilayer, independently of the lipid system. Interestingly, in POPC:POPG this orientation does not change with concentration, while in DMPC:DMPG and, in particular, in the lyso-lipid containing sample there is a change in helix orientation and/or dynamics with increasing peptide concentration.

A comparison of the splittings measured at sample inclinations of 0° and 90° shows that the splittings at 90° are always close to $\frac{1}{2}$ of the splittings at 0° . This relationship indicates that the peptides are engaged in fast lateral diffusion around the membrane normal, which leads to an averaging on the NMR time scale. This rapid rotation is seen even at a very high concentration of the peptide ($P/L = 1:10$), in all lipid systems, suggesting a high mobility of the short peptide in all cases.

For α -helical peptides it is possible to roughly estimate the helix tilt angle in a membrane using ^{15}N -NMR on a ^{15}N -labelled peptide backbone, using an oriented sample at 0° inclination. (Aisenbrey, Michalek, Salnikov, & Bechinger, 2013; Bechinger, Gierasch, Montal, Zasloff, & Opella, 1996; Strandberg et al., 2013) Thus, ^{15}N -NMR experiments were performed on BP100 with a ^{15}N -Leu8, prepared in the same way as those used for the extensive ^2H -NMR analysis. A P/L ratio of 1:20 was chosen, not only because of sensitivity issues, but also because here the ^2H quadrupolar splittings showed interesting variations in the different lipid systems. The corresponding ^{15}N -NMR spectra are shown in Figure 3.14. In POPC:POPG, BP100 gives a narrow signal at 88 ppm, corresponding to peptides in a surface-bound orientation (a so-called “S-state”) with the helix tilt angle τ close to 90° (Bechinger et al., 1991). In DMPC:DMPG, the signal is shifted to 95 ppm, indicating that the peptide is slightly tilted in the lipid bilayer. In DMPC:DMPG:lyso-MPC, the signal is found at 137 ppm, indicating a more tilted peptide. These ^{15}N -NMR results support the changes observed in ^2H -NMR and already give a rough idea on the helix re-alignment, but they cannot provide more detailed information on the helix tilt angle, the azimuthal rotation angle, or peptide dynamics.

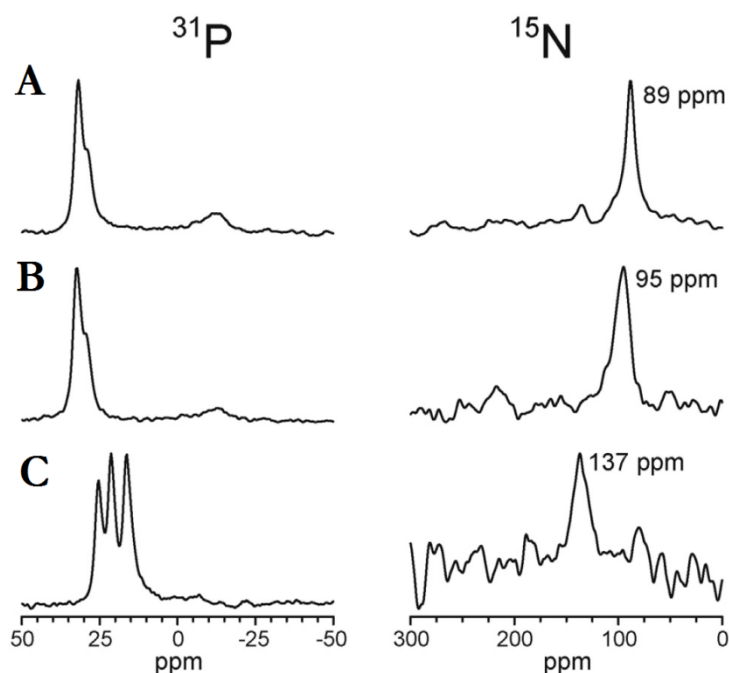


Figure 3.14. ^{31}P and ^{15}N NMR spectra of oriented samples of BP100, labeled with ^{15}N at the backbone amide of position Leu-8, at P/L = 1:20 in bilayers of POPC:POPG (3:1) (A); DMPC:DMPG (3:1) (B); and DMPC:DMPG:lyso-MPC (1:1:1) (C). The ^{15}N chemical shift of the main signal is given.

To obtain an accurate description of BP100 in the different lipid systems, the ^2H -NMR data of the full series of eleven Ala- d_3 labelled analogues at P/L = 1:20 were analysed. The spectra measured at 0° sample inclination are shown in Figure 3.15. All samples gave good signals, but for BP100L3A and BP100I7A a broad central peak is seen, such that a splitting cannot be resolved. The two samples were prepared anew and measured again, but did not give better resolved spectra. These spectra can be compared with that of BP100K2A, which has only a sharp central peak, and where the splitting is clearly close to zero. By subtracting this spectrum from those of BP100L3A and BP100I7A, the corresponding splittings could be estimated in these cases. The splittings of all samples are listed in Table 3.7. BP100K5A had already been measured in the concentration series described above, but as a control the samples for BP100K5A were prepared afresh and measured again, hence both splittings are given in Table 3.7. From these repeated samples, the experimental error in the splittings is estimated to be ± 1 kHz, and the average of the two splittings of BP100K5A was used in the data analysis.

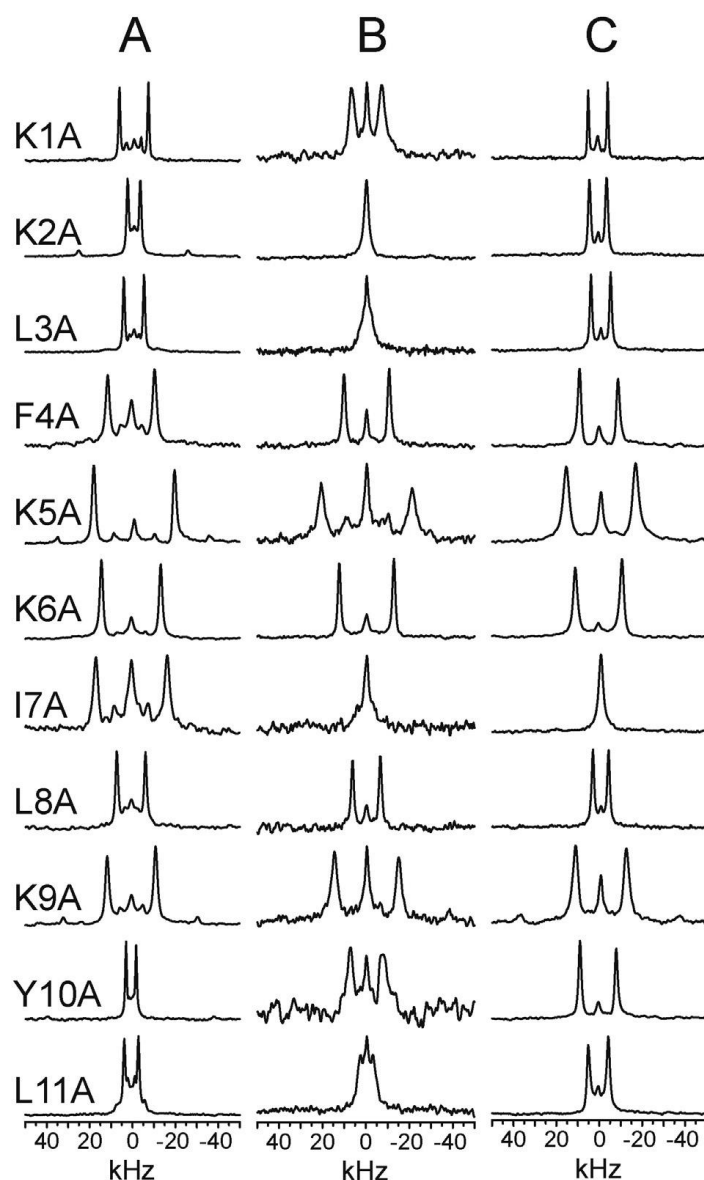


Figure 3.15. ^2H NMR spectra of oriented samples of BP100, labeled with Ala- d_3 at the positions indicated in each row, at P/L=1:20 in bilayers of POPC:POPG (3:1) (A); DMPC:DMPG (3:1) (B); and DMPC:DMPG:lyso-MPC (1:1:1) (C).

Table 3.7. ^2H quadrupolar couplings (in kHz) of Ala- d_3 labelled BP100 analogues in different oriented lipid bilayers, at P/L = 1:20, measured with the membrane normal parallel to the magnetic field.

Peptide	Lipid system		
	POPC:POPG (3:1)	DMPC:DMPG (3:1)	DMPC:DMPG:lyso-MPC (1:1:1)
BP100K1A	13.5	12.0	9.1
BP100K2A	5.9	0.0	8.0
BP100L3A	9.4	5.0	9.2
BP100F4A	21.9	20.8	17.9
BP100K5A	37.6/39.7	41.8/39.6	32.3/33.0
BP100K6A	26.7	25.2	21.8
BP100I7A	33.3	7.0	0.0
BP100L8A	13.4	12.8	7.3
BP100K9A	22.6	29.6	23.7
BP100Y10A	4.7	14.8	16.9
BP100L11A	6.6	5.6	9.3

4. ORIENTATION OF BP100 IN LIPID BILAYERS

From the ^2H -NMR quadrupolar splittings, the helix orientation and peptide dynamics in the lipid bilayer were evaluated, as described in detail in previous publications of the group. (Glaser et al., 2004; Strandberg et al., 2008; Strandberg et al., 2004; van der Wel et al., 2002) To obtain reliable structural results on peptides bound to liquid crystalline membranes, dynamics must be taken into account in the NMR data analysis. (Esteban-Martin, Strandberg, Fuertes, Ulrich, & Salgado, 2009; Esteban-Martin, Strandberg, Salgado, & Ulrich, 2010; Grage et al., 2012; Strandberg et al., 2009; Strandberg, Esteban-Martin, Ulrich, & Salgado, 2012) The short BP100 helix with a compact shape can be expected to be even more mobile than the longer α -helical peptides studied before; hence the dynamical model is particularly important here. The ^2H -NMR data were analysed using an “explicit” dynamical model, where helix motions are explicitly described in terms of a whole-body wobble, which leads to fluctuations of the τ and ϱ angles. It can be assumed that the angles fluctuate fast on the timescale of the NMR experiment, such that the ^2H quadrupolar splittings are averaged by these fluctuations. The data analysis model assumes that fluctuations can be described with Gaussian probability distributions of angles, i.e. by an average value τ_0 (or ϱ_0) and by a width σ_τ (or σ_ϱ) of the distribution. (Strandberg et al., 2009) These types of Gaussian distributions have been observed in several molecular dynamics simulations of peptides in membranes, (Esteban-Martin & Salgado, 2007; Ulmschneider, Smith, Ulmschneider, Ulrich, & Strandberg, 2012; Y. Wang et al., 2014) and can thus be assumed to be a good approximation of the real peptide behaviour.

Table 3.8. Best-fit orientations of BP100 in lipid bilayers from ^2H NMR data at P/L=1:20. For BP100K5A average splittings from two experiments were used. Different combinations of splittings were included in the fit as indicated by the positions used.

Lipid	Positions used	τ (°)	ϱ (°)	σ_τ (°)	σ_ϱ (°)	RMSD (kHz)
POPC:POPG (3:1)	1-11	85	168	0	22	4.3
	1-6,8-11	83	169	4	25	2.5
	1-10	85	168	0	22	4.5
	2-11	86	166	0	22	3.2
DMPC:DMPG (3:1)	1-11	76	168	21	23	3.7
	1-10	69	169	30	19	2.9
	2-11	74	167	27	21	3.6
DMPC:DMPG:lyso-MPC (1:1:1)	1-11 ^a	42	166	41	26	2.3
	1-11 ^b	72	166	20	33	3.2

^a Dynamics parameters not restricted; ^b σ_τ restricted to values up to 20°.

In POPC:POPG, a fit of the splittings from all 11 positions gave an orientation defined by $\tau = 85^\circ$ and $\varrho = 168^\circ$ (Table 3.8). The helix wobble is described by a very small variation of τ , $\sigma_\tau = 0^\circ$, and a larger variation of ϱ , $\sigma_\varrho = 22^\circ$. Similar values had been previously found in POPC:POPG for MSI-103, a somewhat longer amphiphilic helix with 21 amino acids. (Strandberg, Tiltak, et al., 2012) However, the RMSD of 4.3 kHz between the experimental and calculated splittings of BP100 was large. A detailed analysis showed that the splitting from position 7 did not fit with the other data points (Figure 3.16A and 3.15B). A second sample was prepared with this labelled peptide, but it produced the same splitting. A fit without position 7 gave essentially the same orientation and dynamics, but

with a better RMSD of 2.5 kHz (Table 3.8), which is similar to the values found previously for other peptides. (Strandberg, Tiltak, et al., 2012; Strandberg et al., 2006) On the other hand, removing the splitting from position 1 or position 11, at the termini, did not improve the fit much (Table 3.8). In Figure 3.16 we show the result for BP100. All positions, from 1 to 11, had experimental splittings that were well matched by the best-fit curve, with the exception of position 7, as mentioned above (Figures 3.16A and 3.16B). A plot of the RMSD as a function of the values of τ and ϱ used in the calculations shows a single, well-defined minimum (Figure 3.16C), indicating that there is no other orientation that could give rise to the measured splittings. A corresponding plot of the RMSD as a function of dynamics, in the form of σ_τ and σ_ϱ values used in the calculations, shows a broader minimum in which σ_ϱ is better defined than σ_τ (Figure 3.16D).

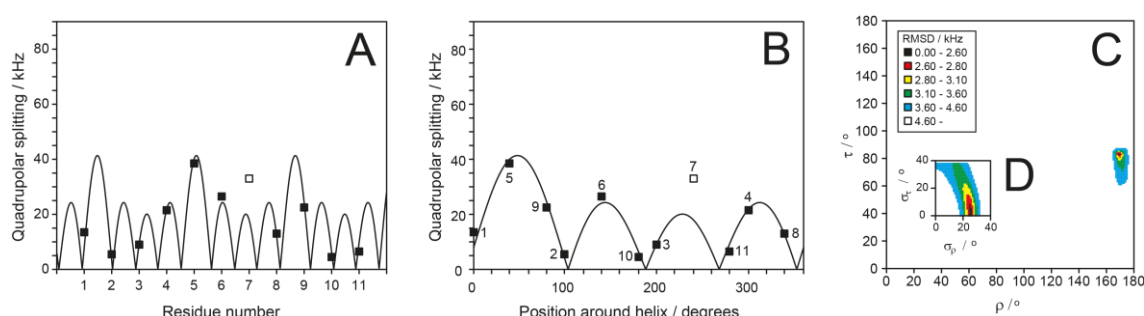


Figure 3.16. Analysis of ^2H NMR data of BP100 in POPC:POPG (3:1) bilayers at P/L=1:20 using the explicit dynamical model. (Data from position 7 was not used in the fit, see text). **A.** Best-fit helical wave plotted against the experimental quadrupolar couplings measured at different labelled positions in the helix. **B.** The same helical wave plotted around one turn of the helix, with labelled positions marked at the angle corresponding to the position around the helix axis, according to a helical wheel view (see Figure 3.8). **C.** RMSD between experimental and calculated splittings, for all τ and ϱ values used in the calculation. For each τ - ϱ pair, the best-fit values of σ_τ and σ_ϱ are used, which can be different for different pairs of angles. The RMSD is color-coded, and the lowest RMSD is marked in black. **D.** RMSD between experimental and calculated splittings, for all σ_τ and σ_ϱ values used in the calculation. For each σ_τ - σ_ϱ pair, the best-fit values of τ and ϱ are used. The RMSD is colour-coded, using the same scale as in **C**.

When the splittings measured in DMPC:DMPG are analysed, a different helix orientation is found for BP100, with the same ϱ value but a smaller $\tau = 76^\circ$ (Figures 3.17C and 3.17D). This result is supported by the ^{15}N -NMR data above. The value of σ_τ (21°) indicates a larger fluctuation of the helix tilt angle than in POPC:POPG, while the variation in ϱ is similar ($\sigma_\varrho = 23^\circ$) (Table 3.8). However, this fit based on all 11 data points is not very good, with an RMSD of 3.7 kHz. Closer inspection showed that this time the splitting of the C-terminal residue at position 11 did not fit well (Figures 3.17A and 3.17B). When this data point was not used in the fit, the RMSD was reduced to 2.9 kHz, while the helix tilt angle changed to 69° and dynamics increased. It can be noted from the fitting plots (Figures 3.17C and 3.17D) that the fit in this case gave a less well-defined τ , while ϱ was better defined and almost unchanged from the value in POPC:POPG. The dynamics is also in this case harder to pinpoint.

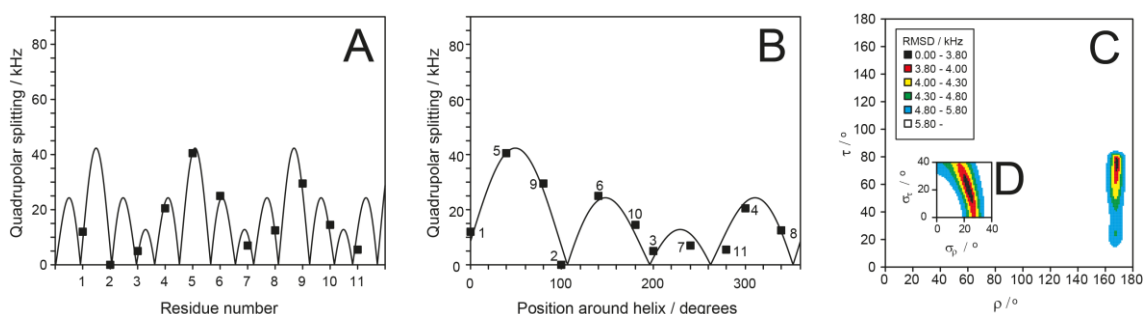


Figure 3.17. Analysis of ^2H NMR data of BP100 in DMPC:DMPG (3:1) bilayers at P/L=1:20. **A.** Best-fit helical wave plotted against quadrupolar couplings measured at different labelled positions in the helix. **B.** The same helical wave plotted around one turn of the helix. **C.** RMSD between experimental and calculated splittings, for all τ and ρ values used in the calculation. **D.** The lower insert gives the RMSD plot (experimental versus calculated splittings) for all σ_τ and σ_ρ values used in the calculation. The RMSD is color-coded, using the same scale in **C** and **D**.

In the case of DMPC:DMPG:lyso-MPC, a fit using all 11 data points gave a good RMSD of 2.3 kHz (Figures 3.18E and 3.18F, Table 3.8). This low value could mean that the peptide structure was closer to an ideal α -helix across the full length of the sequence (compared to a slightly unravelled situation in the other lipid systems). The value of ρ was the same as in the other lipid systems, but the tilt angle τ was even smaller (42°), and also less well defined. This could partly be due to the high variability in τ seen in the dynamic analysis ($\sigma_\tau = 41^\circ$), while the variation of ρ was similar to the other lipid systems ($\sigma_\rho = 26^\circ$). The corresponding plots (Figures 3.18A and 3.18B) show a very good fit of all data points, but the helix tilt angle is not so well defined, and the dynamic parameters can take several possible values with similar RMSD. In fact, the value of σ_τ appears to be unrealistically high, hence another fit was performed in which σ_τ was limited to a maximum of 20° (Figures 3.18E, 3.18F, 3.18G, and 3.18H). Then the minimum in the τ - ρ -plot is much better defined, but it gives a tilt of 72° , which is no longer compatible with the ^{15}N -NMR data. The original value of 42° would be in line with the ^{15}N chemical shift, but a helix tilt of 72° should give a chemical shift of around 100 ppm, as in DMPC:DMPG. On the other hand, a smaller tilt angle of 42° is not so well compatible with small values of σ_τ .

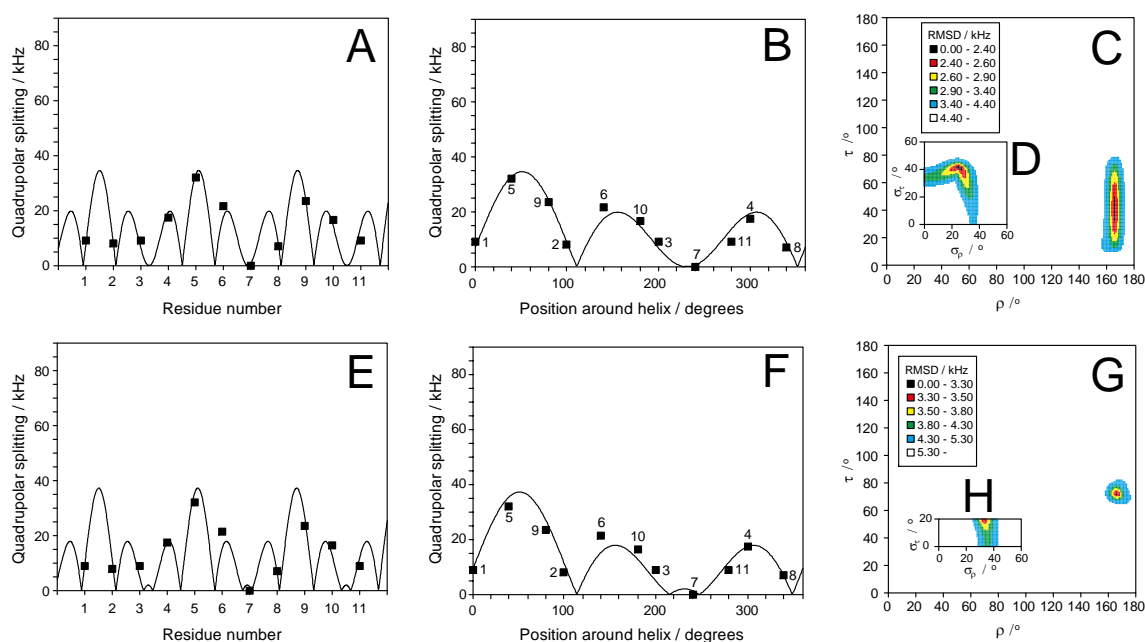


Figure 3.18. Analysis of ^2H NMR data of BP100 in DMPC:DMPG:lyso-MPC (1:1:1) bilayers at P/L=1:20. **A.** Best-fit helical wave plotted against quadrupolar couplings measured at different labelled positions in the helix. **B.** The same helical wave plotted around one turn of the helix. **C.** RMSD between experimental and calculated splittings, for all τ and ρ values used in the calculation. **D.** The lower insert gives the RMSD (experimental versus calculated splittings) for all σ_τ and σ_ρ values used in the calculation. The RMSD is colour-coded, using the same scale in **C** and **D**. **E–H.** same as **A–D**, but restricting the dynamics to $\sigma_\tau \leq 20^\circ$.

5. DISCUSSION

The first aim of the Ala- d_3 scan was to ascertain the importance of each residue in BP100 with regard to its biological function, in order to find peptides with improved therapeutic index. The second aim was to finally obtain an accurate picture of the orientation and dynamics of this short peptide in membranes, since the previous ^{19}F -NMR analysis performed had remained ambiguous due to geometrical reasons. (Wadhvani et al., 2014)

In the absence of lipids, all peptides adopt random coil conformations as the parent peptide. The replacement of any residue of BP100 with Ala led to a higher degree of α -helicity compared to the parent peptide, especially when the substitution was made on Lys residues. This finding might be attributed to variations in the membrane-binding affinity of the different analogues. Highly charged peptides are electrostatically attracted to the membranes, but they can saturate the surface and thereafter generate electrostatic repulsion. However, CD measurements were performed in DMPC:DMPG (3:1) vesicles at a P/L ratio of 1:50, hence there are 12.5 negative charges on the vesicles for each peptide, which is enough to bind them all. (Wadhvani et al., 2014) It thus seems that the most hydrophobic peptides, with less charge, can insert more deeply into the membrane and therefore have higher helix content, which is favoured in a more hydrophobic environment. (White & Wimley, 1999; Zelezetsky & Tossi, 2006)

BP100 was designed to have a high antimicrobial activity and low cytotoxicity. (Badosa et al., 2007) In the original study, it was developed to be active against plant pathogens, hence only some relevant Gram-negative bacteria were tested. (Badosa et al., 2007) BP100 was later shown to be even more active against Gram-negative than Gram-positive bacteria (Torcato et al., 2013), which is also noted in this investigation (Table 3.3), while another study found a higher activity against Gram-positive bacteria. (P. Wadhvani, R. F. Epanand, et al., 2012) The charge of peptides is known to be important for the antimicrobial activity, as increasing positive charge has been shown to give a higher activity, up to a point. (Zelezetsky & Tossi, 2006) For instance, in a 19-mer α -helix, increasing the charge up to +8 increased the activity, but a further increase to +9 led to a lower helicity, and any additional charges did not improve activity, suggesting that about 50 % charged residues may be optimal for activity. (Giangaspero, Sandri, & Tossi, 2001) In the 11-mer BP100, it seems that a charge of + 5 and + 6 gives the same antimicrobial activity, because replacing a single Lys with an Ala does not reduce the activity (Table 3.3). This is true for all Lys positions, which indicates that none of them is individually critical for the function, and four instead of five Lys residues still give an equally active peptide. In the original design of BP100, the starting sequence (Pep3: **W**KL**F**KKIL**K**V**L**-NH₂, bold residues differ from BP100) with 4 Lys was less active (R. Ferre et al., 2006), just like a similar peptide (BP105: **L**KL**F**KKIL**K**Y**L**-NH₂, bold residue differs from BP100). (Badosa et al., 2007) Yet another derivative (BP16: **K**KL**F**KKIL**K****K**L-NH₂, bold residue differs from BP100) with 6 Lys was inactive (R. Ferre et al., 2006), so it seems that activity will not improve by further reducing or increasing the charge.

Given the substantial reduction in antimicrobial activity that was observed when hydrophobic residues were replaced with Ala, it can be concluded that hydrophobicity is crucial for function in BP100 (Table 3.3). Haemolysis was also found to be much more sensitive to Ala-substitutions of hydrophobic residues than of charged ones (Figure 3.10 and Table 3.4). It has previously been reported that haemolysis can either increase or decrease with the hydrophobicity of a peptide. (Dathe et al., 1997; Strandberg et al., 2008; Wieprecht, Dathe, Beyermann, et al., 1997; Wieprecht, Dathe, Krause, et al., 1997; Zelezetsky & Tossi, 2006) In the case of BP100, the total hydrophobicity is augmented when Lys is replaced by Ala, but it is reduced when the large hydrophobic residues are substituted. Therefore, for BP100, haemolysis and hydrophobicity are not perfectly correlated since a reduction of the hydrophobicity leads to a decrease in haemolysis, as expected, but an increase of the hydrophobicity does not augment haemolysis. In a previous study of MSI-103, another amphipathic AMP, only residues on the hydrophobic face of the peptide had been replaced, and in all these cases haemolysis augmented with increasing hydrophobicity, and diminished when hydrophobicity was decreased. (Strandberg et al., 2008) This behaviour has been observed in this study on the hydrophobic face of BP100. However, on the polar face, the behaviour is different, and haemolysis is not increased when a Lys is replaced with an Ala on the polar face. Combining the antimicrobial and haemolysis results, the best candidates for peptides with an increased therapeutic potential identified in this study are BP100K2A and BP100K9A.

The membrane alignment of a number of helical peptides has been previously characterised with solid state ^2H -NMR using selective deuterium labels in the side chains, with the so-called GALA approach (geometric analysis of labelled alanines). (Daily, Greathouse, van der Wel, & Koeppe, 2008; Grage et al., 2012; Ozdirekcan, Rijkers, Liskamp, & Killian, 2005; Strandberg, Esteban-Martin, et al., 2012; Strandberg et al., 2008; Strandberg et al., 2004; Strandberg, Tiltak, et al., 2012; Strandberg et al., 2006; Tremouilhac, Strandberg, Wadhwani, & Ulrich, 2006a, 2006b; van der Wel et al., 2002; Vostrikov, Grant, Daily, Opella, & Koeppe, 2008) In previous studies of the Prof. Ulrich's group, only hydrophobic residues or small ones, like Gly and Ser, were replaced with Ala- d_3 in order to preserve the charge and avoid perturbing the function. BP100 is the first case where charged residues have been replaced, and interestingly it seems that the function is not strongly affected by such replacement. Even the structure is not affected, as seen from the good quality of the fit along the whole length of the helix. Almost all data points fit to a continuous helical wave curve, irrespective of the residue replaced. Thus, Ala- d_3 may also be a useful NMR label for selectively replacing polar or charged residues in other membrane-active peptides.

The good fit to the helical wave curves seems to indicate that the whole peptide forms an almost ideal α -helix from position 1 to 11. A previous NMR study of BP100 bound to lipid vesicles found an α -helix from position 3 to 11, but NMR signals from Lys1 and Lys2 could not be assigned, so those two residues could also be part of the helix. (Manzini et al., 2014) It was noted that the helicity implied from NMR was higher than that found from CD under the same conditions, which might be due to vesicle aggregation affecting the CD spectra and problems to interpret the CD results of short peptides. (Manzini et al., 2014) Also, in the CD experiments the vesicles were present in a large excess of water, while in the NMR samples the degree of hydration is just near saturation. This could mean that in CD a larger proportion of peptides are not bound to the membranes and therefore they are not helical. Thus, the CD results might underestimate the helicity of the membrane-bound state.

In the present study, a series of experimental Ala- d_3 splittings for all 11 positions of the sequence were obtained, which gave a more reliable fit compared to previous studies. (Wadhwani et al., 2014) The ϱ angle was found to be very similar to the value calculated from the earlier ^{19}F -NMR analysis (166° – 169°), while the τ angle was somewhat different. The results are summarized in Figure 3.19.

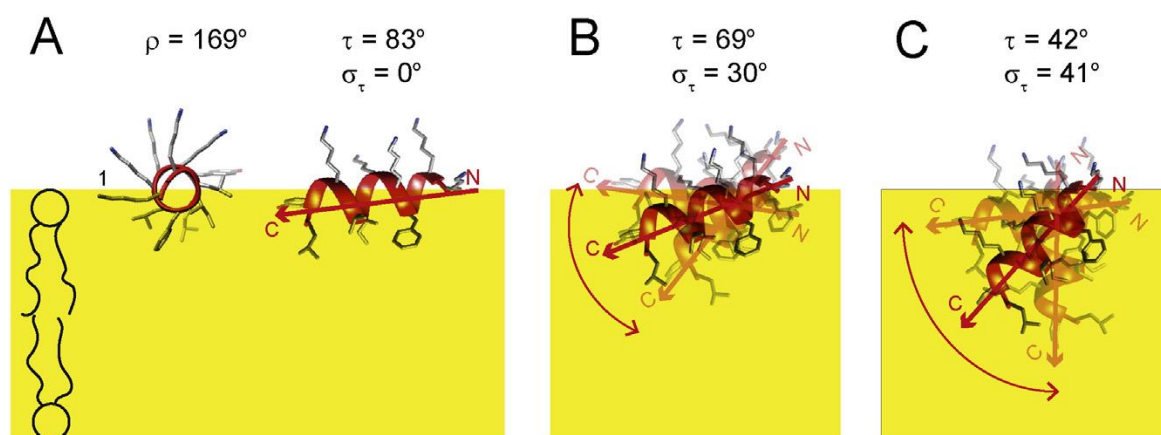


Figure 3.19. Orientation and dynamical behaviour of BP100 in lipid bilayers, as determined from solid state ^2H NMR. In POPC:POPG (3:1) (**A**), in DMPC:DMPG (3:1) (**B**), and in DMPC:DMPG:lyso-MPC (1:1:1) (**C**). The azimuthal angle ρ is in all cases the same, but the tilt angle τ and the dynamics parameter σ_τ varies depending on the lipid system. In all cases, the C-terminal is more deeply inserted in the membrane than the N-terminus, and the charged lysine side chains point towards the water phase. With increasing spontaneous curvature of the lipids ($A < B < C$) the peptide is able to insert more deeply and becomes more mobile.

In POPC:POPG, a system previously not used for BP100, an essentially flat helix orientation with a tilt angle of 83° was obtained (Figure 3.19A). This picture fits well with studies of other helical AMPs in POPC:POPG and POPC bilayers, which all showed flat orientations (S-state). (Strandberg & Ulrich, 2015; Strandberg et al., 2013) In particular, MSI-103 had been comprehensively studied using ^2H -NMR in POPC, POPE:POPG and POPE:POPC, and in all cases the same S-state was found, with $\tau = 93\text{--}95^\circ$. (Strandberg et al., 2008; Strandberg, Tiltak, et al., 2012) In general, it has been observed that in lipids with a negative spontaneous curvature, such as POPC:POPG, amphipathic peptides are always oriented flat on the membrane surface. (Strandberg & Ulrich, 2015)

In lipids with a positive spontaneous curvature, such as DMPC, amphipathic peptides tend to be oriented flat on the membrane surface at low concentration, but at higher concentration they can attain a more tilted orientation, or even become fully inserted in a transmembrane alignment (corresponding to an oligomeric pore). (Glaser et al., 2005; Strandberg et al., 2008; Strandberg, Tiltak, et al., 2012; Strandberg & Ulrich, 2015; Strandberg et al., 2006; Tremouilhac et al., 2006a; Wadhwani et al., 2008) It thus seems that the likelihood of more inserted orientations is higher in lipid systems with a larger positive spontaneous curvature, like those containing lyso-lipids. This behaviour was also observed here for BP100. However, a very high peptide concentration was needed for BP100 to reach a tilted state. In the case of PGLa, a tilted state was observed in DMPC already at $P/L = 1:100$, and for MSI-103 a tilted state was seen even at a concentration as low as $1:200$. (Bürck et al., 2008; Strandberg et al., 2008; Tremouilhac et al., 2006a) In contrast, for BP100 the S-state was retained at $P/L = 1:50$ even in DMPC:DMPG:lyso-MPC, and a tilted state only appeared at $1:20$. Compared to the other amphiphilic peptides studied so far, BP100 thus seems to have the highest threshold concentration to undergo re-alignment and insert into the membrane.

In DMPC:DMPG at P/L = 1:20, an orientation of BP100 that is seemingly similar to that previously found using ^{19}F -NMR and ^{15}N -NMR was observed, i.e. with a largely surface-bound helix that is only slightly tilted into the bilayer by around 20° (Figure 3.19B). In the previous ^{19}F -NMR studies, an ambiguous tilt angle τ of around 110° in combination with $\varrho \approx 160^\circ$ were reported, while here the more precise combination of $\tau \approx 70^\circ$ and $\varrho \approx 170^\circ$ was calculated from ^2H -NMR data. The values of both tilt angles (which are defined relative to the membrane normal, see Figure 3.20) are fully compatible with the ^{15}N -NMR data, but the actual τ - ϱ combinations contradict one another with regard to which terminus of the helix is inserted more deeply into the membrane. Notably, ^{15}N -NMR can only provide the tilt angle of the helix axis with respect to the membrane normal, but it cannot discriminate between the pseudo-symmetrical situation of 110° being seemingly equivalent to 70° ($= 180^\circ - 110^\circ$). On the other hand, in ^2H -NMR and ^{19}F -NMR the three-dimensional arrangement of the reporter groups around the helix backbone provides a strict criterion to decide how a helix gets tilted into the membrane. It can be discriminated whether the N-terminus or the C-terminus gets inserted first into the hydrophobic bilayer, by taking into account the azimuthal rotation of the polar and hydrophobic faces of an amphiphilic peptide. The orientational parameters and their meanings are explained in Figure 3.20. In brief, the combination of $\tau \approx 110^\circ$ and $\varrho \approx 160^\circ$ from the old ambiguous ^{19}F -NMR data had implied the N-terminus to go in first. In contrast, the new reliable ^2H -NMR data with $\tau \approx 70^\circ$ and $\varrho \approx 170^\circ$ (which is equivalent to the combination of $\tau \approx 110^\circ$ and $\varrho \approx -10^\circ$) demonstrate clearly that the C-terminus of BP100 gets inserted more deeply into the membrane. The latter, more trustworthy scenario found here by ^2H -NMR was also observed in a recent molecular dynamics simulation, which gave a tilt angle around 70° with the C-terminus dipping in (Y. Wang et al., 2014).

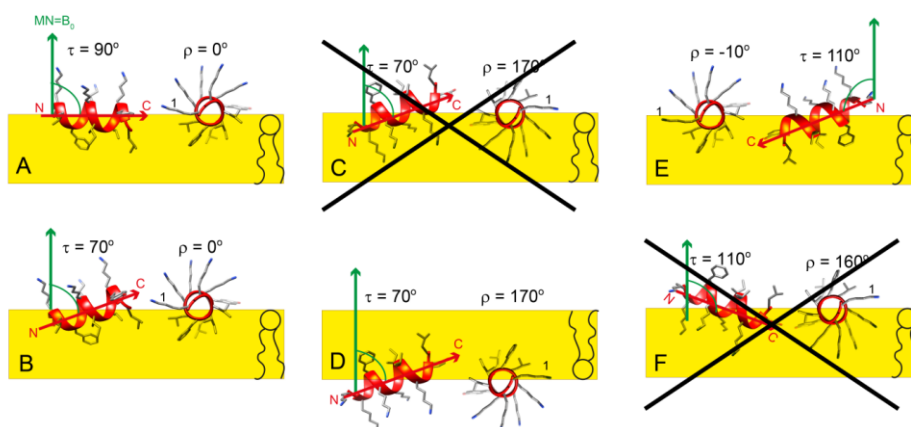


Figure 3.20. Definition of peptide orientation and meaning of the angular parameters. The tilt angle τ is defined between the membrane normal (MN, which is along the external magnetic field B_0) and the peptide axis, directed from the N- to the C-terminus. The azimuthal angle ϱ describes the right-handed rotation around the helix axis. **A.** A completely flat orientation ($\tau = 90^\circ$) on the membrane surface. **B.** A slightly tilted orientation with $\tau = 70^\circ$, and still aligned with the azimuthal starting point of $\varrho = 0^\circ$. **C.** The orientation found here by ^2H -NMR in DMPC:DMPG is illustrated for the combination of $\tau = 70^\circ$ and $\varrho = 170^\circ$. **D.** With the same combination of $\tau = 70^\circ$ and $\varrho = 170^\circ$ the peptide can be just as well located on the other side of the membrane, with lysines pointing out of the membrane as expected. **E.** A rotation of panel **D** by 180° , with the combination of $\tau = 110^\circ$ and $\varrho = -10^\circ$, which is equivalent to $\tau = 70^\circ$ and $\varrho = 170^\circ$. **F.** The previously reported combination of $\tau \approx 110^\circ$ and $\varrho \approx 160^\circ$. It is now superseded by the new ^2H -NMR results in panels **D** and **E**.

It is, in fact, much more energetically probable that the uncharged (amidated) C-terminus can get inserted more deeply into the membrane with the charged lysines pointing towards the polar environment. An insertion of the doubly charged N-terminus of Lys1, as suggested in the previous ^{19}F -NMR study, is clearly a less favourable situation than the one shown here. It can be concluded that the ^{19}F -NMR value was incorrect, probably due to the uncertainty in the dynamical fit from the insufficient access to suitably ^{19}F -labelled positions. As comprehensively discussed in that paper, when having access to NMR reporter groups only on one face of the peptide (i.e. when using the hydrophobic CF_3 -Bpg labels), this does not necessarily lead to ambiguous results. However, under certain conditions ambivalent solutions can arise, and the ^{19}F -NMR study of BP100 was challenged by such unfavourable conditions. (Wadhvani et al., 2014) Nonetheless, it is important to emphasize that ^{19}F -NMR analysis based on CF_3 -Bpg or CF_3 -Phg labels has the same intrinsic reliability as ^2H -NMR based on Ala- d_3 , because in all three cases the reporter groups are attached to the peptide backbone with the same local geometry. In many other peptides, the ^{19}F -NMR approach has indeed yielded steadfast results with unprecedented sensitivity, which are fully supported by independent data from OCD, ^2H -NMR, and ^{15}N -NMR. (Afonin, Mikhailiuk, Komarov, & Ulrich, 2007; Bürck et al., 2008; Strandberg et al., 2008; Strandberg et al., 2006) The current limitations of ^{19}F -NMR can be overcome as soon as some hydrophilic and/or charged ^{19}F -labelled amino acids are available as reporter groups. (Tkachenko et al., 2013)

It has been proposed that the MIC of AMPs against bacteria is related to P/L^* , the threshold P/L ratio at which the peptide shows disruptive activity in model membranes, and some calculations have been published for BP100 in this regard. (M. N. Melo & Castanho, 2012) In the present study, a gradual re-orientation of BP100 is observed as a function of P/L. When BP100 is lying flat on the membrane surface it most likely cannot induce much damage, but when it gets more inserted this can be a sign of membrane disrupting activity. The critical P/L^* was estimated from the NMR results. In Figure 3.13, the ^2H -NMR spectra of all lipid systems show the same splitting at $\text{P/L} = 1:100$, which can be taken as an inactive state. In POPC:POPG (3:1) no change is seen up to $\text{P/L} = 1:10$. In DMPC:DMPG (3:1) there is a visible change between 1:20 and 1:10, for which we can take 1:15 as a threshold P/L^* . In DMPC:DMPG:lyso-MPC (1:1:1) a change is seen between 1:50 and 1:20, so we can estimate P/L^* to be around 1:35. To estimate MIC from P/L^* , a binding constant K_p is needed, which has been characterised for BP100 and varies between vesicles according to charge. (Rafael Ferre et al., 2009; Torcato et al., 2013) NMR samples characterised here do not contain any excess of water, so peptides are more likely to be completely bound in these oriented bilayers compared to an aqueous suspension of vesicles of the same lipid composition. Even if the K_p values from vesicles might not be fully applicable, as a first approximation it can be assumed that binding is at least as good as to vesicles of POPC:POPG (1:2), which had been studied previously. (M. N. Melo & Castanho, 2012) A MIC value of around $3\text{ }\mu\text{M}$ from the P/L^* in DMPC:DMPG can be predicted, and $1.2\text{ }\mu\text{M}$ from the P/L^* in DMPC:DMPG:lyso-MPC. This is slightly lower than the MIC values determined against *E. coli* or *P. aeruginosa* (Table 3.3). Of course, a minor reorientation of BP100 may not perturb the membranes sufficiently to inhibit bacterial growth, but at higher concentrations the disturbance increases, together with the

lateral pressure in the outer lipid monolayer. If, for example, a four-fold higher P/L^* was used, this would give 4-fold higher predicted MIC values, which are still in a reasonable range. The reorientation and increase in dynamics observed for BP100 in the model systems used here would be indeed related to the bacterial killing action found for this peptide.

If BP100 is compared to the other peptides mentioned above (PGLa and MSI-103) the value of P/L^* as defined from peptide reorientation in the membrane is much higher for BP100, which should correspond to a higher MIC. However, the MIC of BP100 is actually lower than for PGLa and only slightly higher than for MSI-103. (P. Wadhwani, R. F. Epand, et al., 2012) This distinction may be due to a different mechanism of action, because PGLa and MSI-103 have been proposed to form transmembrane pores. By contrast, BP100 is too short to traverse the lipid bilayer and must have a different mode of action, which can shift the relationship between P/L^* and MIC. Indeed, PGLa has a higher P/L^* than MSI-103, and it is also less antimicrobially active (Blazyk et al., 2001; P. Wadhwani, R. F. Epand, et al., 2012), so for peptides with the same mechanism of action their relative activity seems to correlate with P/L^* .

When the orientation of BP100 in a lipid system with larger positive spontaneous curvature, such as DMPC:DMPG:lyso-MPC, is studied, a pronounced helix tilt is observed using ^{15}N -NMR. From the ^2H -NMR analysis, a tilt angle of about 40° away from the membrane normal was found, with the uncharged C-terminus inserted deep into the membrane core (Figure 3.19C). A very large Gaussian value of σ_τ (41°) was observed in this case (Figure 3.18C), which might seem unrealistically large. In the MD simulation mentioned above, a smaller distribution width of $10\text{--}15^\circ$ was observed, but this simulation was done in DMPC, and only a single peptide was present in the membrane. (Y. Wang et al., 2014) In a soft DMPC:DMPG:lyso-MPC membrane, BP100 may well behave differently and indeed seems to be unusually mobile compared to other, longer peptides. Even at a high concentration ($P/L = 1:10$) BP100 retained its rapid lateral diffusion (see Figure 3.13), which was not the case for longer peptides, e.g. in the 21-mers MSI-103 (Strandberg et al., 2008) or PGLa. (Tremouilhac et al., 2006a) As explained before, and after performing different fits, the strong fluctuations detected in the τ angle can be considered to be realistic, and this is a possible explanation: The highly charged N-terminus (with a charge of +3 on the two first residues) would be very unlikely to get immersed into the hydrophobic part of the membrane. Therefore, in POPC:POPG, when the surface-bound helix is hardly tilted into the membrane ($\tau \approx 85^\circ$), the fluctuation in tilt angle is no more than about 10° . In DMPC:DMPG, where the helix is slightly tilted into the membrane by around 20° (i.e. $\tau \approx 70^\circ$ relative to the membrane normal), a larger variation of the tilt angle is possible over the range $0\text{--}40^\circ$. Finally, in DMPC:DMPG:lyso-MPC, where the helix is steeply tilted on average ($\tau \approx 40^\circ$), the fluctuation range may well be as large as $0\text{--}90^\circ$, i.e. the full range of tilt angles may be present from flat to completely inserted. In all the cases, the N-terminus remains on the membrane surface, while the C-terminus can get inserted into the hydrophobic core, presumably with the charged groups on the lysines snorkelling to stay in an environment as polar as possible. The ability of a peptide to penetrate into the lipid bilayer and the ensuing dynamical situation thus seem to

be regulated by the spontaneous curvature of the lipids, which either seal or soften the membrane for the peptide to enter more deeply.

Even with a tilt angle of 180° (i.e. upright in the membrane), BP100 is much too short to span the membrane and form a pore. It has been recently demonstrated that amphipathic helices need to be at least 17 residues long to assemble into oligomeric transmembrane pores that kill bacteria such as *E. coli* (Grau-Campistany et al., 2015), and 11 residues is clearly not enough. However, if all the BP100 molecules are bound to the membrane surface and are able to partly insert in a highly mobile and monomeric state as proposed above, it is likely that this will lead to an increased permeability of the membrane. Hence, it can be settled that the short amphiphilic BP100 helix utilizes this kind of carpet mechanism to induce membrane leakage, without the need to assemble into an oligomer or to form a transmembrane pore.

CONCLUSIONS

1. All the studied BP100 variants behave as the parent peptide, that is, they show a random coil conformation in aqueous solution, and adopt an α -helical structure in the presence of lipid vesicles.
2. Helix population increases when a Lys residue is substituted by Ala, suggesting that a higher hydrophobicity may favour the insertion into the membrane, with a subsequent augmentation of the helix population.
3. In BP100, not all the positive charges are necessary for the antibacterial activity, but certain hydrophobicity is required.
4. BP100 haemolytic activity is strongly reduced when a hydrophobic residue is replaced by Ala. Substitution of Lys residues lead to a reduction of the haemolysis that can be significant (Lys2 and Lys9), or minor (Lys1, Lys5, and Lys6).
5. BP100K2A and BP100K9A variants possess better therapeutic indexes than BP100, as they maintain the original antimicrobial ability but show a reduced haemolytic activity in the MIC concentration range.
6. Preliminary ^{15}N -ssNMR results on BP100 variants indicated a surface-bound orientation ($\tau \sim 90^\circ$) in POPC:POPG 3:1; a slightly tilted orientation in DMPC:DMPG 3:1; and more tilted orientation in DMPC:DMPG:lyso-MPC 1:1:1.
7. ^2H -ssNMR results are in agreement with ^{15}N -ssNMR at P/L = 1:50 and lower peptide concentrations. At higher peptide concentrations, BP100 can tilt into the bilayer when it is constituted by lipids with positive spontaneous curvature.
8. The P/L value at which BP100 starts to get inserted in the bilayer could be related to the MIC value.
9. BP100 has a high mobility, which is not compatible with the formation of an oligomeric complex.
10. A carpet model of action has been proposed for BP100, from where it is able to insert dynamically into the bilayer core and induce a substantial disorder that leads to an increased membrane permeability and ultimately, to bacterial cell death.

Part 2

**Crotalicidin, a rattlesnake venom
cathelicidin with antimicrobial and
antitumour activity**

INTRODUCTION

1. CATHELICIDINS

Together with defensins, cathelicidins are one of the two major AMPs families present in the host defence system of mammals. The common feature of cathelicidins is the presence of a conserved N-terminal prosequence (~100 amino acids) in the unprocessed form of these peptides (Figure 3.21). That prosequence is called “cathelin” domain. (Zanetti, 2004; Zanetti, Gennaro, & Romeo, 1995)

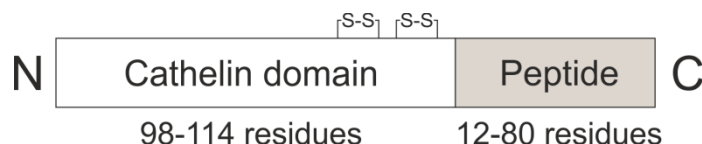


Figure 3.21. Schematic representation of a cathelicidin holoprotein.

Diverse cathelicidins have been isolated from numerous mammalian organisms, such as mice, rats, rabbits, guinea pigs, goats, sheep, pigs, horses, cattle, monkeys and humans. But they also have been described in other animals, like birds, amphibians, reptiles, fish, or even in living fossils, such as the jawless craniate hagfish. The presence of cathelicidins in such primitive vertebrates demonstrates that the genes of these peptides originated early in the course of biological evolution. (Falcao et al., 2014; Zanetti, 2004)

Cathelicidins are peptides containing typically 12–80 amino acids, and they show a wide range of structures that can be classified in three main groups: (Falcao et al., 2014; Zanetti, 2004)

- Linear peptides of 23–37 residues which fold into amphipathic α -helices in the presence of environments mimicking biological membranes (Figure 3.22A). They belong to AMPs class II (see section 2.1).
- Small-sized peptides (12 -18 amino acids) that form β -hairpin structures stabilized by one or two disulphide bonds (Figure 3.22B). They belong to AMPs class IV (see section 2.1).
- Larger peptides (39–80 residues) with an overrepresentation of one or two amino acids, such as proline or tryptophan (Figure 3.22C). They belong to AMPs class III (see section 2.1).

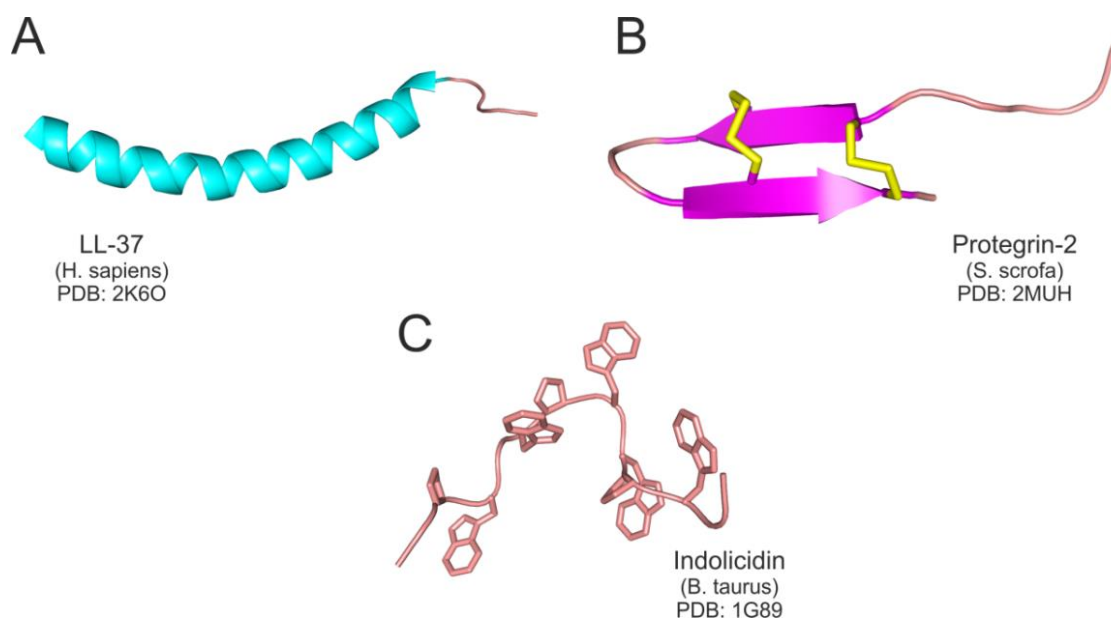


Figure 3.22. Example of the structural families of cathelicidins. For explanation, see the text. Cartoon representation follows this colour code: cyan for α –helices, magenta for β –strands, salmon for loops and unstructured regions. Disulphide bonds are highlighted in yellow. For every structure name, species and PDB code is indicated.

Cathelicidins *in vitro* activity against a variety of bacteria, fungi, and enveloped viruses has been published in several works, showing a broad overlap in specificity and important differences in potency. (Zanetti, 2004)

2. VIPERICIDINS

Cathelicidin–related peptides have been purified from the venom of some elapid snake species. Amongst those peptides, some show potent broad–spectrum antimicrobial and antitumour activities. Recent investigations of orthologous genes of cathelicidin–related peptides in some South American pit vipers have led to the identification of four novel venom–derived cathelicidins, collectively named “viperidins” (Table 3.9.). (Falcao et al., 2014)

Table 3.9. Viperidins identified in the venom of South American pit vipers and their characteristics.*

Peptide name	Source	ORF size (bp)	Peptide length	
			Prepro	Mature
Crotalicidin	<i>Crotalus durissus terrificus</i>	585	194	34
Lachesidicin	<i>Lachesis muta rhombeata</i>	585	194	34
Batroxicidin	<i>Bothrops atrox</i>	570	189	34
Lutzidicin	<i>Bothrops lutzii</i>	570	189	34

*Data from (Falcao et al., 2014).

Viperidins are closely related to the cathelicidin–related antimicrobial peptides (CRAMPs) found in the venom of Asian elapid snake species, and they have a potent antimicrobial activity against Gram–negative bacteria with low cytotoxicity. These properties make these peptides good candidates as antimicrobial agents. (Falcao et al.,

2014) In this work, one of this viperidins called “crotalycin” has been investigated, and its structure, and antimicrobial, antitumour, and cytotoxic properties explored minutely.

OBJECTIVES

In this part, crotallicidin –a cathelicidin from a South American pit viper– was subjected to a structural dissection in order to demarcate the minimal segments responsible for the peptide activity. This work was carried out in collaboration with Prof. Andreu's group, which was in charge of the biological part of the experimentation, as well as the CD spectroscopy.

Antimicrobial and antitumour properties have been reported in cathelicidins, and thus a thorough examination of crotallicidin is of potential interest to find peptide sequences with good therapeutic indexes. In this case, the microbiological assays performed on crotallicidin and its fragments were complemented by structural studies by solution NMR and CD, in order to unveil the function–structure relationship and better understand the mechanism of action of this kind of membrane–active peptides.

The specific objectives proposed in this part are:

- To perform a rational dissection of crotallicidin to define the minimally active regions of the peptide.
- To evaluate the antimicrobial and antitumour activity of crotallicidin and its fragments, as well as their cytotoxicity and stability.
- To investigate the function–structure relationship of crotallicidin and its fragments.

MATERIALS AND METHODS

1. BIOINFORMATIC ANALYSES

The Ctn sequence (Falcao et al., 2014) was processed with the “Peptide cutter” software (http://web.expasy.org/peptide_cutter/) using neutrophil elastase as model protease. Sequences of Ctn and its Ctn_{1–14} and Ctn_{15–34} fragments were also processed with the “Peptide property calculator” (<http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp>) and “Heliquet” software (Gautier, Douguet, Antonny, & Drin, 2008) (<http://heliquet.ipmc.cnrs.fr>) to determine molecular mass, net charge, hydrophobicity, and hydrophobic moment data (Table 3.10), as well as the helical–wheel plots (Figure 3.23).

Table 3.10. Primary structure and physicochemical properties of crotalicidin and fragments.

Peptide	Sequence ^a	Molecular mass ^b	Purity ^c	Charge ^d	Hydrophobicity, <i>H</i> ^e	Hydrophobic momentum, μH
Crotalicidin (Ctn)	KRFKKFFKKVKKSVKK RLKKIFKKPMVIGVTIPF	4151.39 (4151.41)	98 %	+16	0.263	0.440
Ctn _{1–14}	KRFKKFFKKVKKSV	1797.31 (1797.30)	96 %	+9	-0.012	0.763
Ctn _{15–34}	KKRLKKIFKKPMVIGVTIPF	2371.11 (2371.10)	98 %	+8	0.455	0.311

^aElastase putative cleavage sites are highlighted in green (http://web.expasy.org/peptide_cutter/). Peptides are C-terminal amides. ^bDetermined by LC–MS. Theoretical mass, in parentheses, from <http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp>. ^cBy analytical HPLC.

^dAt neutral pH, from <http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp>. ^eFrom <http://heliquet.ipmc.cnrs.fr/>.

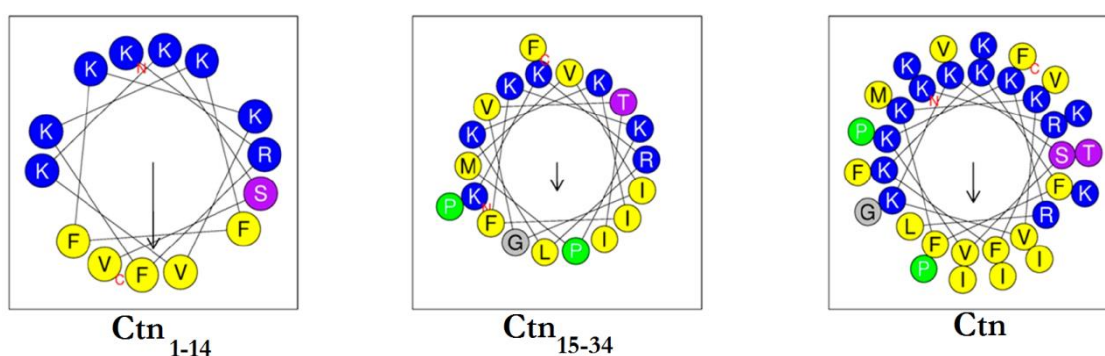


Figure 3.23. Helical –wheel projections for each peptide. Hydrophobic momenta are represented by arrows. Residue colour code: blue for positively–charged; yellow for hydrophobic; magenta for polar; Pro in green; Gly in grey. N– and C– termini are indicated in red.

2. PEPTIDE SYNTHESIS

Ctn, Ctn_{1–14} and Ctn_{15–34} were synthesized in C-terminal carboxamide form in an ABI433 (Applied Biosystems, CA, USA), running Fmoc (FastMoc) SPPS protocols. Synthesis products were purified by analytical reverse–phase HPLC using C18 columns

reaching purities over 95 %. The identity of the peptides was confirmed in all cases with a LC–MS Shimadzu 2010EV system (Kyoto, Japan).

Crotalicidin (Ctn) (KRFKKFFKKVKKS VKRLKKIFKKPMVIGVTIPF): RP–HPLC: $t_R = 10.5$ min, purity 98 % (buffer A: 0.045 % TFA in H₂O; buffer B: 0.036 % TFA in CH₃CN; linear B buffer gradient). HRMS: Theoretical molecular mass: 4151.39; found: 4151.41 [M + H]⁺.

Ctn_{1–14} (KRFKKFFKKVKKS V): RP–HPLC: $t_R = 5.4$ min, purity 96 % (buffer A: 0.045 % TFA in H₂O; buffer B: 0.036 % TFA in CH₃CN; linear B buffer gradient). HRMS: Theoretical molecular mass: 1797.31; found: 1797.30 [M + H]⁺.

Ctn_{15–34} (KKRLKKIFKKPMVIGVTIPF): RP–HPLC: $t_R = 7.1$ min, purity 98 % (buffer A: 0.045 % TFA in H₂O; buffer B: 0.036 % TFA in CH₃CN; linear B buffer gradient). HRMS: Theoretical molecular mass: 2371.11; found: 2371.10 [M + H]⁺.

3. CIRCULAR DICHROISM

CD spectra of the peptides were acquired in a J–815 spectropolarimeter (Jasco, Tokyo, Japan) at 25 °C in the 190–250 nm wavelength range, with a bandwidth of 1 nm and a scan speed of 50 nm·min^{–1}, using a 0.1 cm quartz cell. 70 µM peptide solutions were prepared in 10 mM sodium phosphate buffer (pH 7.4) in the absence or presence of 6 mM DMPG vesicles (Sigma–Aldrich, Madrid, Spain). Vesicles were prepared by dissolving DMPG powder in a chloroform/methanol (2:1) mixture, and then removing the solvents under reduced pressure with a N₂ flow, and drying the resulting lipid films overnight in a freeze–dryer. Next day, lipid films were suspended in 10 mM sodium phosphate using vortex shaker, and sonicated until clear.

The final spectra for each peptide were the average of three consecutive scans per sample after subtraction of buffer and vesicle baselines. Results were expressed as mean residue ellipticity ($[\theta]_{MRE}$) (deg·cm²·dmol^{–1}), as follows:

$$[\theta]_{MRE} = \frac{\theta_{obs} \cdot MRW}{10 \cdot d \cdot C}$$

where θ_{obs} is the observed ellipticity in degrees, MRW is the mean residue weight, d is the cell path length and C is the peptide molar concentration. The percentage of α –helix structure was estimated as:

$$\%_{\alpha\text{-helix}} = \frac{-[\theta]_{222\text{ nm}} + 3000}{39000}$$

where $[\theta]_{222\text{ nm}}$ is the mean residue ellipticity at 222 nm. (Rodrigues et al., 2012) (Torrent, Beatriz, Nogués, Andreu, & Boix, 2009)

4. NMR SPECTROSCOPY

NMR samples were prepared by dissolving the lyophilized peptides (1–2 mg) in 0.5 mL of a fresh solution of 30 mM $[D_{38}]$ -DPC (98 % deuteration; Cambridge Isotope Laboratories, Tewksbury, MA) at pH 3.0 in either H_2O/D_2O (9:1 v/v) or pure D_2O . Peptide concentrations were approximately 1 mM. DSS was added as an internal reference. pH was measured with a glass microelectrode and not corrected for isotope effects and adjusted, if necessary, by adding minimal amounts of NaOD or DCl.

NMR spectra were recorded in a Bruker AV-600 spectrometer operating at a 600.13 MHz proton frequency and equipped with a cryoprobe. Probe temperature was calibrated using a methanol sample. 1D and 2D spectra, i.e., phase-sensitive correlated spectroscopy (COSY), total correlated spectroscopy (TOCSY), nuclear Overhauser enhancement spectroscopy (NOESY), and 1H - ^{13}C heteronuclear single quantum coherence spectra (HSQC), were acquired using standard pulse sequences and processed with the TOPSPIN program, as reported. (Mirassou, Santiveri, Pérez de Vega, González-Muñiz, & Jiménez, 2009) TOCSY and NOESY mixing times were 60 and 150 ms, respectively. 1H - ^{13}C HSQC spectra were acquired at natural heteronuclear abundance. The ^{13}C δ -values were indirectly referenced using the IUPAC-recommended ^{13}C : 1H ratio 0.251 449 53. (Markley et al., 1998)

1H and ^{13}C assignment was achieved by standard sequential analysis (Wüthrich, 1986; Wüthrich, Billeter, & Braun, 1984) of 2D COSY, TOCSY, and NOESY spectra acquired at 25 and 35 °C, examined in combination with the corresponding 2D 1H - ^{13}C HSQC spectra. The ^{13}C chemical shift values served to confirm assignment of side chains, in particular the repeated Lys residues (13, 7, and 6 in Ctn, Ctn₁₋₁₄, and Ctn₁₅₋₃₄, respectively; see “Appendices”, Tables A11, A12, and A13). Chemical shifts for the three peptides have been deposited at the BioMagResBank (<http://www.bmrb.wisc.edu>) with accession codes 25363 (Ctn), 25366 (Ctn₁₋₁₄), and 25370 (Ctn₁₅₋₃₄).

5. NMR STRUCTURE CALCULATION

Structure calculations were performed with the CYANA 2.1 program (Güntert, 2004; Güntert, Mumenthaler, & Wüthrich, 1997). Upper limit distance restraints were obtained from the NOE cross-peaks present in 2D NOESY spectra recorded at 25 and 35 °C, which were integrated using the automatic integration subroutine of the SPARKY3 program. (Goddard & Kneller, 2004) Restraints for the ϕ and ψ dihedral angles were derived from the $^1H_\alpha$, $^{13}C_\alpha$, and $^{13}C_\beta$ chemical shifts using the TALOS program. (Cornilescu, Delaglio, & Bax, 1999) For Ctn and Ctn₁₋₁₄, the standard iterative protocol for automatic NOE assignment implemented in CYANA 2.1 (Güntert, 2004) was used. It consists of seven cycles of combined NOE assignment and structure calculation of 100 conformers per cycle. The distance restraints resulting from the seventh cycle were checked by re-examination of the NOESY spectra, and if necessary, ambiguous constraints were removed or relaxed to generate the final list used as input for a standard CYANA2.1 simulated

annealing calculation of 100 conformers. For each peptide, the final NMR structure corresponds to the ensemble of the 20 conformers with the lowest target function value. The Ctn structure has been deposited at the PDB Data Bank with accession code 2MWT. Statistics for the structural ensembles of the peptides are given in Table 3.12.

As automatic NOE assignment is not applicable to mainly random peptides, the structure of Ctn_{15–34} was calculated by a standard CYANA 2.1 protocol, using as distance restraints only the non-sequential ones (Tables 3.12 and 3.13). In poorly structured peptides, random conformers contribute to the intensity of intraresidual and sequential NOE signals. MOLMOL (Koradi, Billeter, & Wüthrich, 1996) was used to visualize the structures of the three peptides.

6. BACTERIAL STRAINS AND ANTIBACTERIAL ASSAYS

Minimum inhibitory concentration (MIC) assays were performed utilizing reference strains of *E. coli* (ATCC 25922), *P. aeruginosa* (ATCC 27853), *E. faecalis* (ATCC 29212), and *S. aureus* (ATCC 29213). Clinical strains of *E. faecalis*, *S. aureus*, *E. coli*, *P. aeruginosa*, *K. pneumoniae*, *A. baumannii*, and *S. pyogenes* were also used. Isolates were thawed and transferred at least twice on sheep blood agar to ensure purity and good growth and incubated for 24 h prior to testing. Inocula were prepared by direct suspension of cells into saline to achieve the turbidity of the 0.5 McFarland standard. The assays were performed by the microdilution method in Müller–Hinton broth according to Clinical and Laboratory Standards Institute (CLSI) guidelines. ((CLSI), 2012) For *S. pyogenes*, the MIC assay was performed in Müller–Hinton broth supplemented with lysed horse blood (5 %). Gentamicin was used as a positive control.

7. CELL CULTURE

Human HeLa S3 (cervix epithelial carcinoma), leukaemia Jurkat E6.1 (T-cell lineage), HL–60 (promyelocyte lineage), U937, THP–1, and MM6 (monocyte–macrophage lineage) cancer cells, and 1BR3G human fibroblasts were obtained from the Cell Line Repository of the Institut Municipal d’Investigació Mèdica (Barcelona, Spain). Cells were cultured in Dulbecco’s modified Eagle medium (DMEM) supplemented with 10 % foetal bovine serum (FBS) and 1 % penicillin/streptomycin solution and maintained in T–25 cm² flasks at 37 °C in a humidified atmosphere with 5 % CO₂. Cultures were maintained at 10⁵–10⁶ cell·mL^{–1} densities. For 1BR3G fibroblasts, cells were split every time they reached 80–90 % confluence after being harvested with PBS containing 0.025 % (w/v) trypsin and 0.01 % EDTA.

8. PEPTIDE CYTOTOXICITY

About 60,000 cells were added to different microfuge tubes containing 2-fold serial dilutions of the peptides (final concentrations in the 0.1–100 µM range) in DMEM containing 2 % FBS. After 30 min of incubation at 37 °C and 5 % CO₂, 50 µL of medium

containing approximately 10,000 treated cells were transferred to 96-well plates. Then, an amount of 15 μ L of Cell Titer Blue dye (Promega, Madison, WI, USA) was added to each well, and plates were reincubated for up to 24 h. Fluorescence at 4 and 24 h after dye addition was measured in an Infinite 200 (Tecan, Männedorf, Switzerland) reader, with $\lambda_{\text{exc}} = 530$ nm and $\lambda_{\text{em}} = 590$ nm. For 1BR3G fibroblasts, 5,000 cells/well were seeded into 96-well plates. After 24 h of incubation at 37 °C and 5 % CO₂, the medium was removed and fresh medium containing 2 % FBS and the various serial dilutions of the peptides was added. After 30 min of additional incubation, 15 μ L of Cell Titer Blue was added to each well and readings were done as above. Relative cell viability was calculated with cells treated with only DMEM containing 2 % FBS as controls (~100 % viability). Assays were carried out in triplicate.

9. HAEMOLYTIC ACTIVITY

Fresh human blood (10 mL) was collected in EDTA tubes and centrifuged at 1000 g for 10 min at 4 °C. After plasma removal, the pellet containing erythrocytes was washed three times with PBS and resuspended in PBS to obtain an 8 % (v/v) suspension. 100 μ L aliquots of erythrocyte suspension were added to microcentrifuge tubes, each containing 100 μ L of 2-fold serially diluted peptide (0.2–800 μ M) to final concentrations of 4 % (v/v) and 0.1–400 μ M, respectively. The suspensions were incubated for 30 min at 37 °C with gentle agitation, and then centrifuged for 2 min at 1000 g. Supernatants were transferred to 96-well plates, and haemoglobin release was measured by absorbance at 540 nm in an Infinite 200 (Tecan) multiplate reader. Triton X-100 at 1 % and 4 % (v/v) erythrocytes in PBS with no peptides (non-treated) were used as positive and negative controls, respectively. Measurements were carried out in triplicate. Percentage haemolysis was determined as:

$$\% \text{ haemolysis} = \frac{\text{Abs}_{540 \text{ nm}}(\text{peptide-treated}) - \text{Abs}_{540 \text{ nm}}(\text{peptide non-treated})}{\text{Abs}_{540 \text{ nm}}(1 \% \text{ Triton X-100}) - \text{Abs}_{540 \text{ nm}}(\text{peptide non-treated})} \cdot 100$$

10. SERUM STABILITY

Amounts of 0.5 mL each of human serum (Sigma) and peptide (1 mM in H₂O) were combined and incubated for 24 h at 37 °C, with gentle swirling. 120 μ L aliquots were taken at 0, 1, 5, 10, 30, 120, 360, and 1,440 min, treated with 20 μ L of trichloroacetic acid (15 % v/v in H₂O) for 30 min at 4 °C, and centrifuged at 13,000 rpm for 10 min to remove serum proteins. The supernatant was analysed by LC-MS in XBridge C18 or C8 columns (4.6 mm \times 150 mm, 3.5 μ m, Waters), eluting with linear gradients of HCOOH/MeCN (0.08 % v/v) into HCOOH/H₂O (0.1 %, v/v) over 15 min at 1 mL·min⁻¹.

RESULTS AND DISCUSSION

1. STRUCTURAL DISSECTION OF CTN, CTN₁₋₁₄ AND CTN₁₅₋₃₄ FRAGMENTS

The mature Ctn sequence (Table 3.10) was subjected to an *in silico* proteolysis (http://web.expasy.org/peptide_cutter/) with neutrophil elastase, which cuts at the carboxyl side of Val residues. (Hedstrom, 2002) Out of four possible sites, cleavage at Val14 (Table 3.10) gave two fragments of similar length, Ctn₁₋₁₄ and Ctn₁₅₋₃₄, that were selected along with Ctn for this study. Interestingly, while this cleavage generated a similar distribution of basic (Lys, Arg) residues on each fragment (net charges of +9 and +8 for Ctn₁₋₁₄ and Ctn₁₅₋₃₄ at physiological pH, respectively), it produced a marked contrast in the hydrophobicities of both peptides (Table 3.10).

Synthetic Ctn₁₋₁₄ and Ctn₁₅₋₃₄ were prepared in C-terminal amide form by solid phase methods, as described for full-length Ctn. (Falcao et al., 2014) The HPLC-purified peptides were satisfactorily characterised for purity by HPLC (all three peptides above 95 % pure, see Table 3.10) and identity by electrospray MS (Table 3.10). 1 mM stock solutions of all three peptides were prepared in deionized water and stored for up to 6 weeks at 4 °C without any detectable degradation.

2. CIRCULAR DICHROISM

The solution secondary structures of Ctn and its Ctn₁₋₁₄ and Ctn₁₅₋₃₄ fragments were first investigated by CD. In aqueous buffer, all three peptides displayed spectra (Figure 3.24, solid lines) with a 200 nm negative band, typical of aperiodic conformation. In the presence of anionic lipid vesicles, the negative bands shifted to ~220 nm (Figure 3.24, dashed lines), suggesting adoption of α -helical structure. The α -helix shift was more pronounced for Ctn and Ctn₁₋₁₄ (26 % and 31 % increase, respectively; Figure 3.24A and 3.24C) than for the less environment-sensitive Ctn₁₅₋₃₄ (10 % increase, Figure 3.24B). Helical wheel plots (Figure 3.23) show near-ideal amphipathic structure for Ctn₁₋₁₄, consistent with its high calculated hydrophobic moment of 0.763 (Table 3.10), while for Ctn ($\langle\mu_H\rangle = 0.440$) and Ctn₁₅₋₃₄ ($\langle\mu_H\rangle = 0.311$) the amphipathic distributions were far less perfect.

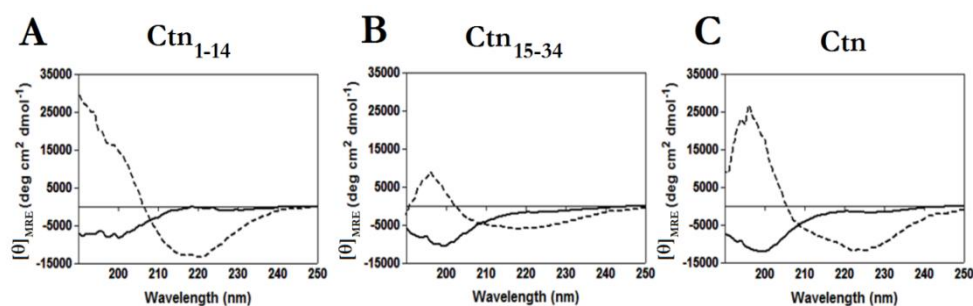


Figure 3.24. CD spectra of Ctn and its Ctn₁₋₁₄ and Ctn₁₅₋₃₄ fragments in aqueous buffer (solid lines) and in the presence of DMPG vesicles (dashed lines).

3. NMR STRUCTURES OF CTN AND DERIVED PEPTIDES IN DPC MICELLES

The structures of Ctn and its Ctn_{1–14} and Ctn_{15–34} fragments were next investigated by NMR spectroscopy in a membrane-like environment (DPC micelles). Spectra were acquired at pH 3, which is more suitable than pH 7.0–7.5 for observing NH resonances. As the peptides do not contain any residues titrating in the 2–9 pH range (Asp, Glu, His), the pH change is unlikely to alter the structural behaviour. Since the spectra of Ctn and Ctn_{15–34}, each with two Pro residues in its sequence (Table 3.10), showed signals attributable to minor species resulting from Pro *cis*–*trans* isomerism, the rotameric state of Pro25 and Pro33 in the major species of both Ctn and Ctn_{15–34} was first examined. The *trans* conformation of both Pro residues in each peptide was readily confirmed by the small difference in chemical shift between the ¹³C_β and ¹³C_γ of the two Pro residues ($\Delta\delta^{\text{Pro}} = \delta_{\text{C}\beta} - \delta_{\text{C}\gamma}$; 4.2 and 3.9 ppm for Pro25 and Pro33 of Ctn; 4.6 and 4.1 ppm for Ctn_{15–34}, respectively, at 35 °C), (Schubert, Labudde, Oschkinat, & Schmieder, 2002) and by the characteristic sequential NOE signals between the δδ' protons of Pro25 and Pro33 and the H_α protons of Pro–preceding residues, Lys24 and Ile32, respectively. Since very few signals attributable to the three possible minor rotamers (one with Pro25 in *cis* and Pro33 in *trans*; one with Pro25 in *trans* and Pro33 in *cis*; one with both Pro25 and Pro33 in *cis*) were detectable in the NMR spectra, these were not assigned. Hereafter, only NMR parameters of the major species are referred here.

After ¹H and ¹³C chemical shift assignment (see “Appendices”, Tables A11, A12, and A13), the first hints about the structural behaviour of the three peptides came from the well-established empirical relationship between conformational shifts ($\Delta\delta = \delta^{\text{observed}} - \delta^{\text{random coil}}$, ppm) of both H_α protons and C_α carbons and dihedral ϕ and ψ angles. For full-length Ctn, plots of ΔδH_α and ΔδC_α vs sequence (Figure 3.25) exhibited two clearly distinct regions: a long stretch of negative ΔδH_α and positive ΔδC_α values spanning residues 3–21, and a shorter C–terminal segment with either null or very small (ΔδH_α < 0.05 ppm, and ΔδC_α < 0.5 ppm) values except for Val30 and, as expected, the two Pro–preceding residues (Lys24 and Ile32). This profile suggested that the N–terminal region of Ctn was adopting a helical structure, whereas the C–terminal region was largely disordered, with perhaps some local non-random conformation around Val30. Further support for this conclusion came from the set of helix-characteristic NOE signals observed for the 3–22 segment, i.e., intense sequential HN–HN, and medium range αN(*i*, *i*+3) and αβ(*i*, *i*+3) (Figure 3.26). In contrast, only a few non-sequential NOEs were observed for residues 23–34.

The ΔδH_α and ΔδC_α profiles and NOE signals for Ctn_{1–14} (Figure 3.24, grey bars) were quite similar to the corresponding segments in Ctn, the main difference being the smaller magnitudes of ΔδH_α and ΔδC_α values relative to Ctn (Figure 3.24, black bars). From this it was concluded that Ctn_{1–14} formed a helical structure, though less populated than in the full-length peptide. On the basis of the averaged ΔδH_α values, (Jiménez, Barrachi-Saccilotto, Valdivia, Maqueda, & Rico, 2005; Vila, Ponte, Suau, & Rico, 2000) the helix populations of Ctn (residues 3–21) and Ctn_{1–14} (residues 3–13) in DPC micelles at 25 °C, pH 3.0, were estimated as 82 % and 57 %, respectively (Table 3.12).

For Ctn_{15–34}, $\Delta\delta H_\alpha$ and $\Delta\delta C_\alpha$ values were mostly within the random coil range (Figure 3.24, white bars), with only a few non-sequential NOE signals involving the side chains of Phe22 and Phe34 (Table 3.11, Figure 3.26) and rather small negative $\Delta\delta H_\alpha$ values for residues 18–21 suggestive of a low populated helix (estimated as 17 %, at 25 °C, Table 3.12).

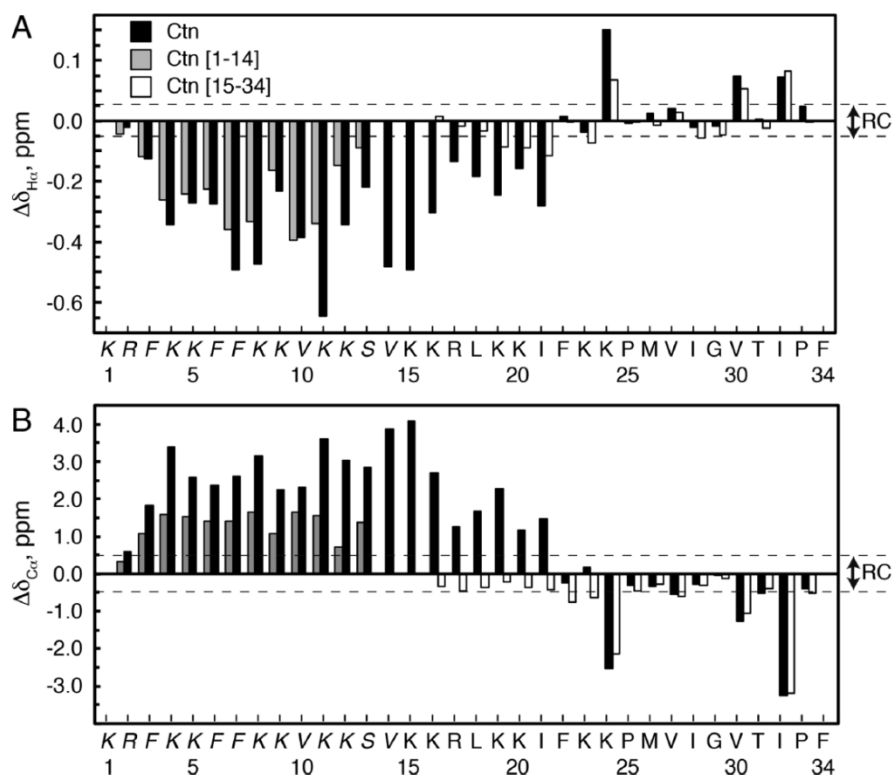


Figure 3.25. **A.** $\Delta\delta H_\alpha$ ($\Delta\delta H_\alpha = \delta H_{\alpha, \text{observed}} - \delta H_{\alpha, \text{RC}}$, ppm) and **B.** $\Delta\delta C_\alpha$ ($\Delta\delta C_\alpha = \delta C_{\alpha, \text{observed}} - \delta C_{\alpha, \text{RC}}$, ppm) values plotted as a function of sequence for Ctn (black bars), Ctn_{1–14} (grey bars), and Ctn_{15–34} (white bars) in 30 mM [D₃₈]-DPC at pH 3.0 and 25 °C. Ctn_{1–14} residues are in italics. Values for N- and C-terminal residues are not shown. Random coil (RC) values were taken from (Wishart, Bigam, Holm, Hodges, & Sykes, 1995). Dashed lines indicate RC ranges.

Table 3.11. Non-sequential NOE signals observed in Ctn_{15–34}. Those also observed for Ctn are shown in bold.

Phe22				Phe34			
	Residues i/j	Proton i	Proton j		Residues i/j	Proton i	Proton j
i, i+2	Lys20/Phe22	H α	H $\delta\delta'$	i, i+2	Ile32/Phe34	C δ 1H ₃	H $\beta\beta'$
		H $\beta\beta'$	H $\delta\delta'$			C δ 1H ₃	H $\delta\delta'$
		H $\beta\beta'$	H ζ			C δ 1H ₃	H $\epsilon\epsilon'$
	Phe22/Lys24	H $\delta\delta'$	H γ			Cγ2H₃	HN
		H $\delta\delta'$	H $\beta\beta'$			Cγ2H₃	H$\delta\delta'$
i, i+3	Phe22/Pro25	H $\delta\delta'$	H $\gamma\gamma'$	i, i+3	Thre31/Phe34	H α	H $\delta\delta'$
i, i+4	Leu18/Phe22	H$\beta\beta'$	H$\delta\delta'$				
		Hγ	H$\delta\delta'$				
		H γ	H ζ				
		Cδ1H₃	H$\delta\delta'$				
		Cδ2H₃	H$\delta\delta'$				
	Phe22/Met26	H $\delta\delta'$	H $\gamma\gamma'$				

Table 3.12. Estimated helix populations in 30 mM DPC at pH 3.0 and 25 °C, and relevant structure calculation data for peptides Ctn, Ctn₁₋₁₄, and Ctn₁₅₋₃₄.

	Ctn	Ctn ₁₋₁₄	Ctn ₁₅₋₃₄
Helix length	3–21	3–13	16–21
Averaged $\Delta\delta H_\alpha$, ppm	–0.32	–0.22	–0.07
% Helix	82	57	17
Number of distance restraints	461	200	24
Number of dihedral angle constraints (ϕ , ψ)	59	24	29
Pairwise RMSD (Å)			
Backbone atoms	0.5 ± 0.2 (3–21) 2.5 ± 0.9 (23–33)	0.5 ± 0.2 (3–13)	0.5 ± 0.6 (17–21) 2.4 ± 0.9 (23–33)
All heavy atoms	1.2 ± 0.2 (3–21) 3.5 ± 1.1 (23–33)	1.7 ± 0.3 (3–13)	2.0 ± 0.7 (3–21) 3.3 ± 0.9 (23–33)

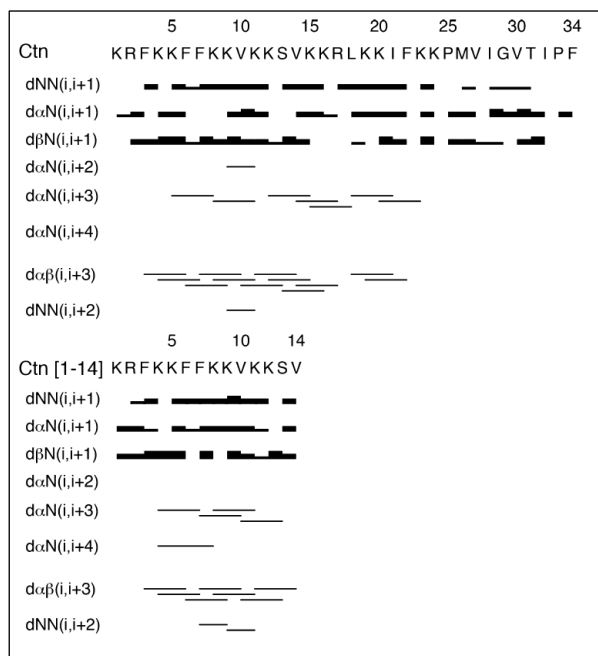


Figure 3.26. Summary of NOE signals observed for Ctn and Ctn₁₋₁₄. Intensities of sequential NOEs classified as strong, medium, and weak are indicated by the thickness of the lines.

To better visualize the 3D features of the peptides, structure calculations were performed as explained before in the “Materials and methods” section. Structural data is summarized in Table 3.13. For Ctn, the structural ensemble was poorly defined (RMSD = 4.8 ± 1.8 Å for backbone atoms) if all residues were taken into account. However, on the basis of the above chemical shift and NOE data, it was possible to outline a well-defined N-terminal α -helix spanning residues 3–21 (Figure 3.27A, RMSD = 0.5 ± 0.2 Å for backbone atoms; Table 3.13) and a poorly structured C-terminal part (Figure 3.27G; RMSD = 2.5 ± 1.9 Å for backbone atoms of residues 23–33). Side chains of the α -helix were ordered (χ_1 variability of less than $\pm 30^\circ$) and distributed as a rather amphipathic helix (Figure 3.27, panels B and D). For Ctn₁₋₁₄, the structural ensemble was a well-defined, amphipathic α -helix spanning residues 3–13 (Figure 3.27, panels C and E; RMSD = 0.6 ± 0.3 Å for backbone atoms of residues 2–13; Table 3.12), very similar to that of full-length Ctn (RMSD for backbone atoms of residues 3–13, Ctn₁₋₁₄ vs Ctn is 0.4 Å; Figure 3.27H).

For Ctn_{15–34}, as expected from the small number of observed non-sequential NOEs, the calculated structure was poorly defined, with a short amphipathic helix at the N-terminal end (Figure 3.27F) and some minor local structure around Pro33, with the side chains of Ile32 and Phe34 in close proximity (Figure 3.28), a situation reminiscent of the corresponding region in full-length Ctn albeit with some variability among the structural ensembles.

Table 3.13. Structural quality data for the ensemble of the 20 lowest target function structures calculated for peptides Ctn, Ctn_{1–14}, and Ctn_{15–34}.

	Peptide		
	Ctn	Ctn _{1–14}	Ctn _{15–34}
Number of distance restraints			
Intraresidue ($i - j = 0$)	219	85	0 ^a
Sequential ($ i - j = 1$)	126	48	5 ^a
Medium range ($1 < i - j < 5$)	116	67	19
Total number	461	200	24
Averaged total number per residue	13.6	14.3	1.2
Number of dihedral angle constraints			
Number of restricted φ angles	30	12	16
Number of restricted ψ angles	29	12	13
Total number	59	24	29
Averaged maximum violations per structure			
Distance (Å)	0.001	0.001	0.001
Dihedral angle (°)	0.2	0.2	0.2
Averaged CYANA target function value	0.2 ± 0.01	0.01 ± 0.007	0.06 ± 0.03
Number of close contacts	0	0	0
Deviations from ideal geometry			
RMSD for bond lengths (Å)	0.001	0.001	0.001
RMSD for bond angles (°)	0.2	0.2	0.2
Pairwise RMSD (Å)			
Backbone atoms ^b	0.5 ± 0.2 (3–21)	0.5 ± 0.2 (3–13)	0.5 ± 0.6 (17–21)
	2.5 ± 0.9 (23–33)		2.4 ± 0.9 (23–33)
All heavy atoms ^b	1.2 ± 0.2 (3–21)	1.7 ± 0.3 (3–13)	2.0 ± 0.7 (17–21)
	4 ± 1 (23–33)		3 ± 1 (23–33)
Ramachandran plot (%)			
Most favoured regions	91.7	99.6	72.7
Additionally allowed regions	8.3	0.4	27.3
Generously allowed regions	0	0	0
Disallowed regions	0	0	0

^aSince random structures can greatly affect the intensities of intra-residue and sequential NOEs, they were excluded for structure calculation, except for the helix-characteristic NH–NH($i, i+1$) observed for the N-terminal region. ^bResidues taken into consideration to calculate RMSD are indicated between parentheses.

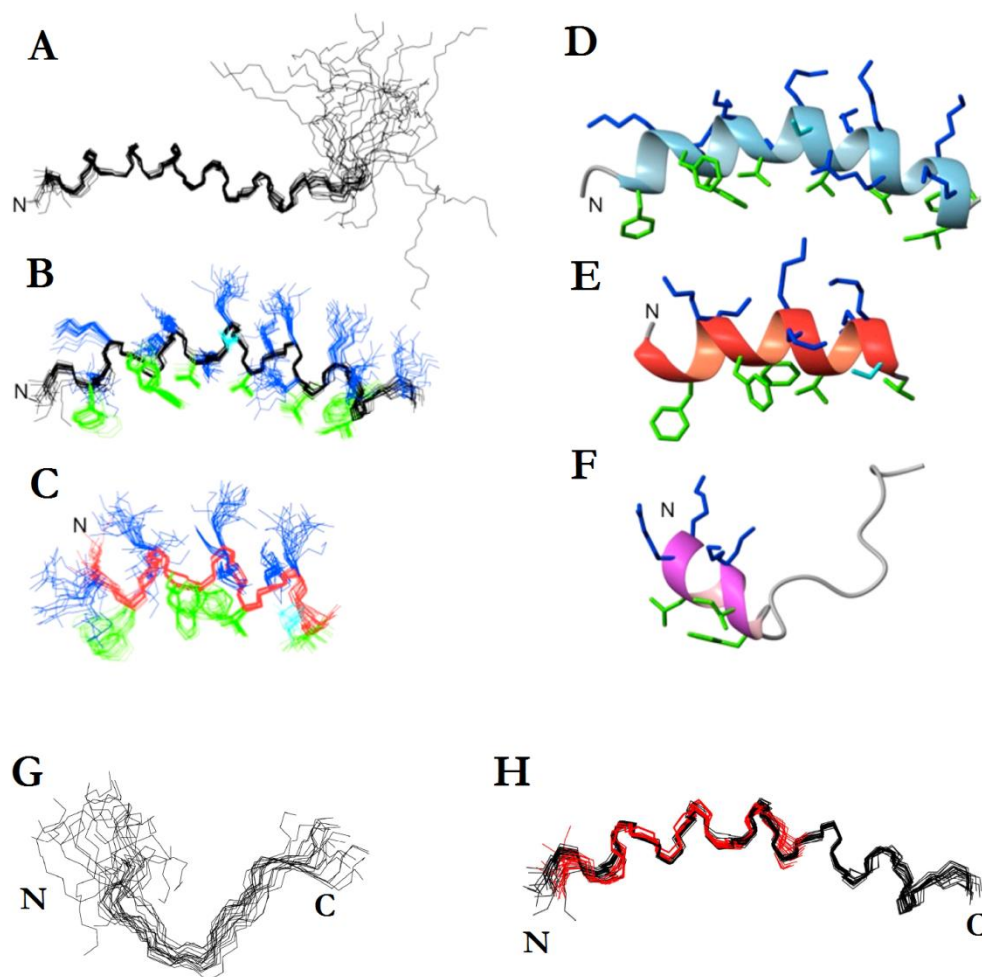


Figure 3.27. Ctn structures in DPC micelles. Representations of the 20 lowest target function calculated structures for all backbone atoms of Ctn superposed over residues 3–21 (**A**); the N–terminal region, residues 1–21, of Ctn (**B**) showing backbone atoms in black and side chains in blue if positively charged, in green if hydrophobic, and in cyan if polar; and Ctn_{15–14} (**C**) showing backbone atoms in red and side chains coloured as in panel B. Ribbon representations for the lowest target function structures of Ctn (**D**), Ctn_{1–14} (**E**), and Ctn_{15–34} (**F**). **G**. Superposition of the backbone atoms of the C–terminal region (residues 22–34) of the Ctn structure. **H**. Backbone atoms of Ctn_{1–14} structure (in red) superimposed onto those of Ctn (in black). In all the panels, “N” and “C” label the N– and C–terminus, respectively.

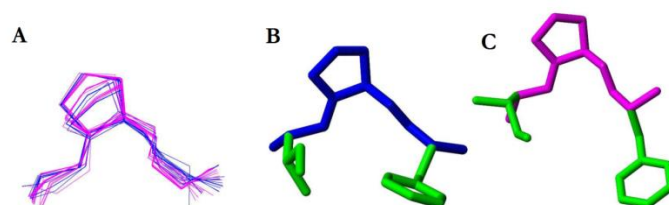


Figure 3.28. Local structure around Pro33 in Ctn and Ctn_{15–34}. **A**. Superposition of backbone atoms for the 20 lowest target function structures calculated for Ctn (blue) and Ctn_{15–34} (magenta). **B**. Representative structure of Ctn. Backbone and Pro side chain are displayed in blue, and Ile32 and Phe34 side chains in green. **C**. Representative structure of Ctn_{15–34}. Backbone and Pro side chain are displayed in magenta, Ile32 and Phe34 side chains in green.

4. ANTIMICROBIAL ACTIVITY

Considering the good antimicrobial profile already reported for Ctn, (Falcao et al., 2014) the question of to what extent it was preserved in the Ctn₁₋₁₄ and Ctn₁₅₋₃₄ fragments arose. Table 3.14 summarizes the activity of Ctn, Ctn₁₋₁₄, and Ctn₁₅₋₃₄ against various Gram-positive and Gram-negative bacterial strains. As previously observed, (Falcao et al., 2014) Ctn is an effective antimicrobial, particularly against Gram-negative organisms, with MICs in the low μM range, often better than the gentamicin control, against both standard and clinical strains of *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Escherichia coli*, and *Acinetobacter baumannii*. Ctn₁₋₁₄, for its part, did not show any antimicrobial activity. On the other hand, Ctn₁₅₋₃₄, despite > 40 % size reduction relative to Ctn, showed an antimicrobial profile only slightly inferior to the parent peptide, again better in molar terms than gentamicin against *S. pyogenes* and three of the Gram-negatives, *K. pneumoniae*, *E. coli*, and *A. baumannii*.

Table 3.14. Minimal inhibitory concentration (MIC) of Ctn and fragments Ctn₁₋₁₄ and Ctn₁₅₋₃₄ against standard and clinical strains of different bacterial species.^a

Microorganism	Gram	MIC, μM ($\mu\text{g}\cdot\text{mL}^{-1}$)			
		Ctn ₁₋₁₄	Ctn ₁₅₋₃₄	Ctn	Gentamicin
<i>E. faecalis</i> (ATCC 29212)	+	71 (> 128)	54 (> 128)	7.7 (32)	17 (8)
<i>E. faecalis</i> (CI)	+	71 (> 128)	54 (> 128)	31 (128)	34 (16)
<i>S. aureus</i> (ATCC 29213)	+	71 (> 128)	54 (> 128)	7.7 (32)	0.5 (0.25)
<i>S. aureus</i> (CI)	+	71 (> 128)	54 (> 128)	7.7 (32)	0.5 (0.25)
<i>S. pyogenes</i> (CI)	+	36 (64)	3.4 (8)	3.8 (16)	17 (8)
<i>P. aeruginosa</i> (ATCC 27853)	-	71 (> 128)	27 (64)	0.24 (1)	2.1 (1)
<i>P. aeruginosa</i> (CI)	-	71 (> 128)	27 (64)	3.8 (16)	17 (8)
<i>K. pneumoniae</i> (CI)	-	71 (> 128)	3.4 (8)	1.9 (8)	4.2 (2)
<i>E. coli</i> (ATCC 25922)	-	71 (> 128)	0.11 (0.25)	0.06 (0.25)	1 (0.5)
<i>E. coli</i> (CI)	-	71 (> 128)	3.4 (8)	3.8 (16)	34 (16)
<i>A. baumannii</i> (CI)	-	71 (> 128)	1.7 (4)	3.8 (16)	17 (8)

^aCI = clinical isolate

5. ANTITUMOUR ACTIVITY

Then, it was investigated whether the activity of Ctn and its two fragments against bacteria was paralleled against tumour cells. For these experiments, the leukaemia cell lines U937, THP-1, MM6, HL-60, and Jurkat E6.1, as well as HeLa S3, were used and, in

addition to the three isolated peptides, an equimolar mixture of Ctn₁₋₁₄ and Ctn₁₅₋₃₄ was also tested to verify if non-covalent association of both peptides could be as effective as the parent peptide. Cell viabilities were determined by measuring the decrease in metabolic activity with the non-fluorescent dye resazurin, which is reduced to fluorescent resorufin by live viable cells only. As shown in Figure 3.29, peptides displayed selective toxicity in a concentration-dependent manner. Ctn was again the most toxic peptide, with IC₅₀ values below 1 µM for HeLa S3 and U937, in the low µM range for THP-1 (IC₅₀ ≈ 1.56 µM) and MM6 (IC₅₀ ≈ 3.12 µM), and above 12.5 µM for HL-60 and Jurkat E6.1. As with bacteria, both Ctn₁₋₁₄ and Ctn₁₅₋₃₄ had differing behaviours toward tumour cells. While Ctn₁₅₋₃₄ was toxic for the most Ctn-sensitive cells (IC₅₀ values of 1, 6.25, and 25 µM for HeLa S3, U937, and THP-1, respectively), Ctn₁₋₁₄ was toxic to HeLa S3 cells only above 10 µM and had practically no effects on leukaemia cells. For the equimolar Ctn₁₋₁₄ + Ctn₁₅₋₃₄ mixture, practically no change in activity against most tumour cells was observed, except for HeLa S3, and MM6, where a slight enhancement over that achieved with Ctn₁₅₋₃₄ alone was noted, albeit never reaching that of full-length Ctn. After 24 h of incubation, peptide toxicity against all tumour cell lines exhibited similar profiles to those at 4 h (compare Figures 3.29 and 3.30), although at the highest concentrations and longest incubation times a further decrease in viabilities could be observed, as expected.

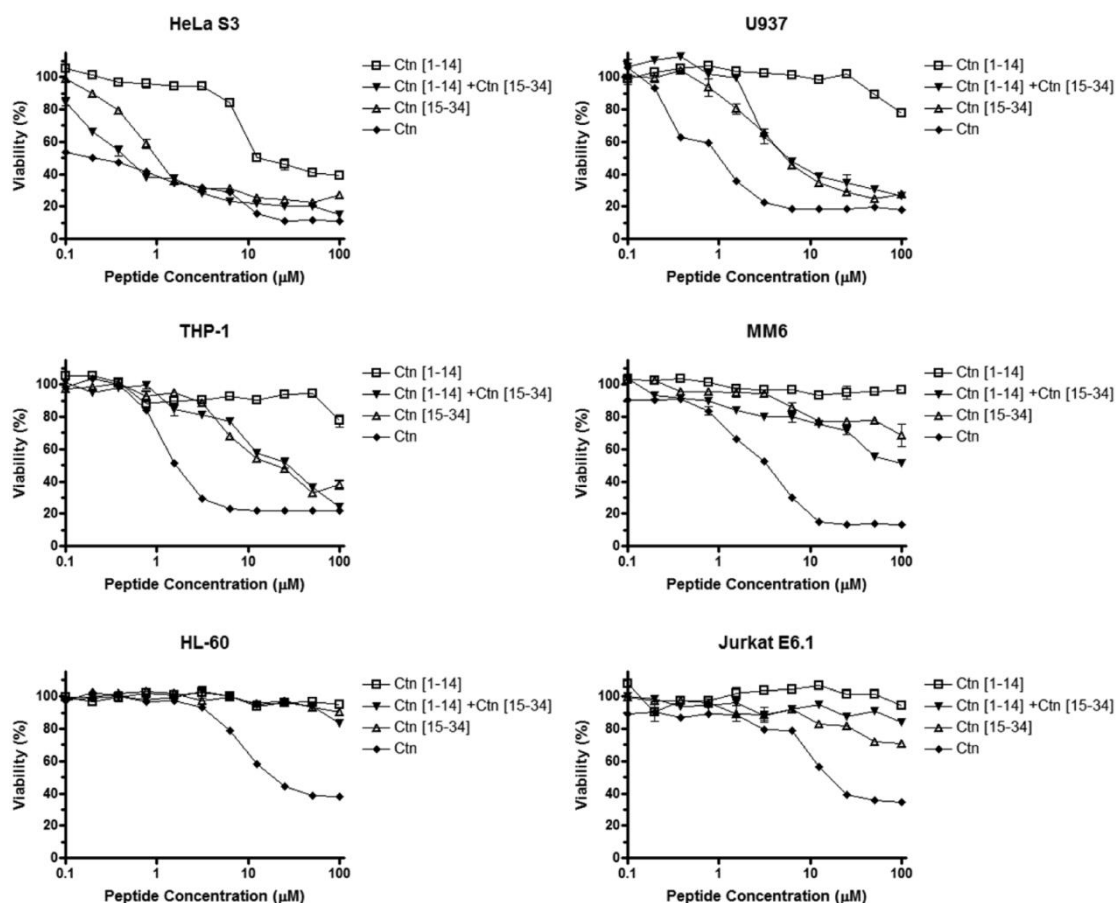


Figure 3.29. Viability of HeLa and leukaemia cells upon treatment with Ctn and fragments for 4 h.

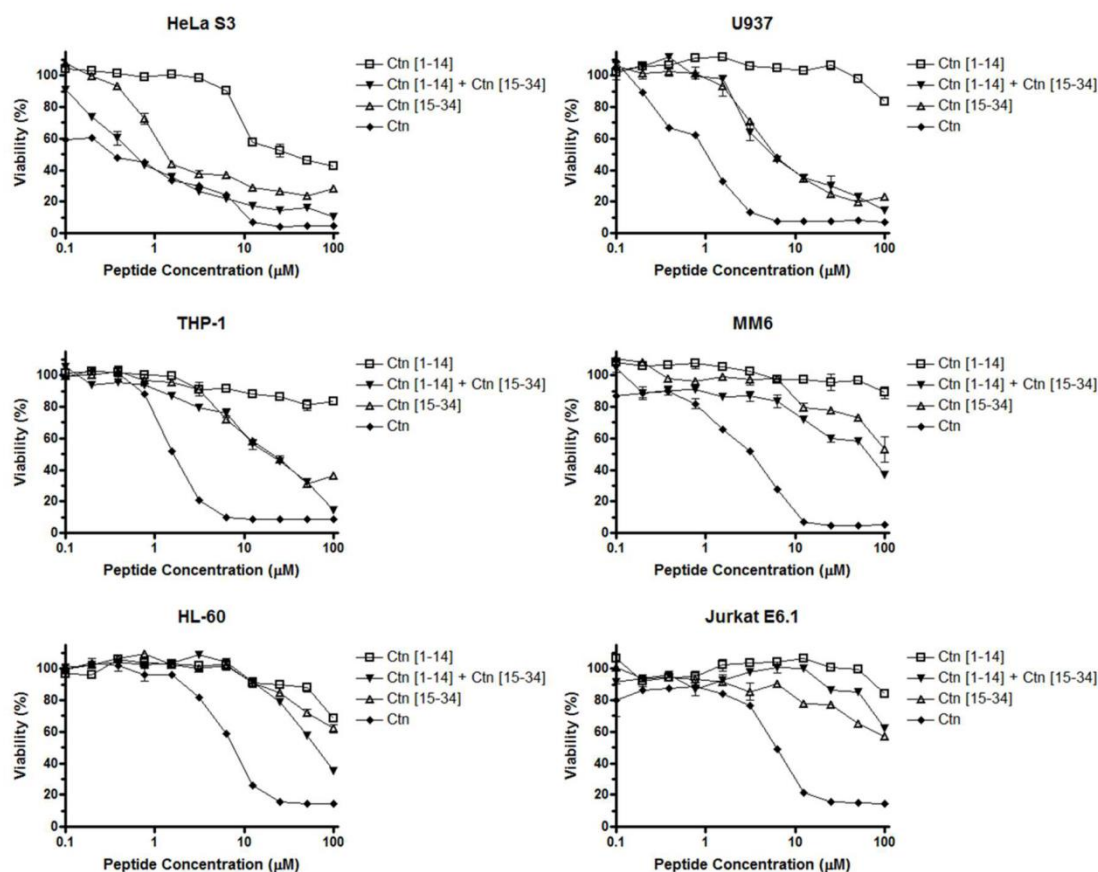


Figure 3.30. Relative cell viabilities (mean \pm SEM, with $n = 3$) of tumour cells after treatment with Ctn and fragments for 24 h.

6. TOXICITY TO EUKARYOTIC CELLS

The toxicity of Ctn and its fragments Ctn₁₋₁₄ and Ctn₁₅₋₃₄ towards human fibroblasts and human erythrocytes, as representative healthy eukaryotic cells, was next investigated. Fibroblasts (1BR3G line) were submitted to the same metabolic assay in the same concentration range as tumour cells above. Figure 3.31 (top) shows that, as for tumour cells, Ctn was again the most toxic peptide, with $IC_{50} \approx 6.25 \mu M$ after 4 h of incubation. Fragments Ctn₁₋₁₄ and Ctn₁₅₋₃₄ and their equimolar mixture had little toxicity towards fibroblasts, with only a 20 % reduction in viability at 100 μM , the highest concentration tested. Toxicity profiles were similar after 24 h, though again only Ctn had noticeable effect on cell viability (Figure 3.30).

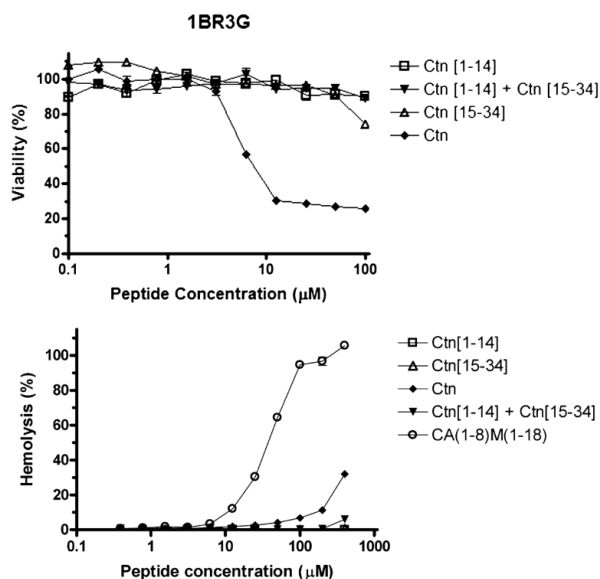


Figure 3.31. Toxicity of Ctn and fragments to eukaryotic cells: **(top)** 1BR3G fibroblast viability after 4 h; **(bottom)** hemolysis data.

The haemolytic effect of the peptides was tested on human erythrocytes and showed Ctn as a moderately lytic peptide (7 % and 33 % haemolysis at 100 and 400 μM, respectively), although far less than standard AMPs, such as the cecropin A–melittin hybrid CA(1–8)M(1–18) (Figure 3.31, bottom). In contrast, Ctn_{1–14} and Ctn_{15–34} were totally non-haemolytic up to 400 μM, and their equimolar mixture at that concentration caused only 7 % haemolysis.

7. SERUM STABILITY

To further explore their therapeutic potential, Ctn, Ctn_{1–14} and Ctn_{15–34} were tested for stability against human serum by LC–MS. For Ctn_{1–14}, relatively fast breakdown was observed (Figure 3.32), with 100 % clearance after 6 h ($t_{1/2}$ = 21 min). Most cleavage products reflected trypsin–like cleavage, consistent with the presence of 7 Lys + 1 Arg residues.. Predictably, the protease lability of the N–terminal also caused a rather short lifetime for the full–length sequence, with $t_{1/2}$ = 71 min. In contrast, Ctn_{15–34} proved remarkably resilient to serum degradation, with ~10 % peptide still remaining after 24 h incubation ($t_{1/2}$ = 770 min) (Figure 3.32).

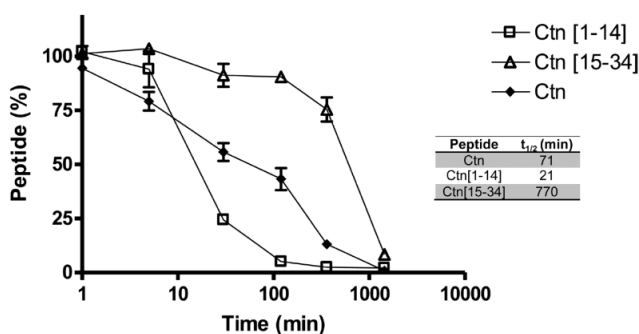


Figure 3.32. Time–course of the treatment of Ctn, Ctn_{1–14}, and Ctn_{15–34} with human serum and (inset table) the corresponding half–lives.

8. DISCUSSION

In the unrelenting search for new medicines that can curb fast-spreading antibiotic resistance, AMPs have for some time been regarded as a promising alternative. (B Hadley & EW Hancock, 2010; R. Hancock & Sahl, 2013; R. E. Hancock & Sahl, 2006; Yeung, Gellatly, & Hancock, 2011) Until now, however, few AMPs have reached clinical trials and their use is mainly confined to topical applications. Despite some recognized advantages, clinical development of AMPs as anti-infective agents is still hampered by issues such as low protease stability, potential toxicity, and high manufacturing costs. Protease susceptibility, with ensuing poor bioavailability, is particularly challenging and has been only partially addressed by strategies such as N- or C-terminal alteration (acetylation or amidation), site-specific changes (D- or artificial amino acid residues, peptidomimetics), prodrug approaches, (R. Hancock & Sahl, 2013; Yeung et al., 2011) or encapsulation, which can also decrease toxicity. (Torchilin, 2005) However, as with other sequence-dependent properties, protease susceptibility or toxicity can be predicted only incompletely by bioinformatic approaches (Gupta, Ravi, & Sharma, 2013; Sharma, Singla, Rashid, & Raghava, 2014) and experimental validation is mandatory.

To address these toxicity and protease stability issues often preventing AMPs from advancing in the drug pipeline, a structural dissection of Ctn has been performed, which has led to Ctn₁₅₋₃₄, a substantially (60 %) downsized analogue with potential therapeutic application. Earlier works had already unveiled a potent antibacterial activity in Ctn, (Falcao et al., 2014) but peptide size, at 34 residues, was a drawback, hence minimally active substructures were searched for. Similar structure-guided approaches with human cathelicidin LL-37 managed to downsize the parent molecule to fragments retaining amphipathic α -helical structure with the same or improved antimicrobial activity. (Dannehl, Gutschmann, & Brezesinski, 2013; Nagant et al., 2012; Wang, Mishra, Epand, & Epand, 2014) Also, Asian elapid CRAMPs have been modified in their α -helical domains to enhance antibiotic properties by the addition or removal of basic and/or hydrophobic residues. (Chen et al., 2011; Juba, Porter, Dean, Gillmor, & Bishop, 2013; Zhang et al., 2010) In the present case, the known conversion of LL-37 by sweat proteases into fragments retaining the activity of the parent molecule (Murakami, Lopez-Garcia, Braff, Dorschner, & Gallo, 2004) was a motivation to simulate Ctn cleavage using neutrophil elastase as a model enzyme. For a snake venom peptide such as Ctn, the choice of elastase was justified in that neutrophils are present and release protease-containing granules at inflammation sites, such as those in a bitten prey. (Meyer-Hoffert, 2008) The *in silico* dissection at four possible cleavage sites generated eight pairwise fragments of which only the largest pair, Ctn₁₋₁₄ and Ctn₁₅₋₃₄, resulting from cleavage at Val14, was considered worth evaluating along with Ctn. All three peptides were synthesized as C-terminal carboxamides, a modification that entails a modest enhancement in proteolytic stability and has been previously applied to a fragment of LL-37 (Strömstedt, Pasupuleti, Schmidtchen, & Malmsten, 2009) as well as to other AMPs. (Cao, Zhou, Ma, Luo, & Wei, 2005; R Dennison, HG Morton, & A Phoenix, 2012)

CD examination of the secondary structure of the three peptides in anionic DMPG micelles that resemble bacterial cell membranes showed a strong α -helical tendency for Ctn₁₋₁₄, also observed for the entire Ctn sequence, and a more disorganized structure, with lower helix propensity, for Ctn₁₅₋₃₄. NMR experiments in the presence of DPC micelles simulating the neutral zwitterionic membranes of healthy eukaryotic cells corroborated CD results, revealing two distinct structural regions in Ctn (a well-defined helix at residues 3–21 followed by a rather unstructured C-terminal part) roughly around the putative elastase cleavage site.

Antimicrobial screening of Ctn and its fragments against a representative panel of Gram-positive and Gram-negative organisms showed that the earlier reported strong bactericidal activity of the full-length peptide (Falcao et al., 2014) is largely confined to the C-terminal Ctn₁₅₋₃₄ section, while the N-terminus, regardless of its strong cationic (8 basic residues out of 14) and amphipathic α -helical nature (predictable from the calculated $\mu_H = 0.763$ (Table 3.10), the highest of all three peptides and experimentally confirmed by NMR (Figure 3.27C and 3.27E), was devoid of significant AMP properties. This lack of activity of a strongly cationic, amphipathic helical peptide such as Ctn₁₋₁₄ is somewhat puzzling, at least in terms of the conventional paradigm whereby AMP action is associated with amphipathic α -helical structures rich in basic (Lys, Arg) residues (e.g., cecropins, magainins, melittin). Thus, when the Ctn sequence is run on an AMP prediction algorithm, (Torrent et al., 2012) the 22-residue segment from the N-terminus, comprising Ctn₁₋₁₄, is (inaccurately) predicted as antimicrobial.

As a possible explanation for such non-compliance of Ctn₁₋₁₄ with the predictions, it might be argued that the high μ_H value (Terwilliger, 1982) calculated for the N-terminal segment does not realistically estimate its AMP potential, since its low content in hydrophobic residues (only 1/3 of total, $H = -0.012$, Table 3.10) may affect its global ability to interact with lipid bilayers. Indeed, the simplistic view of cationic, amphipathic α -helices as mandatory for AMP action has been convincingly challenged (Papo, Oren, Pag, Sahl, & Shai, 2002; Shai & Oren, 1996) by showing that AMPs lacking α -helical structure but with amphipathic features (i.e., diastereomers of canonical AMPs made by educated D-amino acid replacements) can be designed as effective antimicrobials; hence, amphipathicity though not α -helicity appears to be the requirement. In any event, these data on cationic, α -helical, amphipathic yet inactive Ctn₁₋₁₄, as well as on poorly structured but active Ctn₁₅₋₃₄ (see below), constitute a salutary reminder that even the best structure-based predictions of AMP activity are no substitute for experimental validation.

While for Gram-positive bacteria, except *S. pyogenes* (Table 3.14), Ctn₁₅₋₃₄ significantly underperformed Ctn; for Gram-negatives it matched to a reasonable extent the antimicrobial behaviour of the parental structure, despite the substantial size reduction (42 %) relative to Ctn and the low helicity and non-ideal amphipathicity revealed by CD and NMR. Of note are the low- μ M MICs of Ctn₁₅₋₃₄ against clinical isolates of *K. pneumoniae* and, particularly, *E. coli* and *A. baumannii*, which outdid a reference antibiotic such as gentamicin by about one order of magnitude on a molar basis. Hence, it seems that preserving one-half of the positive charge of Ctn but substantially increasing

hydrophobicity (0.455 vs 0.263, Table 3.10) turns Ctn_{15–34} into a peptide that, while less amphipathic than the parent Ctn, can effectively target the anionic membranes and/or the lipopolysaccharide of Gram–negative bacteria although it loses all activity toward Gram–positives. Interestingly, the rather equivalent antimicrobial profiles of Ctn and Ctn_{15–34} against Gram–negatives (except *P. aeruginosa*) did not apply to eukaryotic cells such as fibroblasts or erythrocytes, for both of which the full–length peptide proved to be rather toxic whereas Ctn_{15–34} was practically innocuous. It would appear that the size reduction in Ctn_{15–34}, plus the loss of half the positive charges and the increase in hydrophobicity relative to Ctn, caused weaker interactions with the membranes of healthy eukaryotic cells, composed mainly of zwitterionic phospholipids and cholesterol.

As antitumour activity is often observed alongside microbicidal properties in many AMPs, (Gaspar et al., 2014; Schweizer, 2009) the three peptides of this study were accordingly tested against various tumour cell lines (Figure 3.29). In tune with the above–mentioned toxicity to healthy eukaryotic cells, Ctn was also rather cytotoxic against both HeLa and leukaemia cells, whereas Ctn_{1–14} and Ctn_{15–34} were, respectively, practically ineffective (active only against HeLa at >10 µM) and active on those cell types most susceptible to Ctn, i.e., HeLa S3, U937, and THP–1, with IC₅₀ values not widely different from those of the parent peptide. The toxicity of Ctn_{15–34} towards cancer cells did not significantly improve by admixture with Ctn_{1–14}, strongly suggesting that the bioactive structure (hence maximal antitumour activity) is achieved only by the full Ctn sequence, not by a combined effect of its fragments.

Taken together, the evidence collected thus far (bacterial MICs, antitumour activity, toxicity to non–tumour cells) portrays Ctn as a toxic peptide with a rather indiscriminate killing effect on bacteria, tumour, and non–tumour cells, hence with limited therapeutic potential. The N–terminal Ctn_{1–14} segment, for its part, has lost practically all the antimicrobial and antitumour activity of the parental sequence and thus lacks interest for either anti–infective or antitumour applications. In contrast, Ctn_{15–34} preserves most of the antimicrobial activity of Ctn, particularly against Gram–negatives, and its slightly inferior cytotoxicity to tumour cells relative to Ctn is more than made up by its practically nil toxicity toward healthy eukaryotic cells (Figure 3.31). All these features plus a more convenient size (20 vs 34 residues), hence ease of production, and, last but not least, a definitely remarkable stability in human serum (Figure 3.32) combine to make Ctn_{15–34} a rather promising peptide lead for potential development into an anti–infective or (more likely) an antitumour agent.

CONCLUSIONS

1. Ctn and its fragments adopt random coil conformation in aqueous solution, and they become helical in the presence of lipid vesicles, except Ctn₁₅₋₃₄, which seems to be less environment-sensitive.
2. In the presence of micelles, Ctn shows a long helical structure. The helix is conserved in the whole length of the Ctn₁₋₁₄ peptide. In Ctn₁₅₋₃₄ fragment, structure is poorly defined, except for the helical conformation maintained in the N-terminus and a minor local structuration around Pro33.
3. Ctn is an effective antimicrobial, particularly against Gram-negative bacteria. It was the most toxic for tumour cells (IC₅₀ values $\sim\mu\text{M}$ or lower), and also for healthy eukaryotic cells. Ctn showed a short lifetime in human serum. All these features taken together make of Ctn a peptide with limited therapeutical potential.
4. Ctn₁₋₁₄ did not show any antibacterial activity, it has minor effects on tumour cells, and low serum stability, and thus, it has no therapeutical interest. An equimolar mixture of Ctn₁₋₁₄ and Ctn₁₅₋₃₄, showed an antitumour activity similar to that of Ctn₁₅₋₃₄ alone and low cytotoxicity.
5. Ctn₁₅₋₃₄ is only slightly less effective as bactericidal than the parent peptide, Ctn, and it was toxic for the most Ctn-sensitive tumour cells. In addition, it was remarkable stable in human serum and showed a negligible cytotoxicity in healthy cells. These good properties and its reduced size make this peptide a promising therapeutic molecule.
6. It was demonstrated that the poorly structured fragment, Ctn₁₅₋₃₄, retained a high effectiveness against bacteria, whereas the helical fragment, Ctn₁₋₁₄, did not show any antibacterial activity. This fact reveals that classic ideas about structure-function relationships have to be considered always together with experimental evidences.

BIBLIOGRAPHY

- (CLSI), C. L. S. I. (2012). Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically M07-A9.
- (WHO), W. H. O. (2012). Disease and injury regional mortality estimates, 2000-2012. http://www.who.int/entity/healthinfo/global_burden_disease/GHE_DthWHOR_eg6_2000_2012.xls?ua=1
- (WHO), W. H. O. (2014). Antimicrobial Resistance: Global Report on Surveillance. http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748_eng.pdf?ua=1
- Afonin, S., Mikhailiuk, P. K., Komarov, I. V., & Ulrich, A. S. (2007). Evaluating the amino acid CF₃-bicyclopentylglycine as a new label for solid-state ¹⁹F-NMR structure analysis of membrane-bound peptides. *J Pept Sci*, 13(9), 614-623. doi:10.1002/psc.854
- Aisenbrey, C., Michalek, M., Salnikov, E. S., & Bechinger, B. (2013). Solid-state NMR approaches to study protein structure and protein-lipid interactions. *Methods Mol Biol*, 974, 357-387. doi:10.1007/978-1-62703-275-9_16
- Alves, C. S., Melo, M. N., Franquelim, H. G., Ferre, R., Planas, M., Feliu, L., . . . Castanho, M. A. (2010). Escherichia coli cell surface perturbation and disruption induced by antimicrobial peptides BP100 and pepR. *J Biol Chem*, 285(36), 27536-27544. doi:10.1074/jbc.M110.130955
- Ammann, C., Meier, P., & Merbach, A. (1982). A simple multinuclear NMR thermometer. *Journal of Magnetic Resonance (1969)*, 46(2), 319-321.
- Andreu, D., Ubach, J., Boman, A., Wahlin, B., Wade, D., Merrifield, R. B., & Boman, H. G. (1992). Shortened cecropin A-melittin hybrids. Significant size reduction retains potent antibiotic activity. *FEBS Lett*, 296(2), 190-194.
- B Hadley, E., & EW Hancock, R. (2010). Strategies for the discovery and advancement of novel cationic antimicrobial peptides. *Current topics in medicinal chemistry*, 10(18), 1872-1881.
- Badosa, E., Ferre, R., Planas, M., Feliu, L., Besalu, E., Cabrefiga, J., . . . Montesinos, E. (2007). A library of linear undecapeptides with bactericidal activity against phytopathogenic bacteria. *Peptides*, 28(12), 2276-2285. doi:10.1016/j.peptides.2007.09.010
- Balla, M. S., Bowie, J. H., & Separovic, F. (2004). Solid-state NMR study of antimicrobial peptides from Australian frogs in phospholipid membranes. *Eur Biophys J*, 33(2), 109-116. doi:10.1007/s00249-003-0342-7
- Bechinger, B., Gierasch, L. M., Montal, M., Zasloff, M., & Opella, S. J. (1996). Orientations of helical peptides in membrane bilayers by solid state NMR spectroscopy. *Solid State Nucl Magn Reson*, 7(3), 185-191.
- Bechinger, B., Kim, Y., Chirlian, L. E., Gesell, J., Neumann, J. M., Montal, M., . . . Opella, S. J. (1991). Orientations of amphipathic helical peptides in membrane bilayers determined by solid-state NMR spectroscopy. *Journal of Biomolecular NMR*, 1(2), 167-173.
- Blazyk, J., Wiegand, R., Klein, J., Hammer, J., Epand, R. M., Epand, R. F., . . . Kari, U. P. (2001). A novel linear amphipathic β -sheet cationic antimicrobial peptide with enhanced selectivity for bacterial lipids. *Journal of Biological Chemistry*, 276(30), 27899-27906.
- Boman, H. G., Wade, D., Boman, I. A., Wählin, B., & Merrifield, R. B. (1989). Antibacterial and antimalarial properties of peptides that are cecropin-melittin hybrids. *FEBS Letters*, 259(1), 103-106. doi:10.1016/0014-5793(89)81505-4

- Brogden, K. A. (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol*, 3(3), 238-250. doi:10.1038/nrmicro1098
- Brusselsaers, N., Vogelaers, D., & Blot, S. (2011). The rising problem of antimicrobial resistance in the intensive care unit. *Annals of intensive care*, 1(1), 1-7.
- Bürck, J., Roth, S., Wadhwani, P., Afonin, S., Kanithasen, N., Strandberg, E., & Ulrich, A. S. (2008). Conformation and membrane orientation of amphiphilic helical peptides by oriented circular dichroism. *Biophysical journal*, 95(8), 3872-3881.
- Caccavo, D., Pellegrino, N. M., Altamura, M., Rigon, A., Amati, L., Amoroso, A., & Jirillo, E. (2002). Review: Antimicrobial and immunoregulatory functions of lactoferrin and its potential therapeutic application. *Journal of Endotoxin Research*, 8(6), 403-417. doi:10.1177/09680519020080060901
- Cao, W., Zhou, Y., Ma, Y., Luo, Q., & Wei, D. (2005). Expression and purification of antimicrobial peptide adenoregulin with C-amidated terminus in *Escherichia coli*. *Protein expression and purification*, 40(2), 404-410.
- Cavallarin, L., Andreu, D., & San Segundo, B. (1998). Cecropin A-derived peptides are potent inhibitors of fungal plant pathogens. *Mol Plant Microbe Interact*, 11(3), 218-227. doi:10.1094/mpmi.1998.11.3.218
- Cornilescu, G., Delaglio, F., & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR*, 13(3), 289-302.
- Chen, W., Yang, B., Zhou, H., Sun, L., Dou, J., Qian, H., . . . Han, J. (2011). Structure-activity relationships of a snake cathelicidin-related peptide, BF-15. *Peptides*, 32(12), 2497-2503.
- Daily, A. E., Greathouse, D. V., van der Wel, P. C., & Koeppe, R. E., 2nd. (2008). Helical distortion in tryptophan- and lysine-anchored membrane-spanning alpha-helices as a function of hydrophobic mismatch: a solid-state deuterium NMR investigation using the geometric analysis of labeled alanines method. *Biophys J*, 94(2), 480-491. doi:10.1529/biophysj.106.097543
- Dannehl, C., Gutschmann, T., & Brezesinski, G. (2013). Surface activity and structures of two fragments of the human antimicrobial LL-37. *Colloids and Surfaces B: Biointerfaces*, 109, 129-135.
- Dathe, M., Wieprecht, T., Nikolenko, H., Handel, L., Maloy, W. L., MacDonald, D. L., . . . Bienert, M. (1997). Hydrophobicity, hydrophobic moment and angle subtended by charged residues modulate antibacterial and haemolytic activity of amphipathic helical peptides. *FEBS Lett*, 403(2), 208-212.
- Davis, J., Jeffrey, K., Bloom, M., Valic, M., & Higgs, T. (1976). Quadrupolar echo deuteron magnetic resonance spectroscopy in ordered hydrocarbon chains. *Chemical Physics Letters*, 42(2), 390-394.
- De Lucca, A. J., & Walsh, T. J. (1999). Antifungal Peptides: Novel Therapeutic Compounds against Emerging Pathogens. *Antimicrobial agents and chemotherapy*, 43(1), 1-11.
- Dennison, S. R., Harris, F., Bhatt, T., Singh, J., & Phoenix, D. A. (2009). A theoretical analysis of secondary structural characteristics of anticancer peptides. *Molecular and Cellular Biochemistry*, 333(1), 129-135. doi:10.1007/s11010-009-0213-3
- Eggenberger, K., Mink, C., Wadhwani, P., Ulrich, A. S., & Nick, P. (2011). Using the peptide BP100 as a cell-penetrating tool for the chemical engineering of actin filaments within living plant cells. *Chembiochem*, 12(1), 132-137. doi:10.1002/cbic.201000402
- Esteban-Martin, S., & Salgado, J. (2007). The dynamic orientation of membrane-bound peptides: bridging simulations and experiments. *Biophys J*, 93(12), 4278-4288. doi:10.1529/biophysj.107.113043

- Esteban-Martin, S., Strandberg, E., Fuertes, G., Ulrich, A. S., & Salgado, J. (2009). Influence of whole-body dynamics on 15N PISEMA NMR spectra of membrane proteins: a theoretical analysis. *Biophys J*, 96(8), 3233-3241. doi:10.1016/j.bpj.2008.12.3950
- Esteban-Martin, S., Strandberg, E., Salgado, J., & Ulrich, A. S. (2010). Solid state NMR analysis of peptides in membranes: Influence of dynamics and labeling scheme. *Biochim Biophys Acta*, 1798(2), 252-257. doi:10.1016/j.bbamem.2009.08.010
- Falcao, C., de La Torre, B., Pérez-Peinado, C., Barron, A., Andreu, D., & Rádis-Baptista, G. (2014). Viperidins: a novel family of cathelicidin-related peptides from the venom gland of South American pit vipers. *Amino acids*, 46(11), 2561-2571.
- Fennell, J. F., Shipman, W. H., & Cole, L. J. (1967). *Antibacterial action of a bee venom fraction (melittin) against a penicillin-resistant staphylococcus and other microorganisms*. Retrieved from
- Fennell, J. F., Shipman, W. H., & Cole, L. J. (1968). Antibacterial Action of Melittin, a Polypeptide from Bee Venom. *Experimental Biology and Medicine*, 127(3), 707-710. doi:10.3181/00379727-127-32779
- Ferre, R., Badosa, E., Feliu, L., Planas, M., Montesinos, E., & Bardaji, E. (2006). Inhibition of plant-pathogenic bacteria by short synthetic cecropin A-melittin hybrid peptides. *Appl Environ Microbiol*, 72(5), 3302-3308. doi:10.1128/aem.72.5.3302-3308.2006
- Ferre, R., Melo, M. N., Correia, A. D., Feliu, L., Bardaji, E., Planas, M., & Castanho, M. (2009). Synergistic effects of the membrane actions of cecropin-melittin antimicrobial hybrid peptide BP100. *Biophysical journal*, 96(5), 1815-1827.
- Fields, G. B., & Noble, R. L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int J Pept Protein Res*, 35(3), 161-214.
- Fox, J. L. (2013). Antimicrobial peptides stage a comeback. *Nat Biotech*, 31(5), 379-382. doi:10.1038/nbt.2572
- French, G. (2005). Clinical impact and relevance of antibiotic resistance. *Advanced drug delivery reviews*, 57(10), 1514-1527.
- Fung, B. M., Khitrin, A. K., & Ermolaev, K. (2000). An improved broadband decoupling sequence for liquid crystals and solids. *J Magn Reson*, 142(1), 97-101. doi:10.1006/jmre.1999.1896
- Gaspar, D., Veiga, A. S., & Castanho, M. A. (2014). From antimicrobial to anticancer peptides. A review. *New edge of antibiotic development: antimicrobial peptides and corresponding resistance*, 24.
- Gautier, R., Douguet, D., Antonny, B., & Drin, G. (2008). HELIQUEST: a web server to screen sequences with specific α -helical properties. *Bioinformatics*, 24(18), 2101-2102.
- Giangaspero, A., Sandri, L., & Tossi, A. (2001). Amphipathic alpha helical antimicrobial peptides. *Eur J Biochem*, 268(21), 5589-5600.
- Glaser, R. W., Sachse, C., Durr, U. H., Wadhwani, P., Afonin, S., Strandberg, E., & Ulrich, A. S. (2005). Concentration-dependent realignment of the antimicrobial peptide PGLa in lipid membranes observed by solid-state 19F-NMR. *Biophys J*, 88(5), 3392-3397. doi:10.1529/biophysj.104.056424
- Glaser, R. W., Sachse, C., Durr, U. H., Wadhwani, P., & Ulrich, A. S. (2004). Orientation of the antimicrobial peptide PGLa in lipid membranes determined from 19F-NMR dipolar couplings of 4-CF3-phenylglycine labels. *J Magn Reson*, 168(1), 153-163. doi:10.1016/j.jmr.2004.02.008
- Goddard, T., & Kneller, D. (2004). SPARKY 3. *University of California, San Francisco*, 14, 15.
- Grage, S. L., Strandberg, E., Wadhwani, P., Esteban-Martin, S., Salgado, J., & Ulrich, A. S. (2012). Comparative analysis of the orientation of transmembrane peptides using solid-state (2)H- and (15)N-NMR: mobility matters. *Eur Biophys J*, 41(5), 475-482. doi:10.1007/s00249-012-0801-0

- Grau-Campistany, A., Strandberg, E., Wadhwani, P., Reichert, J., Bürck, J., Rabanal, F., & Ulrich, A. S. (2015). Hydrophobic mismatch demonstrated for membranolytic peptides, and their use as molecular rulers to measure bilayer thickness in native cells. *Scientific reports*, 5.
- Güell, I., Micaló, L., Cano, L., Badosa, E., Ferre, R., Montesinos, E., . . . Planas, M. (2012). Peptidotriazoles with antimicrobial activity against bacterial and fungal plant pathogens. *Peptides*, 33(1), 9-17. doi:<http://dx.doi.org/10.1016/j.peptides.2011.12.003>
- Güntert, P. (2004). Automated NMR structure calculation with CYANA *Protein NMR Techniques* (pp. 353-378): Springer.
- Güntert, P., Mumenthaler, C., & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol*, 273(1), 283-298.
- Gupta, P., Ravi, I., & Sharma, V. (2013). Induction of β -1,3-glucanase and chitinase activity in the defense response of *Eruca sativa* plants against the fungal pathogen *Alternaria brassicicola*. *Journal of Plant Interactions*, 8(2), 155-161. doi:10.1080/17429145.2012.679705
- Habermann, E. (1954). Zur Pharmakologie des Melittin. *Naunyn-Schmiedeberg's Archiv für experimentelle Pathologie und Pharmakologie*, 222(1), 173-175. doi:10.1007/bf00249295
- Habermann, E., & Jentsch, J. (1966). Über die Struktur des toxischen Bienengiftpeptids Melittin und deren Beziehung zur pharmakologischen Wirkung. *Naunyn-Schmiedeberg's Archiv für experimentelle Pathologie und Pharmakologie*, 253(1), 40-41. doi:10.1007/bf00259241
- Hancock, R., & Sahl, H.-G. (2013). New strategies and compounds for anti-infective treatment. *Current opinion in microbiology*, 16(5), 519.
- Hancock, R. E., & Sahl, H. G. (2006). Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nat Biotechnol*, 24(12), 1551-1557. doi:10.1038/nbt1267
- Hancock, R. E. W., & Chapple, D. S. (1999). Peptide Antibiotics. *Antimicrobial agents and chemotherapy*, 43(6), 1317-1323.
- Hedstrom, L. (2002). Serine protease mechanism and specificity. *Chemical reviews*, 102(12), 4501-4524.
- Heitz, F., Morris, M. C., & Divita, G. (2009). Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *British Journal of Pharmacology*, 157(2), 195-206. doi:10.1111/j.1476-5381.2009.00057.x
- Hoskin, D. W., & Ramamoorthy, A. (2008). Studies on anticancer activities of antimicrobial peptides. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(2), 357-375. doi:<http://dx.doi.org/10.1016/j.bbamem.2007.11.008>
- Hultmark, D., Steiner, H., Rasmuson, T., & Boman, H. G. (1980). Insect Immunity. Purification and Properties of Three Inducible Bactericidal Proteins from Hemolymph of Immunized Pupae of *Hyalophora cecropia*. *European Journal of Biochemistry*, 106(1), 7-16. doi:10.1111/j.1432-1033.1980.tb05991.x
- Jiménez, M., Barrachi-Saccilotto, A. C., Valdivia, E., Maqueda, M., & Rico, M. (2005). Design, NMR characterization and activity of a 21-residue peptide fragment of bacteriocin AS-48 containing its putative membrane interacting region. *Journal of Peptide Science*, 11(1), 29-36.
- Johnson, W. C. (1999). Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins*, 35(3), 307-312.
- Juba, M., Porter, D., Dean, S., Gillmor, S., & Bishop, B. (2013). Characterization and performance of short cationic antimicrobial peptide isomers. *Peptide Science*, 100(4), 387-401.

- Koradi, R., Billeter, M., & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics*, 14(1), 51-55.
- Kreil, G. (1965). Zur Reindarstellung und Charakterisierung von Melittin, dem Haupttoxin des Bienengiftes. *Monatshefte für Chemie und verwandte Teile anderer Wissenschaften*, 96(6), 2061-2063. doi:10.1007/bf01185937
- Last, N. B., Schlamadinger, D. E., & Miranker, A. D. (2013). A common landscape for membrane-active peptides. *Protein Science*, 22(7), 870-882. doi:10.1002/pro.2274
- Levitt, M., Suter, D., & Ernst, R. (1986). Spin dynamics and thermodynamics in solid-state NMR cross polarization. *The Journal of chemical physics*, 84(8), 4243-4255.
- Li, Y., Xiang, Q., Zhang, Q., Huang, Y., & Su, Z. (2012). Overview on the recent study of antimicrobial peptides: Origins, functions, relative mechanisms and application. *Peptides*, 37(2), 207-215. doi:<http://dx.doi.org/10.1016/j.peptides.2012.07.001>
- Lindgren, M., Hällbrink, M., Prochiantz, A., & Langel, Ü. (2000). Cell-penetrating peptides. *Trends in Pharmacological Sciences*, 21(3), 99-103. doi:[http://dx.doi.org/10.1016/S0165-6147\(00\)01447-4](http://dx.doi.org/10.1016/S0165-6147(00)01447-4)
- Lobley, A., Whitmore, L., & Wallace, B. A. (2002). DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, 18(1), 211-212.
- Madani, F., Lindberg, S., Langel, #220, lo, Futaki, S., . . . slund, A. (2011). Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. *Journal of Biophysics*, 2011. doi:10.1155/2011/414729
- Manzini, M. C., Perez, K. R., Riske, K. A., Bozelli, J. C., Santos, T. L., da Silva, M. A., . . . Almeida, F. C. (2014). Peptide: lipid ratio and membrane surface charge determine the mechanism of action of the antimicrobial peptide BP100. Conformational and functional studies. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1838(7), 1985-1999.
- Markley, J. L., Bax, A., Arata, Y., Hilbers, C., Kaptein, R., Sykes, B. D., . . . Wüthrich, K. (1998). Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J Mol Biol*, 280(5), 933-952.
- Melo, M. N., & Castanho, M. A. (2012). The Mechanism of Action of Antimicrobial Peptides: Lipid Vesicles vs. Bacteria. *Front Immunol*, 3, 236. doi:10.3389/fimmu.2012.00236
- Melo, M. N., Ferre, R., & Castanho, M. A. (2009). Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations. *Nature Reviews Microbiology*, 7(3), 245-250.
- Meyer-Hoffert, U. (2008). Neutrophil-derived serine proteases modulate innate immune responses. *Frontiers in bioscience (Landmark edition)*, 14, 3409-3418.
- Mirassou, Y., Santiveri, C. M., Pérez de Vega, M. J., González-Muñiz, R., & Jiménez, M. (2009). Disulfide Bonds versus Trp... Trp Pairs in Irregular β -Hairpins: NMR Structure of Vammin Loop 3-Derived Peptides as a Case Study. *Chembiochem*, 10(5), 902-910.
- Misiewicz, J., Afonin, S., Grage, S. L., van den Berg, J., Strandberg, E., Wadhwani, P., & Ulrich, A. S. (2015). Action of the multifunctional peptide BP100 on native biomembranes examined by solid-state NMR. *Journal of Biomolecular NMR*, 61(3-4), 287-298. doi:10.1007/s10858-015-9897-8
- Mohanty, D. P., Mohapatra, S., Misra, S., & Sahu, P. S. (2015). Milk derived bioactive peptides and their impact on human health – A review. *Saudi Journal of Biological Sciences*. doi:<http://dx.doi.org/10.1016/j.sjbs.2015.06.005>
- Mompeán García, M. Á. (2015). *Structural and computational studies of amyloids and noxious folds in biomolecules*. Universidad Autónoma de Madrid.

- Mulder, K. C. L., Lima, L. A., Miranda, V. J., Dias, S. C., & Franco, O. L. (2013). Current scenario of peptide-based drugs: the key roles of cationic antitumor and antiviral peptides. *New edge of antibiotic development: antimicrobial peptides and corresponding resistance*, 4, 40-62.
- Murakami, M., Lopez-Garcia, B., Braff, M., Dorschner, R. A., & Gallo, R. L. (2004). Postsecretory processing generates multiple cathelicidins for enhanced topical antimicrobial defense. *The Journal of Immunology*, 172(5), 3070-3077.
- Nagant, C., Pitts, B., Nazmi, K., Vandenbranden, M., Bolscher, J., Stewart, P., & Dehaye, J.-P. (2012). Identification of peptides derived from the human antimicrobial peptide LL-37 active against biofilms formed by *Pseudomonas aeruginosa* using a library of truncated fragments. *Antimicrobial agents and chemotherapy*, 56(11), 5698-5708.
- Oldfield, E., Meadows, M., Rice, D., & Jacobs, R. (1978). Spectroscopic studies of specifically deuterium labeled membrane systems. Nuclear magnetic resonance investigation of the effects of cholesterol in model systems. *Biochemistry*, 17(14), 2727-2740.
- Ozdirekcan, S., Rijkers, D. T., Liskamp, R. M., & Killian, J. A. (2005). Influence of flanking residues on tilt and rotation angles of transmembrane peptides in lipid bilayers. A solid-state 2H NMR study. *Biochemistry*, 44(3), 1004-1012. doi:10.1021/bi0481242
- Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Science*, 4(11), 2411-2423.
- Papo, N., Oren, Z., Pag, U., Sahl, H.-G., & Shai, Y. (2002). The consequence of sequence alteration of an amphipathic α -helical antimicrobial peptide and its diastereomers. *Journal of Biological Chemistry*, 277(37), 33913-33921.
- Pärn, K., Eriste, E., & Langel, Ü. (2015). The Antimicrobial and Antiviral Applications of Cell-Penetrating Peptides. *Cell-Penetrating Peptides: Methods and Protocols*, 223-245.
- Patel, L. N., Zaro, J. L., & Shen, W.-C. (2007). Cell Penetrating Peptides: Intracellular Pathways and Pharmaceutical Perspectives. *Pharmaceutical Research*, 24(11), 1977-1992. doi:10.1007/s11095-007-9303-7
- Ponnappan, N., Budagavi, D. P., Yadav, B. K., & Chugh, A. (2015). Membrane-Active Peptides from Marine Organisms—Antimicrobials, Cell-Penetrating Peptides and Peptide Toxins: Applications and Prospects. *Probiotics and Antimicrobial Proteins*, 7(1), 75-89. doi:10.1007/s12602-014-9182-2
- Pretzel, J., Mohring, F., Rahlfs, S., & Becker, K. (2013). Antiparasitic Peptides. In A. Vilcinskis (Ed.), *Yellow Biotechnology I: Insect Biotechnology in Drug Discovery and Preclinical Research* (pp. 157-192). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Provencher, S. W., & Glockner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, 20(1), 33-37.
- R Dennison, S., HG Morton, L., & A Phoenix, D. (2012). Effect of amidation on the antimicrobial peptide aurein 2.5 from Australian southern bell frogs. *Protein and peptide letters*, 19(6), 586-591.
- Rance, M., & Byrd, R. A. (1983). Obtaining high-fidelity spin-12 powder spectra in anisotropic media: Phase-cycled Hahn echo spectroscopy. *Journal of Magnetic Resonance* (1969), 52(2), 221-240.
- Reichert, J., Grasnick, D., Afonin, S., Buerck, J., Wadhwani, P., & Ulrich, A. S. (2006). A critical evaluation of the conformational requirements of fusogenic peptides in membranes. *European Biophysics Journal*, 36(4), 405-413. doi:10.1007/s00249-006-0106-2
- Rodrigues, M., Santos, A., Beatriz, G., Rádis-Baptista, G., Andreu, D., & Santos, N. C. (2012). Molecular characterization of the interaction of crotamine-derived nucleolar

- p>targeting peptides with lipid membranes.
- Biochimica et Biophysica Acta (BBA)-Biomembranes*
- , 1818(11), 2707-2717.
- Sato, H., & Feix, J. B. (2006). Peptide–membrane interactions and mechanisms of membrane destruction by amphipathic α -helical antimicrobial peptides. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1758(9), 1245-1256.
- Schubert, M., Labudde, D., Oschkinat, H., & Schmieder, P. (2002). A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ¹³C chemical shift statistics. *Journal of Biomolecular NMR*, 24(2), 149-154.
- Schweizer, F. (2009). Cationic amphiphilic peptides with cancer-selective toxicity. *European journal of pharmacology*, 625(1), 190-194.
- Seo, M.-D., Won, H.-S., Kim, J.-H., Mishig-Ochir, T., & Lee, B.-J. (2012). Antimicrobial peptides for therapeutic applications: a review. *Molecules*, 17(10), 12276-12286.
- Shai, Y. (2002). Mode of action of membrane active antimicrobial peptides. *Peptide Science*, 66(4), 236-248.
- Shai, Y., & Oren, Z. (1996). Diastereomers of cytolysins, a novel class of potent antibacterial peptides. *Journal of Biological Chemistry*, 271(13), 7305-7308.
- Sharma, A., Singla, D., Rashid, M., & Raghava, G. P. S. (2014). Designing of peptides with desired half-life in intestine-like environment. *BMC bioinformatics*, 15(1), 282.
- Sreerama, N., Venyaminov, S. Y., & Woody, R. W. (1999). Estimation of the number of α -helical and β -strand segments in proteins using circular dichroism spectroscopy. *Protein Science*, 8(2), 370-380. doi:10.1110/ps.8.2.370
- Sreerama, N., Venyaminov, S. Y., & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. *Analytical Biochemistry*, 287(2), 243-251. doi:10.1006/abio.2000.4879
- Sreerama, N., & Woody, R. W. (1993). A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Analytical Biochemistry*, 209(1), 32-44. doi:10.1006/abio.1993.1079
- Steiner, H., Hultmark, D., Engstrom, A., Bennich, H., & Barman, H. (1981). Sequence and specificity of two antibacterial proteins involved in insect immunity. *Nature*, 292, 246-248.
- Strandberg, E., Esteban-Martin, S., Salgado, J., & Ulrich, A. S. (2009). Orientation and dynamics of peptides in membranes calculated from 2H-NMR data. *Biophys J*, 96(8), 3223-3232. doi:10.1016/j.bpj.2009.02.040
- Strandberg, E., Esteban-Martin, S., Ulrich, A. S., & Salgado, J. (2012). Hydrophobic mismatch of mobile transmembrane helices: Merging theory and experiments. *Biochim Biophys Acta*, 1818(5), 1242-1249. doi:10.1016/j.bbamem.2012.01.023
- Strandberg, E., Kanithasen, N., Tiltak, D., Burck, J., Wadhwani, P., Zwernemann, O., & Ulrich, A. S. (2008). Solid-state NMR analysis comparing the designer-made antibiotic MSI-103 with its parent peptide PGLa in lipid bilayers. *Biochemistry*, 47(8), 2601-2616. doi:10.1021/bi701944r
- Strandberg, E., Ozdirekcan, S., Rijkers, D. T., van der Wel, P. C., Koeppe, R. E., 2nd, Liskamp, R. M., & Killian, J. A. (2004). Tilt angles of transmembrane model peptides in oriented and non-oriented lipid bilayers as determined by 2H solid-state NMR. *Biophys J*, 86(6), 3709-3721. doi:10.1529/biophysj.103.035402
- Strandberg, E., Tiltak, D., Ehni, S., Wadhwani, P., & Ulrich, A. S. (2012). Lipid shape is a key factor for membrane interactions of amphipathic helical peptides. *Biochim Biophys Acta*, 1818(7), 1764-1776. doi:10.1016/j.bbamem.2012.02.027
- Strandberg, E., Tiltak, D., Ieronimo, M., Kanithasen, N., Wadhwani, P., & Ulrich, A. S. (2007). Influence of C-terminal amidation on the antimicrobial and hemolytic activities of cationic α -helical peptides. *Pure and Applied Chemistry*, 79(4), 717-728.

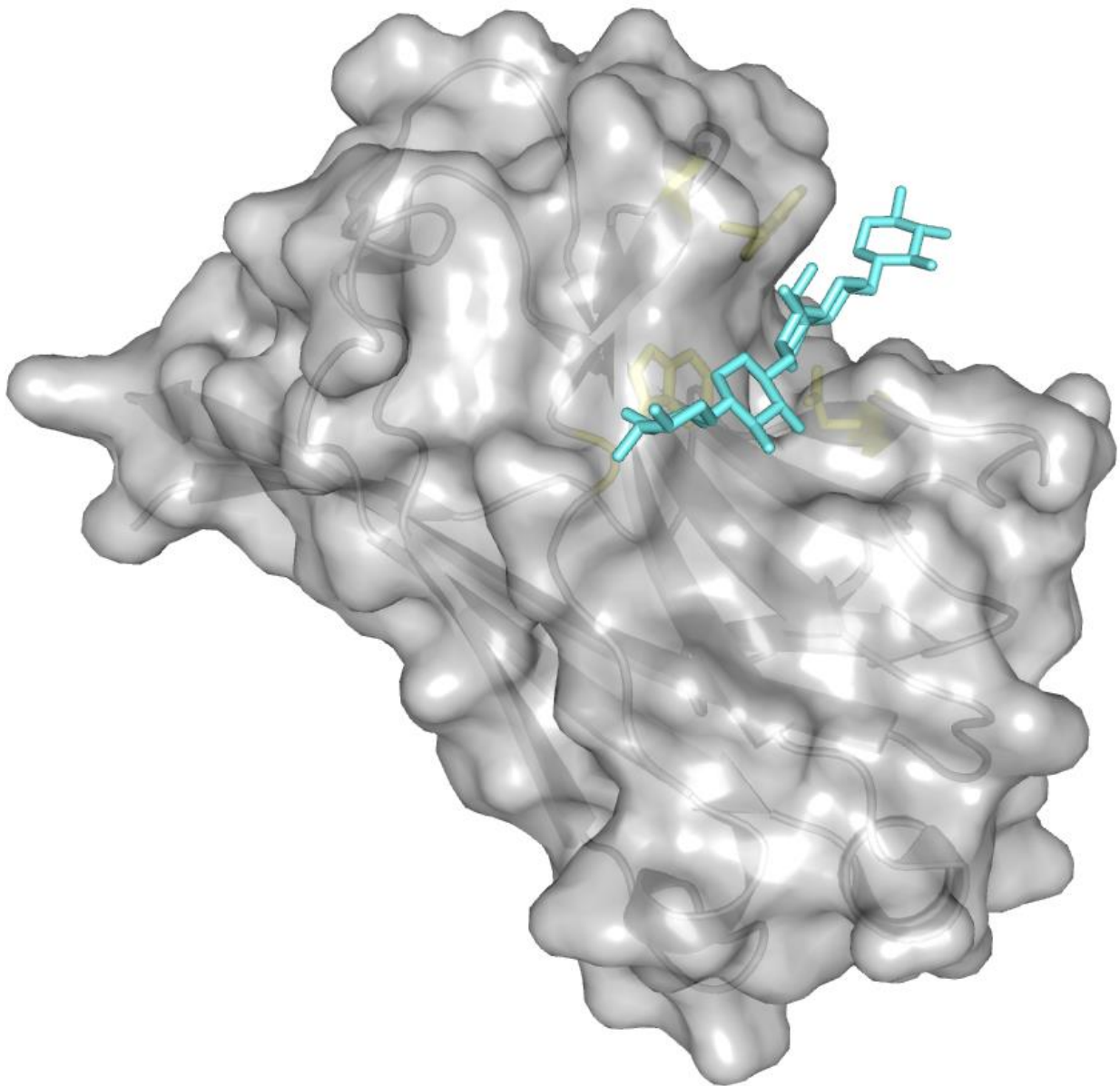
- Strandberg, E., & Ulrich, A. S. (2015). AMPs and OMPs: Is the folding and bilayer insertion of β -stranded outer membrane proteins governed by the same biophysical principles as for α -helical antimicrobial peptides? *Biochimica et Biophysica Acta (BBA)-Biomembranes*.
- Strandberg, E., Wadhwani, P., Tremouilhac, P., Durr, U. H., & Ulrich, A. S. (2006). Solid-state NMR analysis of the PGLa peptide orientation in DMPC bilayers: structural fidelity of ^2H -labels versus high sensitivity of ^{19}F -NMR. *Biophys J*, 90(5), 1676-1686. doi:10.1529/biophysj.105.073858
- Strandberg, E., Zerweck, J., Wadhwani, P., & Ulrich, A. S. (2013). Synergistic insertion of antimicrobial magainin-family peptides in membranes depends on the lipid spontaneous curvature. *Biophys J*, 104(6), L9-11. doi:10.1016/j.bpj.2013.01.047
- Strömstedt, A. A., Pasupuleti, M., Schmidtchen, A., & Malmsten, M. (2009). Evaluation of strategies for improving proteolytic resistance of antimicrobial peptides by using variants of EFK17, an internal segment of LL-37. *Antimicrobial agents and chemotherapy*, 53(2), 593-602.
- Terwilliger, T. C. (1982). The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature*, 299, 371-374.
- Tkachenko, A. N., Mykhailiuk, P. K., Afonin, S., Radchenko, D. S., Kubyshkin, V. S., Ulrich, A. S., & Komarov, I. V. (2013). A ^{19}F NMR Label to Substitute Polar Amino Acids in Peptides: A CF_3 -Substituted Analogue of Serine and Threonine. *Angewandte Chemie International Edition*, 52(5), 1486-1489.
- Torcatto, I. M., Huang, Y.-H., Franquelim, H. G., Gaspar, D., Craik, D. J., Castanho, M. A., & Henriques, S. T. (2013). Design and characterization of novel antimicrobial peptides, R-BP100 and RW-BP100, with activity against Gram-negative and Gram-positive bacteria. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(3), 944-955.
- Torchilin, V. P. (2005). Recent advances with liposomes as pharmaceutical carriers. *Nature reviews Drug discovery*, 4(2), 145-160.
- Torrent, M., Beatriz, G., Nogués, V. M., Andreu, D., & Boix, E. (2009). Bactericidal and membrane disruption activities of the eosinophil cationic protein are largely retained in an N-terminal fragment. *Biochemical Journal*, 421(3), 425-434.
- Torrent, M., Di Tommaso, P., Pulido, D., Nogués, M. V., Notredame, C., Boix, E., & Andreu, D. (2012). AMPA: an automated web server for prediction of protein antimicrobial regions. *Bioinformatics*, 28(1), 130-131.
- Tremouilhac, P., Strandberg, E., Wadhwani, P., & Ulrich, A. S. (2006a). Conditions affecting the re-alignment of the antimicrobial peptide PGLa in membranes as monitored by solid state ^2H -NMR. *Biochim Biophys Acta*, 1758(9), 1330-1342. doi:10.1016/j.bbamem.2006.02.029
- Tremouilhac, P., Strandberg, E., Wadhwani, P., & Ulrich, A. S. (2006b). Synergistic transmembrane alignment of the antimicrobial heterodimer PGLa/magainin. *J Biol Chem*, 281(43), 32089-32094. doi:10.1074/jbc.M604759200
- Ulmschneider, J. P., Smith, J. C., Ulmschneider, M. B., Ulrich, A. S., & Strandberg, E. (2012). Reorientation and dimerization of the membrane-bound antimicrobial peptide PGLa from microsecond all-atom MD simulations. *Biophys J*, 103(3), 472-482. doi:10.1016/j.bpj.2012.06.040
- van der Wel, P. C., Strandberg, E., Killian, J. A., & Koeppe, R. E., 2nd. (2002). Geometry and intrinsic tilt of a tryptophan-anchored transmembrane α -helix determined by (^2H) NMR. *Biophys J*, 83(3), 1479-1488. doi:10.1016/s0006-3495(02)73918-0
- van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R., & Groen, F. C. (1990). Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Analytical Biochemistry*, 191(1), 110-118.

- Vila, R., Ponte, I., Suau, P., & Rico, M. (2000). A helix-turn motif in the C-terminal domain of histone H1. *Protein Science*, 9(4), 627-636.
- Vostrikov, V. V., Grant, C. V., Daily, A. E., Opella, S. J., & Koeppe, R. E., 2nd. (2008). Comparison of "Polarization inversion with spin exchange at magic angle" and "geometric analysis of labeled alanines" methods for transmembrane helix alignment. *J Am Chem Soc*, 130(38), 12584-12585. doi:10.1021/ja803734k
- Wachinger, M., Kleinschmidt, A., Winder, D., von Pechmann, N., Ludvigsen, A., Neumann, M., . . . Brack-Werner, R. (1998). Antimicrobial peptides melittin and cecropin inhibit replication of human immunodeficiency virus 1 by suppressing viral gene expression. *Journal of General Virology*, 79(4), 731-740. doi:doi:10.1099/0022-1317-79-4-731
- Wade, D., Boman, A., Wählin, B., Drain, C. M., Andreu, D., Boman, H. G., & Merrifield, R. B. (1990). All-D amino acid-containing channel-forming antibiotic peptides. *Proceedings of the National Academy of Sciences of the United States of America*, 87(12), 4761-4765.
- Wadhwani, P., Burck, J., Strandberg, E., Mink, C., Afonin, S., & Ulrich, A. S. (2008). Using a sterically restrictive amino acid as a 19F NMR label to monitor and to control peptide aggregation in membranes. *J Am Chem Soc*, 130(49), 16515-16517.
- Wadhwani, P., Epand, R. F., Heidenreich, N., Burck, J., Ulrich, A. S., & Epand, R. M. (2012). Membrane-active peptides and the clustering of anionic lipids. *Biophys J*, 103(2), 265-274. doi:10.1016/j.bpj.2012.06.004
- Wadhwani, P., Reichert, J., Burck, J., & Ulrich, A. S. (2012). Antimicrobial and cell-penetrating peptides induce lipid vesicle fusion by folding and aggregation. *Eur Biophys J*, 41(2), 177-187. doi:10.1007/s00249-011-0771-7
- Wadhwani, P., Strandberg, E., Heidenreich, N., Bürck, J., Fanghänel, S., & Ulrich, A. S. (2012). Self-assembly of flexible β -strands into immobile amyloid-like β -sheets in membranes as revealed by solid-state 19F NMR. *Journal of the American Chemical Society*, 134(15), 6512-6515.
- Wadhwani, P., Strandberg, E., van den Berg, J., Mink, C., Burck, J., Ciriello, R. A., & Ulrich, A. S. (2014). Dynamical structure of the short multifunctional peptide BP100 in membranes. *Biochim Biophys Acta*, 1838(3), 940-949. doi:10.1016/j.bbamem.2013.11.001
- Wang, G., Mishra, B., Epand, R. F., & Epand, R. M. (2014). High-quality 3D structures shine light on antibacterial, anti-biofilm and antiviral activities of human cathelicidin LL-37 and its fragments. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1838(9), 2160-2172.
- Wang, G., Mishra, B., Lau, K., Lushnikova, T., Golla, R., & Wang, X. (2015). Antimicrobial peptides in 2014. *Pharmaceuticals*, 8(1), 123-150.
- Wang, Y., Zhao, T., Wei, D., Strandberg, E., Ulrich, A. S., & Ulmschneider, J. P. (2014). How reliable are molecular dynamics simulations of membrane active antimicrobial peptides? *Biochim Biophys Acta*, 1838(9), 2280-2288. doi:10.1016/j.bbamem.2014.04.009
- White, S. H., & Wimley, W. C. (1999). Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct*, 28, 319-365. doi:10.1146/annurev.biophys.28.1.319
- Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32(Web Server issue), W668-673. doi:10.1093/nar/gkh371
- Wieprecht, T., Dathe, M., Beyermann, M., Krause, E., Maloy, W. L., MacDonald, D. L., & Bienert, M. (1997). Peptide hydrophobicity controls the activity and selectivity of

- magainin 2 amide in interaction with membranes. *Biochemistry*, 36(20), 6124-6132. doi:10.1021/bi9619987
- Wieprecht, T., Dathe, M., Krause, E., Beyermann, M., Maloy, W. L., MacDonald, D. L., & Bienert, M. (1997). Modulation of membrane activity of amphipathic, antibacterial peptides by slight modifications of the hydrophobic moment. *FEBS Lett*, 417(1), 135-140.
- Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., & Sykes, B. D. (1995). ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR*, 5(1), 67-81.
- Wüthrich, K. (1986). *NMR of proteins and nucleic acids* (Vol. 155): Wiley New York.
- Wüthrich, K., Billeter, M., & Braun, W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. *J Mol Biol*, 180(3), 715-740.
- Yeung, A. T., Gellatly, S. L., & Hancock, R. E. (2011). Multifunctional cationic host defence peptides and their clinical applications. *Cellular and Molecular Life Sciences*, 68(13), 2161-2176.
- Zanetti, M. (2004). Cathelicidins, multifunctional peptides of the innate immunity. *Journal of leukocyte biology*, 75(1), 39-48.
- Zanetti, M., Gennaro, R., & Romeo, D. (1995). Cathelicidins: a novel protein family with a common proregion and a variable C-terminal antimicrobial domain. *FEBS Letters*, 374(1), 1-5. doi:10.1016/0014-5793(95)01050-O
- Zasloff, M. (2002). Antimicrobial peptides of multicellular organisms. *Nature*, 415(6870), 389-395.
- Zasloff, M. (2016). Antimicrobial Peptides: Do They Have a Future as Therapeutics? In J. Harder & J.-M. Schröder (Eds.), *Antimicrobial Peptides: Role in Human Health and Disease* (pp. 147-154). Cham: Springer International Publishing.
- Zelezetsky, I., & Tossi, A. (2006). Alpha-helical antimicrobial peptides--using a sequence template to guide structure-activity relationship studies. *Biochim Biophys Acta*, 1758(9), 1436-1449. doi:10.1016/j.bbamem.2006.03.021
- Zhang, Y., Zhao, H., Yu, G.-Y., Liu, X.-D., Shen, J.-H., Lee, W.-H., & Zhang, Y. (2010). Structure-function relationship of king cobra cathelicidin. *Peptides*, 31(8), 1488-1493.
- Zorko, M., & Langel, Ü. (2005). Cell-penetrating peptides: mechanism and kinetics of cargo delivery. *Advanced drug delivery reviews*, 57(4), 529-545. doi:<http://dx.doi.org/10.1016/j.addr.2004.10.010>

CHAPTER 4

CARBOHYDRATE-BINDING MODULES (CBMs)



The work included in this chapter was accomplished thanks to the collaboration with the group of Prof. Mayte Villalba and Prof. Rosalía Rodríguez, from Department of Biochemistry and Molecular Biology I (Universidad Complutense de Madrid, Madrid, Spain). Dr. María Torres, from the aforesaid department, was responsible for the biochemical part of the experimentation. Dr. Noemí Bustamante and Dr. Miguel Mompeán, both from Instituto de Química Física Rocasolano (Madrid, Spain), afforded their assistance for the analytical ultracentrifugation and molecular docking, respectively. Dr. Anjos Macedo from Universidade Nova de Lisboa (Lisboa, Portugal) also collaborated in some preliminary aspects of the internal dynamics and binding of rCtD–Ole e 9.

INTRODUCTION

1. β -1,3-GLUCANASES

β -1,3-glucanases (EC 3.2.1.39) are proteins belonging to the glycosylhydrolase family (GHF), *i.e.* enzymes which are capable of hydrolysing glycosidic bonds. In particular, β -1,3-glucanases break β (1 \rightarrow 3) glycosidic bonds present in non-branched segments of some kinds of D-glucose polysaccharides and oligosaccharides. These proteins are expressed by a wide variety of organisms, including viruses, archaea, bacteria, fungi, algae, superior plants and even in invertebrate animals. Presently, a great number of them have been characterised, showing an extensive diversity of substrate specificities and mechanisms of action. (Martin et al., 2007; Sun, Gurnon, Adams, Graves, & Van Etten, 2000; Wojtkowiak, Witek, Hennig, & Jaskolski, 2012)

1.1 β -1,3-GLUCANS

β -1,3-glucans (Table 4.1) are the natural substrates of these enzymes. They are polysaccharides composed of D-glucose units connected by β (1 \rightarrow 3) glycosidic bonds. (Figure 4.1). Occasionally, these substrates present complex structures (e.g., the curdlan triple helix), branching (e.g., the β (1 \rightarrow 6) branches in the callose) or a combination of different types of glycosidic bonds (e.g., β -1,3/1,4-glucans). Hydrolysis products derived from the β -1,3-glucanase activity are both β -1,3-oligoglucosides formed by 2–9 glucose units, and isolated glucose units. (Martín-Cuadrado et al., 2008; Wojtkowiak et al., 2012)

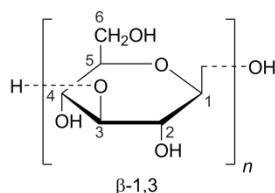


Figure 4.1. Representation of the chemical structure of the general repetitive unit of the β -1,3-glucan backbone (D-glucose connected by β (1 \rightarrow 3) bonds).

These polysaccharides are fundamentally found in yeast and fungi cell walls, reaching up to 50% of the dry mass of these structures. The function of β -1,3-glucans in yeast and fungi is to provide rigidity and mechanic resistance to the cell wall, as this structure is crucial to preserve the osmotic integrity and it determines the cell morphology in the different stages of the cell cycle. Some β -1,3-glucans present in these organisms are found in the cell surface and are thought to be involved in morphogenetic change processes, as aggregation or mycelial strand formation. (Martín-Cuadrado et al., 2008; Stone & Clarke, 1992; Wojtkowiak et al., 2012)

β -1,3-glucans can also be found in other organisms, like in algae, where they are part of the cell walls and act as storage polysaccharides. This is the case of laminarin, the most extensively studied β -1,3-glucan (β -1,3-glucanases are often called “laminarinases” for this reason). Some bacteria (*Agrobacterium* sp, *Rhizobium* sp,...) produce curdlan, a β -1,3-glucan

involved in cell adhesion processes and protection against external agents. In superior plants, a β -1,3-glucan called callose gives rigidity and stability to cell walls. (Bacic, Fincher, & Stone, 2009; Coutinho & Henrissat, 1999; Lashbrook, Gonzalez-Bosch, & Bennett, 1994; Qin, Yan, Yang, & Jiang, 2015)

Table 4.1. Examples of β -1,3-glucans and some of their properties.*

Name	Backbone bonds	Branching [†]	Average DP [‡]	Organism	Function
Curdlan	β (1→3)	-	~200–400	Bacteria	Protection, adhesion,...
Laminarin	β (1→3)	β (1→6) (7:1)	~25	Brown algae	Structural, storage,...
Escleroglucan	β (1→3)	β (1→6) (3:1)	~100–800	Fungi	Structural
Chrysolaminarin	β (1→3)	β (1→6) (11:1)	~20	Phytoplankton	Storage
Lentinan	β (1→3)	β (1→6) (6:1)	~10–20	Fungi	Structural
Lichenan	β (1→3)/ β (1→4)	-	~400	Lichen	Structural
Callose	β (1→3)	-	~500	Plants	Structural

*Data from (Bacic et al., 2009; Beattie, Hirst, & Percival, 1961; Bluhm, Deslandes, Marchessault, Pérez, & Rinaudo, 1982; Him, Pelosi, Chanzy, Putaux, & Bulone, 2001; Martin et al., 2007; Mohd Jamil et al., 2013; Peat, Whelan, & Roberts, 1957; Read, Currie, & Bacic, 1996; Steinbüchel, 2002) [†]Branched bonds:backbone bonds ratio is indicated in parentheses. [‡]DP: degree of polymerization.

1.2 FUNCTIONS OF THE β -1,3-GLUCANASES

The functions performed by β -1,3-glucanases and the biological processes in which they are involved are diverse and depend on the organism, but they are always directly related with the nature of their catalytic activity, *i.e.* the hydrolysis of β -1,3-glucans.

Viral β -1,3-glucanases are thought to be implicated in the host cell wall degradation during the processes of internalization or externalization of the viral particle. In archaea, these enzymes participate in fermentation, whereas in other bacteria they are part of the defence mechanism against fungi and they also present functions related to cell metabolism. (Wojtkowiak et al., 2012)

In algae, the main function of β -1,3-glucanases is degradation of storage polymers to obtain glucose units as an energy source. (Wojtkowiak et al., 2012) Undoubtedly, β -1,3-glucanases play an outstanding role in fungi and yeast due to the huge substrate availability in these organisms, as β -1,3-glucans are widespread in their cell walls and surfaces. Cell wall assembly and reorganization during growth, morphogenesis, germination, sporulation, conjugation, and degradation of storage polymers after exhaustion of external nutrition are some of the processes in which these enzymes take part actively in fungi and yeast. (Bielecki & Galas, 1991; Ishida et al., 2009; Mouyna, Hartl, & Latgé, 2013)

In the animal kingdom, β -1,3-glucanases are restricted to a few invertebrates. This is the case of some kind of nematodes that express these enzymes to degrade the fungi they feed on. (Kikuchi, Shibuya, & Jones, 2005)

β -1,3-glucanases are very important enzymes in the plant kingdom, since they possess a double functionality depending on the origin of the substrate on which they act:

- On one hand, these enzymes can perform their function on β -1,3-glucans present in plant cells, such as callose, that give resistance and firmness to the cell walls.

Degradation of this kind of β -1,3-glucans takes place in the course of several natural events, such as cell division, microsporogenesis, pollen development, seed germination, fertilization, embryogenesis, fruit ripening, or flowering. (Dogra & Sreenivasulu, 2015; Leubner-Metzger, Frundt, Vogeli-Lange, & Meins, 1995; Wojtkowiak et al., 2012)

- On the other hand, they can act on exogenous β -1,3-glucans, principally those from fungi which cause pathologies in plants. Therefore, in plants β -1,3-glucanases not only participate in their metabolism and physiology, but also they play a key role in the plant defence system against fungal phytopathogens. In this context, plant β -1,3-glucanases form the group 2 of the so-called “pathogenesis-related proteins” (PRs), which contribute to the plant defence against fungal infections acting both isolated and, especially, in combination with other enzymes, such as chitinase. These proteins perform their defensive function against fungal pathogens in a double way: directly, weakening and degrading their cell walls, and indirectly, releasing elicitors through the limited hydrolysis of the fungal cell wall that triggers a defence reaction cascade. (Gao et al., 2016; Singh, Ambroise, Haicour, Sihachakr, & Rajam, 2014)

Even though these enzymes are always present in the plant organism, it has been reported the induction of their expression activated by the infections with pathogens in some plant species. Localization of these enzymes in plant cells is diverse, as the presence of basic β -1,3-glucanases inside vacuoles has been confirmed, whereas several acid β -1,3-glucanases are considered extracellular proteins. (Gao et al., 2016; Liu, He, Li, Chen, & Ge, 2012)

1.3 CLASSIFICATION OF THE β -1,3-GLUCANASES

β -1,3-glucanases are categorized in the EC classification in the glycosidase group (EC 3.2.1), and they are also part of the glycosylhydrolase family (GHF). According to the way in which they degrade the substrate, β -1,3-glucanases are divided into two classes: (Lombard, Golaconda Ramulu, Drula, Coutinho, & Henrissat, 2014; Martín-Cuadrado et al., 2008; Mouyna et al., 2013)

- Exo- β -1,3-glucanases (EC 3.2.1.58): they hydrolyse β -1,3-glucans from the non-reducing end of the polymeric chain, releasing glucose units. Within GHF, they are classified in the subfamilies GH-3, GH-5, GH-17, and GH-55.
- Endo- β -1,3-glucanases (EC 3.2.1.39): they hydrolyse β -1,3-glucans from internal positions of the polymeric chain, releasing a mixture of oligosaccharides, being glucose a minor product. Within GHF, they are classified in the subfamilies GH-16, GH-17, GH-55, GH-64, GH-81, and GH-128.

Endo- β -1,3-glucanases grouped in the different GHF subfamilies act on similar substrates, but they are expressed by diverse kinds of organisms and evolved with distinct folds (Table 4.2).

Table 4.2. Subfamilies of the endo- β -1,3-glucanases.*

GH subfamily	Organisms	Fold
16	Archaea, bacteria, plants, viruses	β -sandwich jelly-roll
17	Plants, fungi, bacteria	$(\beta/\alpha)_8$ -TIM barrel
55	Fungi, bacteria	right-handed parallel β -helix
64	Bacteria, fungi	β -barrel / mixed α/β
81	Fungi, bacteria, plants	?
128	Fungi, bacteria, viruses	$(\beta/\alpha)_8$ -barrel

*Data from (Lombard et al., 2014; Wojtkowiak et al., 2012)

1.4 PLANT ENDO- β -1,3-GLUCANASES

As explained before, β -1,3-glucanases perform very significant functions in plant organisms. Plant endo- β -1,3-glucanases belong to the pathogen-related protein group 2 (PR-2). There are four classes of endo- β -1,3-glucanases in this group: those from class I are located in vacuoles, whereas those from classes II–IV are secreted to the apoplastic space of plant cells. (Fujimori et al., 2016)

Most plant endo- β -1,3-glucanases, forming part of GH-17 subfamily, show a big diversity in their specificity, an evidence of the result of a divergent evolution with a preservation of the catalytic mechanism. All these enzymes share a $(\beta/\alpha)_8$ -barrel fold with two preserved catalytic Glu residues located close to the C-terminal ends of β -strands 4 and 7. (Chen, Garrett, Fincher, & Høj, 1995; Leubner-Metzger et al., 1995; Wojtkowiak et al., 2012)

Endo- β -1,3-glucanases are enzymes of notable interest in agriculture with regard to their role in the defence of plant integrity against fungal infections. Numerous plant species of interest, such as cereals, grapevine, some vegetables or tobacco, are prone to the attack of phytopathogenic fungal organisms, what can result in substantial economic losses. One of the strategies explored to fight these infections is the expression of endo- β -1,3-glucanases-encoding genes in transgenic plants, what has proved to confer an additional resistance to the pathogen fungi. (Ceasar & Ignacimuthu, 2012; Fujimori et al., 2016; Kirubakaran & Sakthivel, 2007; Singh et al., 2014)

In the clinical field, plant endo- β -1,3-glucanases are object of continuous investigation, given that they possess allergenic properties and are present in pollen grains. These enzymes have been identified as the most allergenic components among the natural latex rubber proteins, and as cross-reactivity allergens in the latex-fruit syndrome. (Huecas, Villalba, & Rodríguez, 2001; Sunderasan et al., 1995; Wagner et al., 2004)

2. CARBOHYDRATE-BINDING MODULES (CBMs)

β -1,3-glucanases, like other glycosylhydrolases and, in general, many carbohydrate-active enzymes, frequently show a modular structure composed of a catalytic domain, responsible for the enzymatic function, and one to several carbohydrate-binding modules (CBMs). A CBM is defined as a contiguous amino acid sequence within a carbohydrate-

active enzyme with a discrete fold having carbohydrate-binding activity. (Hashimoto, 2006; Shoseyov, Shani, & Levy, 2006) The primary function of CBMs is to increase catalytic efficiency of carbohydrate-active enzymes against soluble and/or insoluble substrates by bringing the catalytic module into prolonged and intimate contact with substrates. (Hashimoto, 2006) In this regard, a new question has arisen in the last times, as it has been suggested that CBMs can also enhance enzyme activity by physically disrupting the structure of the saccharide substrates. (Davies & Williams, 2016)

2.1 CLASSIFICATION AND STRUCTURE OF THE CBMs

CBMs are grouped in families and, currently, there are 71 different families described. This classification comes from a previous one including only cellulose-binding modules and which was based on sequence similarity. Nevertheless, criteria followed for the creation of these families are not very intuitive and some authors have developed less formal classifications which seem to be more useful to sort this kind of modules. One of these classifications distinguishes the different CBMs according to their fold and establishes several “folding families”. (Table 4.3, Figure 4.2) (A. Boraston, D. Bolam, H. Gilbert, & G. Davies, 2004; Lombard et al., 2014)

Table 4.3. CBM folding families.*

Folding family	Fold	CBM families
1	β -sandwich	2, 3, 4, 6, 9, 11, 15–17, 20, 22–23, 25–33–41, 44–48, 51, 53–54, 56–57, 59, 61–63, 66–68, 70–71, 73
2	β -trefoil	13, 42
3	Cystine knot	1
4	Unique	5, 12
5	OB fold	10
6	Hevein fold	18
7	Unique, contains hevein-like fold	14

*Data from (A. Boraston et al., 2004; Carvalho et al., 2004; Hashimoto, 2006; Lombard et al., 2014; Sillitoe et al., 2015)

As observed in table 4.3, β -sandwich folding is the most common among CBMs and its general structure comprises two β -sheets, each consisting of three to six antiparallel β -strands (Figure 4.2, structures a, b, c, d, e, i, j). They usually have at least one bound metal atom which most times play a role as structure stabilizer, but sometimes it is directly involved in the carbohydrate-binding process. In the majority of these proteins, the binding site is located on the same face of the β -sandwich, but occasionally it is found at the edge of the β -sandwich. Some authors have suggested the division of the CBMs with β -sandwich fold into two subfamilies: CBMs with β -sandwich jelly-roll fold (β -barrel like) and CBMs with immunoglobulin-like β -sandwich (two β -sheets forming a sandwich with Greek key topology) (A. Boraston et al., 2004; Hashimoto, 2006; Kister, Fokas, Papatheodorou, & Gelfand, 2006; Woelfson, Evans, Hutchinson, & Thornton, 1993)

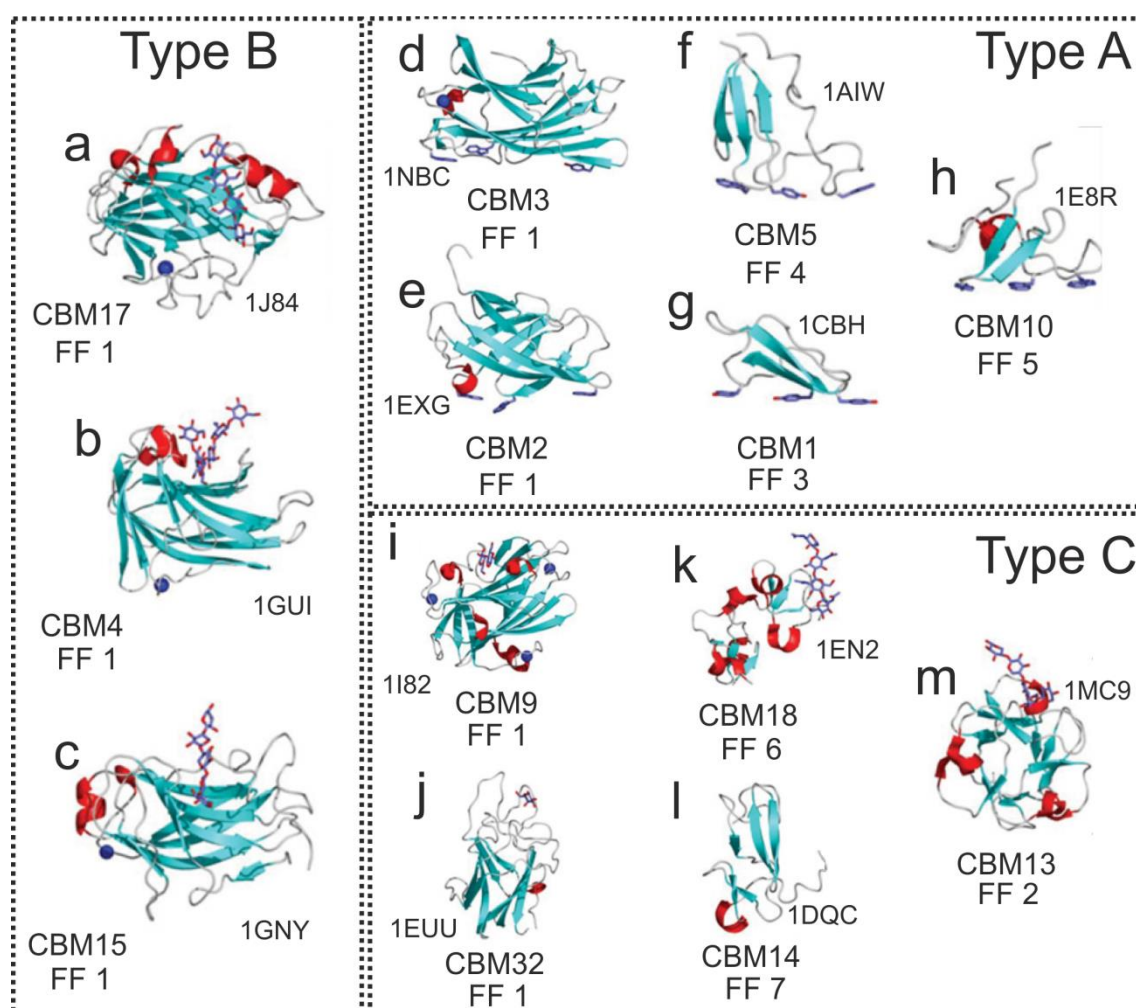


Figure 4.2. Examples of CBMs structures. For each structure PDB code, CBM family and folding family (FF) according to table 4.3 are indicated. The structures are grouped according to their carbohydrate-binding type (see classification in page 233). Figure adapted from (A. Boraston et al., 2004).

β -trefoil fold (Figure 4.2, structure m) comprises twelve β -strands forming six hairpin turns. A β -barrel is formed by six of the strands, corresponding to three hairpin turns. The other three hairpin turns form a triangular cap on one end of the β -barrel (“hairpin triplet”). Each of the three subunits forming this fold is called “trefoil domain” and contributes with one hairpin to the β -barrel and one hairpin to the hairpin triplet. In each domain, functional carbohydrate-binding sites are present. (A. Boraston et al., 2004)

Folding families 3 (cystine knot), 4 (unique) and 5 (OB fold) (Figure 4.2, structures g, f and h, respectively) are CBMs consisting of short polypeptides of 30–60 residues containing only β -sheet and coil regions. They show less diversity in their ligand specificity, since they are more specialized to recognise cellulose and/or chitin. These CBMs usually have carbohydrate-binding sites with a planar geometry comprising aromatic residues. (A. Boraston et al., 2004)

Hevein-like domains (Figure 4.2, structure k) are small CBMs (around 40 amino acids) and they usually are plant chitin-binding proteins. Their structure is predominantly coil, with two small β -sheets and a short helical region, and it has an extended carbohydrate-

binding site. CBM14 family (Figure 4.2, structure I) incorporates some aspects of the hevein domain, but it has an additional fused small β -sheet structure. (A. Boraston et al., 2004)

Many structurally characterised CBMs show folding motifs that cannot be categorized in the existing folding families. This is the case of the CBMs studied in this chapter, belonging to CBM43 family. Probably, the progress in the characterisation of the CBMs described up to now will permit to develop a more detailed and complete classification of the CBMs and to achieve a deeper understanding of the function–structure relationship in these proteins, their evolution and carbohydrate–recognition mechanisms.

2.2 CARBOHYDRATE–CBM INTERACTION

Although CBM families can be grouped according to the folding conservation, this classification is not enough to deduce the functionality of these modules due to big diversity of functional elements, as specific residues or binding–site topographies, existing among the members of each family. Thus, it has been proposed a different classification based on structural and functional similarities, according to which three classes of CBMs can be distinguished: (A. Boraston et al., 2004)

- Type A (surface-binding CBMs)

These CBMs show planar or platform-like carbohydrate-binding sites, which are thought to be complementary to the flat geometry of cellulose and chitin crystal surfaces. They have little or no affinity for soluble carbohydrates. It has been postulated that protein–carbohydrate interactions in these CBMs are driven by thermodynamical forces, with the participation of aromatic residues located in the interaction region. Different members of CBM 1, 2, 3, 5, and 10 families are classified as type A CBMs.

- Type B (glycan-chain-binding CBMs)

In these CBMs the interaction with carbohydrates takes place in extended regions, often described as grooves or clefts, which contain several small subsites able to accommodate the individual sugar units of the polymeric ligand. Polymerization degree (DP) of the ligand affects the binding efficiency, increasing with the chain length up to six units. Oligosaccharides composed by three or less monomers show a negligible interaction. Thus, these CBMs interact preferentially with individual glycan chains rather than crystalline surfaces. Ligand–specificity and affinity of these CBMs is directly conditioned by the aromatic residues located in the binding site, as well as by the formation of direct hydrogen bonds. Different members of CBM 2, 4, 6, 15, 17, 20, 22, 27, 28, 29, 34, and 43 families are classified as type B CBMs. The CBMs studied in this chapter belong to this type.

- Type C (small-sugar-binding CBMs)

These CBMs bind preferentially mono-, di- or trisaccharides, therefore they do not possess extended regions of interaction. These CBMs present some lectin-like properties and interaction takes place by a hydrogen-bonding network, which is more extensive than in type B CBMs. Different members of CBM 9, 13, 14, 18 and 32 families are classified as type C CBMs.

Binding-site topography of CBMs is a key determinant of their binding specificity. The two major factors contributing to model the binding-site are the location of aromatic amino acid side chains and orientation of loop structures. Tryptophan, tyrosine and, less commonly, phenylalanine or histidine side chains form hydrophobic platforms that can show planar, twisted or sandwich-like geometries in the CBMs binding sites. These platforms accommodate the sugar rings present in the carbohydrates interacting with CBMs. CBMs seem, in common with lectins, to contain preformed carbohydrate-recognition sites which mirror the solution conformations of their ligands, minimizing the energetic penalty paid upon binding. (A. Boraston et al., 2004)

Other ligand-specificity-defining interactions in the CBMs are direct hydrogen bonds or calcium-mediated coordination. Carbohydrates are amphipathic molecules with a high capacity for hydrogen-bond formation, due to the presence of multiple hydroxyl groups. Some observations point to the fact that the hydrogen bonds formed between those hydroxyl groups and polar residues located in the binding site of some CBMs may contribute to the ligand specificity and affinity. The role of this hydrogen-bond network seems to be more relevant in CBMs type B and C than in type A. (A. Boraston et al., 2004)

Additionally, several CBMs are metalloenzymes and their function depends on the presence of calcium atoms. The importance of these metal atoms comes from their ability to maintain the adequate protein folding or to establish direct interactions with the ligand, as described in some CBMs. (A. Boraston et al., 2004; Jamal-Talabani et al., 2004)

CBMs have been described as carbohydrate-binding proteins which show a relatively weak interaction with ligands ($K_d < 10^{-6}$ M). These interactions take place in quite open regions and leaving significant portions of the ligand exposed to solvent when bound. This can be understood if we keep in mind the fact that CBM ligands often have a big size and it is not possible to be completely enveloped in the binding site of the protein. Nevertheless, this weak interaction is often compensated by multivalent interactions (high avidity) derived from the presence of many clustered binding sites or the tandem repetition of CBMs, as seen in several glycosylhydrolases. CBMs repetition can increase the affinity from 10- to 100-fold, compared to an isolated CBM. (A. Boraston et al., 2004; Quiocho, 1986)

3. OLE E 9 AND FRA E 9: TWO HOMOLOGUE ENDO- β -1,3-GLUCANASES

As mentioned before, some plant endo- β -1,3-glucanases are found in pollen and they possess allergenic properties that have been reported in many works. A clear example is Ole e 9 glycoprotein from olive tree (*Olea europaea*) pollen, as well as its homologue in the European ash tree (*Fraxinus excelsior*) pollen, Fra e 9, both with endo- β -1,3-glucanase function (EC 3.2.1.39). They belong to the glycosylhydrolase family GH-17 and the pathology-related protein group PR-2. (Torres et al., 2015)

3.1 OLE E 9

Olive tree pollen is one of the main causes of pollinosis in the Mediterranean area, especially in Spain and Italy, where olive trees are very widespread. Among the 11 allergens known in the olive tree pollen, Ole e 9 shows a prevalence over 65% in the allergic population of the high pollen density regions. (Barber et al., 2008; Huecas et al., 2001; Villalba, Rodríguez, & Batanero, 2014)

Ole e 9 is an acid glycoprotein ($pI_{\text{theoretical}} = 5.62$) encoded by a 1380 nucleotide sequence which originates a 460 amino acid polypeptide chain, being the first 26 residues a signal peptide which is post-translationally processed to produce a mature protein with a theoretical molecular mass of 46.044 kDa (434 amino acids). This protein shows a modular structure composed by two independent domains: an N-terminal domain (NtD, MW \approx 35.46 kDa) and a C-terminal domain (CtD, MW \approx 10.58 kDa) connected by a linker.

The Ole e 9 N-terminal domain, produced in *Pichia pastoris*, contains the active site and it is responsible for the endo- β -1,3-glucanase activity of the protein. The C-terminal domain, also produced in *P. pastoris*, is a CBM belonging to CBM43 family, as the allergen Ole e 10, and, as all CBMs, it plays a key role in the regulation of the catalytic action of the complete protein. (Torres et al., 2015)

3.2 FRA E 9

European ash tree (*F. excelsior*) is a common species in Europe, especially in the northern part of the Iberian Peninsula, British islands, France, Central Europe and the southern region of Scandinavia. This tree, belonging to the same family as the olive tree (*Oleaceae*), is responsible for pollinosis in warm latitudes with a prevalence up to 20% among allergic patients of these regions, and therefore it should be regarded as a relevant source of allergenicity. (Torres, 2014)

Up to now, 8 allergens have been identified in European ash tree pollen and the most allergenically relevant are Fra e 1, Fra e 2, Fra e 9 and Fra e 11, being Fra e 1 the most prevalent (> 75%). In the Mediterranean area, most allergic patients show positive skin tests both to ash tree pollen and to olive tree pollen, and this is probably due to a cross-reactivity

between their allergenic proteins, a phenomenon observed before in members of the *Oleaceae* family. (Torres, 2014)

In previous works which revealed the presence of β -1,3-glucanases homologue to Ole e 9 in the European ash tree pollen, Fra e 9 allergen was identified. This allergen is a basic glycoprotein ($pI_{\text{theoretical}} = 8.52$) encoded by a 1383 nucleotide sequence which originates a 461 amino acid polypeptide chain, being the first 29 residues a signal peptide which is posttranslationally processed to produce a mature protein with a theoretical molecular mass of 46.948 kDa (432 amino acids). Like Ole e 9, this protein shows a modular structure composed by two independent domains: a N-terminal domain (NtD, MW \approx 36.14 kDa) and a C-terminal domain (CtD, MW \approx 11.36 kDa) connected by a linker. (O. Palomares, Villalba, Quiralte, Polo, & Rodríguez, 2005; Torres, 2014)

Fra e 9 N-terminal domain, produced in *E. coli*, contains the active site and it is responsible for the endo- β -1,3-glucanase activity of the protein. Fra e 9 C-terminal domain, produced in *P. pastoris*, shows a significant sequence analogy to the corresponding domain of its olive tree pollen homologue, Ole e 9, and thus it has been proposed that it implicated in the regulation of the catalytic activity of the whole enzyme, since its ability to bind polysaccharides has been reported. (Torres et al., 2015)

3.3 CtD-OLE E 9 AND CtD-FRA E 9: TWO HOMOLOGUE CBMs WITH DIFFERENT LIGAND AFFINITY

The high sequence similarity between Fra e 9 C-terminal domain (CtD-Fra e 9) and other CBMs, such as the homologue domain of Ole e 9 (CtD-Ole e 9) (Figure 4.3), aroused the interest in analysing its ability for binding carbohydrates, in the same way as it was done in previous works with the aforementioned CtD-Ole e 9 and Ole e 10. (Barral et al., 2005; Rodríguez, Villalba, Batanero, Palomares, & Salamanca, 2007; Torres, 2014) Using AGE assays in polyacrylamide gels with different carbohydrates embedded in the gel matrix, the affinity of recombinant CtD-Fra e 9 (rCtD-Fra e 9) for different ligands was studied and compared with the results obtained before for recombinant CtD-Ole e 9 (rCtD-Ole e 9).

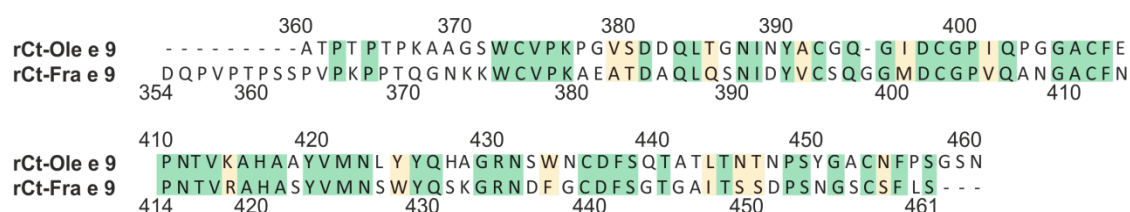


Figure 4.3. Sequence alignment for rCtD-Ole e 9 and rCtD-Fra e 9. Performed using the SIM alignment tool (<http://web.expasy.org/sim/>) (Huang & Miller, 1991)). Identical residues are highlighted in green, whereas similar residues are highlighted in yellow.

The carbohydrates selected for the study were laminarin ([1,3- β -glucan]₂₃₋₂₇), laminaritetraose ([1,3- β -glucan]₄), laminarihexose ([1,3- β -glucan]₆), agarose ([1,3/1,4- α -galactose]_n), CM-cellulose ([1,4- β -glucan]_n) and lichenan ([1,3/1,4- β -glucan]_n). Results showed that rCtD-Fra e 9 is only able to bind laminarin significantly, although a minor

interaction with lichenan was also detected. Specificity displayed by rCtD-Fra e 9 is comparable to that showed by rCtD-Ole e 9, as both are capable of interacting with 1,3- β -glucans and long-chain 1,3/1,4- β -glucans. (Torres, 2014; Zamora-Carreras et al., 2015) Nevertheless, comparing the affinity for laminarin of both domains it is obvious that rCtD-Fra e 9 possesses a binding capacity noticeably lower than that of rCtD-Ole e 9 ($K_d = 1.1 \pm 0.4$ mM vs. $K_d = 0.065 \pm 0.003$ mM, respectively).

It is remarkable the different ligand affinity for a 1,3- β -glucans showed by CtD-Ole e 9 and CtD-Fra e 9, considering that 1,3- β -glucans are presumably the natural substrates of the complete enzymes (as they are endo- β -1,3-glucanases), and their significant sequence identity (55.7%, according to the SIM alignment tool, <http://web.expasy.org/sim/> (Huang & Miller, 1991))

3.4 A MODEL β -1,3-GLUCAN: LAMINARIN

As mentioned before, laminarin is a small storage polysaccharide found in brown algae. It is a β -1,3-D-glucan containing around 25 monomers of glucose (Table 4.1), with a mainly linear geometry and a small proportion of glucosyl units attached with β (1 \rightarrow 6) bonds (Figure 4.4). Laminarin is a polydisperse polymer, showing some structural heterogeneity and it has been used as a model polysaccharide for years due to its small size and simplicity of structure. (Read et al., 1996)

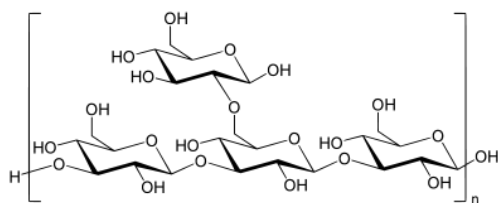


Figure 4.4. Representation of the chemical structure of the repetitive unit of laminarin.

In this work, a commercial laminarin from *Laminaria digitata*, with an average molecular mass of 5.5 kDa, has been utilised as model β -1,3-D-glucan.

OBJECTIVES

In this chapter, a deep study of the C-terminal domain of the allergenic Fra e 9 endo- β -1,3-glucanase, present in the European ash tree pollen, has been tackled. Also a comparative analysis with the C-terminal domain of its homologue Ole e 9, isolated from olive tree pollen, has been carried out. These domains are carbohydrate-binding modules (CBMs) and they represent interesting models for studying the recognition of sugars by proteins. Given the differences observed in the affinity for some polysaccharides between these two domains, which present a significant sequence analogy, the application of a technique with atomic resolution such as solution NMR is very suitable. Solution NMR and other techniques are utilized in this chapter to elucidate the molecular bases that are responsible for the recognition of carbohydrates by the protein systems analysed.

The specific objectives proposed in this chapter are:

- To determine the secondary structure elements and the global folding characteristics of rCtD-Fra e 9, through the assignment and analysis of NMR spectra.
- To study the interaction of rCtD-Fra e 9 with some different ligands by affinity gel electrophoresis (AGE).
- To investigate the interaction of the homologous rCtD-Ole e 9 and rCtD-Fra e 9 with laminarin by solution NMR and identify the carbohydrate-binding regions.
- To characterise the interaction of rCtD-Ole e 9 and rCtD-Fra e 9 with laminarin by relaxation NMR experiments, hydrodynamic behaviour studies and analytical centrifugation experiments.
- To extract information about the binding process and the interactions involved in the recognition by computational techniques, such as molecular docking.
- To propose a mechanism for the carbohydrate recognition based on the obtained results and the comparative analysis between the rCtD-Ole e 9 and rCtD-Fra e 9 results.

MATERIALS AND METHODS

1. rCtD–OLE E 9 AND rCtD–FRA E 9 PRODUCTION AND PURIFICATION

Production and purification of the recombinant C-terminal domains of Ole e 9 and Fra e 9 allergens were performed by the group of Dr. Ó. Palomares, Dr. M. Villalba and Dr. R. Rodríguez (Universidad Complutense de Madrid). rCtD–Ole e 9 domain comprises residues A360–N460 of the whole protein (Figure 4.3), whereas rCtD–Fra e 9 domain is composed of D354–S461 residues of Fra e 9 (Figure 4.3). In both cases, the numbering is referred to the protein sequence before the post-translational processing, this is including the peptide signal. The production of rCtD–Ole e 9 and rCtD–Fra e 9 was carried out using KM71 strain of *Pichia pastoris*. (Oscar Palomares, Villalba, & Rodríguez, 2003; Miguel Á Treviño et al., 2008)

During production process, the cells were cultured in 1 L of BMG (100 mM K_2HPO_4 pH 6, 0.34 % yeast nitrogen base, 1 % $(NH_4)_2SO_4$, $4 \cdot 10^{-5}$ % biotin and 1 % glycerol) for 72 hours at 30 °C. Then, the cells were transferred to 200 mL of BMM induction medium (100 mM K_2HPO_4 pH 6, 0.34 % yeast nitrogen base, 1% $(NH_4)_2SO_4$, $4 \cdot 10^{-5}$ % biotin and 0.5 % methanol). After 4 days of culture, supernatant was dialyzed in the presence of 20 mM NH_4HCO_3 . For protein purification, two chromatographic separations were utilized: one in a Sephadex G–50 column in 0.2 M NH_4HCO_3 , and the other in a ionic exchange DEAE–cellulose column, using a 20 mM to 0.5 M gradient of NH_4HCO_3 . Purity was determined by 15 % SDS–PAGE.

A procedure described previously (Miguel Ángel Treviño et al., 2004) with slight modifications was employed to produce ^{15}N – ^{13}C labelled proteins. Namely $(NH_4)_2SO_4$ was substituted by $(^{15}NH_4)_2SO_4$ (Cambridge Isotopes) in the BMG and BMM media, and 0.5 % ^{13}C –glucose (Cambridge Isotopes) was used instead of glycerol in BMG and methanol was replaced by ^{13}C –methanol (Cambridge Isotopes) in BMM. All the samples were analysed by amino acid analysis, N-terminal sequencing and mass spectroscopy.

2. N–GLYCOSYLATION PREDICTION ANALYSIS

To identify N-glycosylation positions in rCtD–Ole e 9 and rCtD–Fra e 9, a prediction analysis was performed using NetNGlyc webserver (NetNGlyc 1.0 Server, <http://www.cbs.dtu.dk/services/NetNGlyc/>). (Gupta, Jung, & Brunak, 2004)

3. CARBOHYDRATE–BINDING ASSAYS

To test the polysaccharide-binding activity of the rCtD from Fra e 9 and Ole e 9, affinity gel electrophoresis (AGE) was applied, as described previously (Barral et al., 2005). Proteins (2 µg) were electrophoresed in native 15% polyacrylamide gels containing laminarin or lichenan ranging from 0.062 to 1.2 mM or 0.24 to 4.8 mM, respectively. Other different soluble oligosaccharides and polysaccharides (laminaritetraose, laminarihexaose, agarose and

CM-cellulose) were also soaked up in the separating gel mixture at a concentration of $2.5 \text{ mg}\cdot\text{mL}^{-1}$ prior to polymerisation. As controls, gels without ligand and with BSA ($0.7 \text{ }\mu\text{g}$, negative control) were electrophoresed simultaneously. All ligands were purchased from Sigma-Aldrich (USA). The K_d value for the binding of rCtD-Fra e 9 to ligand under the conditions described was determined as the inverse of the absolute value of the intersection of the plot with the abscissa. (Bolam et al., 2004)

4. NMR SPECTROSCOPY AND SPECTRAL ASSIGNMENT

rCtD-Ole e 9 and rCtD-Fra e 9 samples were prepared at 0.4 mM in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 v/v) containing DSS as the internal ^1H chemical shift reference at pH 6.0. NMR spectra were acquired at 298 K on a Bruker AV 800 NMR spectrometer equipped with a triple-resonance TCI cryoprobe and an active shielded z-gradient coil. ^{15}N and ^{13}C chemical shifts were referenced indirectly using the gyromagnetic ratios of $^{15}\text{N}:^1\text{H}$ and $^{13}\text{C}:^1\text{H}$. (Wishart et al., 1995) Standard 2D ^{15}N -HSQC and ^{13}C -HSQC and 3D spectra CBCA(CO)NH, CBCANH, HNCA, HN(CO)CA were acquired and analysed.

The spectra were processed with Bruker Topspin (Bruker, Germany) and spectral analysis was performed with SPARKY3. (Goddard & Kneller, 2004) Backbone assignment of rCtD-Fra e 9 was performed following conventional strategies as previously described for rCtD-Ole e 9 (see “Appendices”, Table A14). (Miguel Á Treviño et al., 2008).

5. TITRATION WITH LAMINARIN MONITORED BY NMR

Increasing amounts of a laminarin (average molecular mass about 5.5 kDa) solution from *Laminaria digitata* (20 mM and pH 6.0) were added to ^{13}C - ^{15}N rCtD-Ole e 9 and ^{13}C - ^{15}N rCtD-Fra e 9 sample solutions (0.4 mM and pH 6.0) and series of ^{15}N -HSQC spectra were recorded at each titration point at 298 K . The final titration point was set at a concentration of $\approx 1 \text{ mM}$ of laminarin because precipitation was observed at higher values. Changes of peak intensity and position were monitored, and both the chemical shift and line width changes were analysed. In all cases, the pH was measured at the final points of the titrations. For the mapping of the interaction surface, average amide ^{15}N and ^1H chemical shift perturbations ($\Delta\delta_{\text{avg}}$, CSP) were calculated according to equation 4.1 (Williamson, 2013).

$$\Delta\delta_{\text{avg}} = \sqrt{\frac{(\Delta\delta_{\text{NH}})^2 + (\Delta\delta_{\text{N}})^2}{2}} \quad (4.1)$$

6. NMR DYNAMICS

NMR relaxation experiments were carried out at the same conditions described above for the laminarin-free and bound proteins. Conventional ^{15}N heteronuclear relaxation rates T_1 , T_2 and NOE data were determined. To this end, a series of 2D heteronuclear

correlated spectra using sensitivity enhanced gradient pulse scheme (Farrow et al., 1994) were recorded. The relaxation delay times were set as follows for T_1 : 5, 50, 150, 300, 600, 800, 1000 and 1200 ms; and for T_2 : 16, 32, 48, 64, 80, 96, 112 and 128 ms. The relaxation time constants T_1 and T_2 were obtained from the exponential fits of the measured cross peak intensities. The uncertainty was taken as the error in the fit of the decay function. For the NOE measurement, the experiments with and without proton saturation were acquired simultaneously in an interleaved manner with a recycling delay of 5 s, and were split during processing into separate spectra for analysis. The values for the heteronuclear NOEs were obtained from the ratio intensities of the resonances according to: $I_{\text{sat}}/I_{\text{ref}}$. Here, the uncertainty was estimated to be about 5 %.

Correlation times (τ_c) were estimated for both rCtD–Fra e 9 and rCtD–Ole e 9 (free and the laminarin complexes) from experimental $\langle T_1/T_2 \rangle$ values using equation 4.2, derived from the equation described in the literature. (Kay, Torchia, & Bax, 1989)

$$\tau_c \approx \frac{1}{4 \cdot \pi \cdot \nu_N} \cdot \sqrt{6 \cdot \left\langle \frac{T_1}{T_2} \right\rangle - 7} \quad (4.2)$$

ν_N is the ^{15}N resonance frequency of the spectrometer, and the rest of parameters were mentioned before.

Additionally, correlation times were also estimated employing the program HydroNMR (de la Torre, Huertas, & Carrasco, 2000). According to the literature (Kay et al., 1989), the experimental $\langle T_1/T_2 \rangle$ ratios were modified by excluding those residues with T_1 , T_2 values that deviate more than one standard deviation from the mean, and with NOE values below 0.65. Calculations with HydroNMR were performed assuming a globular shape and a rigid behaviour for both the isolated domains and complexes. Theoretical correlation times were estimated by a back-calculation procedure based on an iterative method that allows us to compare the theoretical $\langle T_1/T_2 \rangle$ values with the experimental ones obtained as described above.

The molecular masses of the rCtD–Ole e 9 and rCtD–Fra e 9-laminarin complexes were estimated from the corresponding correlation time values through an interpolation based on the least-squares fit of a linear function to experimental correlation times published for different proteins. (Aramini, 2010)

7. ANALYTICAL ULTRACENTRIFUGATION

Sedimentation and velocity experiments were carried out using an Optima XL–A analytical ultracentrifuge (absorption optics) at 25 °C. Samples were prepared in the same conditions used for NMR experiments, $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 v/v) pH 6.0, in the presence of 0.91 mM of laminarin (approx. 5.5 kDa molecular mass) and using concentrations of $3.16 \cdot 10^{-5}$ M and $3.77 \cdot 10^{-5}$ M for rCtD–Ole e 9 (10,509.7 Da) and rCtD–Fra e 9 (11,364.4 Da), respectively.

Equilibrium assays were performed by centrifugation of 80 μL of each sample at 19,000, 22,400 and 33,000 rpm, checking mass conservation for each velocity. Sedimentation profiles were analysed following a single species sedimentation model as previously described. (Varea et al., 2000) The SEDNTERP program (Laue, Shah, Ridgeway, & Pelletier, 1992) was used to calculate the protein specific volume from the sequences ($0.7095 \text{ cm}^3 \cdot \text{g}^{-1}$ for rCtD–Ole e 9 and $0.7076 \text{ cm}^3 \cdot \text{g}^{-1}$ for rCtD–Fra e 9) as well as the buffer viscosity and density. For velocity measurements (45,000 rpm) 400 μL of samples were used. Differential sedimentation coefficients, $c(s)$, were calculated by least squares boundary modelling of sedimentation velocity profiles using the program SEDFIT (Schuck, 2000) and standard sedimentation coefficients ($S_{20,w}$) were calculated by SEDNTERP from experimental values.

8. rCtD–Fra e 9 MODELLING AND MOLECULAR DOCKING

A structural model of rCtD–Fra e 9 was constructed using the SWISS-MODEL server (Guex, Peitsch, & Schwede, 2009). rCtD–Ole e 9 (PDB code: 2JON) (Miguel Á Treviño et al., 2008) reported structure was utilized as a template for the modelling on the basis of the significant sequence and structural similarities. For a more accurate modelling, the unstructured N –terminal tails of rCtD–Ole e 9 (residues A360–S371) and rCtD–Fra e 9 (residues D354–K374) were not considered in the process.

Docking of laminarin and rCtD–Ole e 9 or rCtD–Fra e 9 was performed using HADDOCK. (de Vries, van Dijk, & Bonvin, 2010) Laminarin molecule was first modelled with PyMol (Schrödinger, LLC), and then PRODRG (<http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg>) was used to obtain an initial representation of the carbohydrate. rCtD–Ole e 9 and rCtD–Fra e 9 structures used for the docking process were obtained from the PDB file (2JON) and by molecular modelling, respectively, as explained before. Protein residues highly perturbed upon carbohydrate addition were selected as the active residues to drive the docking. These residues showed chemical shift perturbation (CSP) values higher than the mean plus two times the standard deviation ($2 \cdot \text{SD}$), and therefore, assumed to be directly involved in the protein–carbohydrate interaction. Residues for which CSP was lower than this threshold were excluded from the analysis, on the basis that strong interactions might result in highly cooperative CSP effects.

RESULTS AND DISCUSSION

1. N-GLYCOSYLATION PREDICTION ANALYSIS

Ole e 9 and Fra e 9 have both two classic N-glycosylation NXS/T sequons in positions 355 and 447, and 156 and 454, respectively (numbering corresponds to the full-polypeptide sequences, including signal peptides). These sequons are ³²⁹NPT³³¹ and ⁴²¹NPS⁴²³ in Ole e 9, and ¹⁵⁶NVT¹⁵⁸ and ⁴⁵⁴NGS⁴⁵⁶S in Fra e 9. N-glycosylation is not favoured when X is a Pro, so it is expected to be unlikely in one or both sites in Ole e 9.

A prediction analysis was performed using NetNGlyc webserver (NetNGlyc 1.0 Server, <http://www.cbs.dtu.dk/services/NetNGlyc/>). (Gupta et al., 2004) This prediction server analyses the sequence and for each potential N-glycosylation site (in this case, all Asn residues were considered as potential N-glycosylation sites), it estimates a score (average N-glycosylation potential from the output of nine neural networks). This value together with a jury agreement (number of networks with a positive result) gives a N-glycosylation result (Table 4.4).

Table 4.4. Prediction of N-glycosylation sites for Ole e 9 and Fra e 9.

Ole e 9				Fra e 9			
Position	Potential	Jury agreement	Results	Position	Potential	Jury agreement	Results
³³ NYGQ ³⁶	0.7579	9/9	+++	⁶⁶ NPDI ⁶⁹	0.6583	9/9	++
⁴⁰ NLPS ⁴³	0.7610	9/9	+++	⁹⁴ NMKM ⁹⁷	0.6697	9/9	++
⁴⁹ NLLK ⁵²	0.7629	9/9	+++	¹³¹ NLID ¹³⁴	0.7686	9/9	+++
⁷⁴ NTGV ⁷⁷	0.5791	9/9	++	¹³⁵ NLVG ¹³⁸	0.7522	9/9	+++
⁸⁵ NGDI ⁸⁸	0.7386	9/9	++	¹⁴⁵ NALI ¹⁴⁸	0.6532	9/9	++
⁹⁴ NPNV ⁹⁷	0.7784	9/9	+++	¹⁵⁶ NVT ¹⁵⁹	0.6620	9/9	++
⁹⁶ NVAS ¹⁰³	0.8173	9/9	+++	²²³ NPGV ²²⁶	0.7234	9/9	++
¹⁰⁵ NVMS ¹⁰⁸	0.7727	9/9	+++	²³⁷ NMFD ²⁴⁰	0.7069	9/9	++
¹¹⁴ NIIA ¹¹⁷	0.7913	9/9	+++	⁴⁵⁴ NGSC ⁴⁵⁷	0.4394	7/9	–
¹⁴² NVQN ¹⁴⁵	0.6694	9/9	++				
²⁴¹ NMFD ²⁴⁴	0.7521	9/9	+++				
²⁵⁵ NAMG ²⁵⁸	0.7214	9/9	++				
³⁵⁵ NPTT ³⁵⁵	0.4838	5/9	–				
³⁸⁷ NINY ³⁹⁰	0.6562	9/9	++				
⁴⁴⁷ NPSY ⁴⁵⁰	0.3733	8/9	–				
⁴⁵⁴ NFPS ⁴⁵⁷	0.6646	9/9	++				

*Grey shaded rows correspond to classic sequons.

These results show high N-glycosylation potential in non-canonical sequons for both proteins. In the case of Ole e 9, there are two canonical sequons in which the X position is occupied by Pro, a very unfavourable situation leading to a drastic decrease of the calculated N-glycosylation potential. In a previous experimental study, a limited proteolysis of Ole e 9 was executed using trypsin and two fragments were obtained: N-terminal (residues 1–341) and C-terminal (residues 342–434). Both fragments were separated by SDS-PAGE and transferred to nitrocellulose membranes which were dyed with ConA (a biotinylated lectin which binds to mannose residues). Results demonstrated that only the N-terminal fragment (residues 1–341) reacted with the lectin and thus only that fragment, which

does not contain the canonical sequons, is N-glycosylated. This is in agreement with the prediction, in which some potential N-glycosylation sites in that region are suggested. (Ó. Palomares, 2005; Oscar Palomares et al., 2003) It is unusual to find N-glycosylation out of the canonical sequons, however some cases have been reported. (Chi et al., 2010; Lowenthal, Davis, Formolo, Kilpatrick, & Phinney, 2016; Schulz, 2012; Valliere-Douglass et al., 2010)

In the case of Fra e 9, prediction gives several positions with a high N-glycosylation potential, including one with a canonical sequon ($^{156}\text{NVT}^{158}$). Results indicate that the other canonical sequon is not potentially N-glycosylated. A glycosylation analysis of both Fra e 9 domains expressed in *E. coli* and *P. pastoris*, respectively, was also performed in earlier works. The procedure followed was analogous to that used with Ole e 9, mentioned before. Results indicated that only the C-terminal fragment (residues 350–461) reacted with the lectin and thus only that fragment is N-glycosylated. Regarding the N-glycosylation prediction results, it is probable that glycosylation is located in position 454, where the classic Asn–X–Ser/Thr sequon is present, even though the predicted potential is not very high and the program gives a negative result. (Torres, 2014) In fact, as seen in Figure 4.5, some signals belonging to N-glycosylated residues of rCtD–Fra e 9 are observable in ^1H – ^{15}N –HSQC correlation NMR spectra.

2. NMR SPECTRA ASSIGNMENT, SECONDARY STRUCTURE AND GLOBAL FOLD OF rCtD–FRA E 9

The NMR spectra of the rCtD–Fra e 9 were assigned at pH 6, following the standard NMR heteronuclear methodology. Assignment of the ^{13}C and ^{15}N backbone chemical shifts was facilitated by comparing rCtD–Fra e 9 data with the reported assignment of the homologous rCtD–Ole e 9 (Castrillo et al., 2006), on the basis of the sequence similarity (Figure 4.3). The assignment is practically complete (Figure 4.5), with the exception of eight residues (in bold) in the N-terminus corresponding to a Pro-rich region: $^{354}\text{DQPVPTPSSPVPKP}$ 367. To evaluate the secondary structure, C_α and C_β conformational shifts ($\Delta\delta$) were calculated for every nucleus as the difference between the chemical shift observed experimentally and a reference value obtained for random coil peptides (Wishart et al., 1995). Significant positive C_α conformational shifts indicative of helical structure (Wishart, Sykes, & Richards, 1991) are found in the stretches D384 to S396, V417 to S431 and D436 to G438; whereas significant negative C_α conformational shifts characteristic of β -strand structure were found in C376 to P378 and G445 to T448. These results obtained for the rCtD–Fra e 9 were compared with those published for rCtD–Ole e 9 and reveal the high similarity in the number, type and position of the elements of secondary structure (Figure 4.6). Overall, chemical shift differences between the two proteins are small which strongly suggests that the secondary structure and the global fold of rCtD–Ole e 9 are preserved in the rCtD–Fra e 9.

247

Regarding these structural similarities, a model for rCtD–Fra e 9 was built using the SWISS MODEL server (Guex et al., 2009), in order to facilitate the comparative study between rCtD–Ole e 9 and rCtD–Fra e 9 (Figure 4.7). The structural model was constructed using the previously reported structure of rCtD–Ole e 9 (PDB code: 2JON) (Miguel Á Treviño et al., 2008) as a starting point and the unstructured N-terminal tails of the two domains (residues A360–S371 in rCtD–Ole e 9 and residues D354–K374 in rCtD–Fra e 9) were obviated during the modelling process to gain accuracy. The SWISS MODEL server provides an estimation of the quality of the model created through a parameter called GMQE (global model quality estimation). This parameter takes into account torsion angles, distance-dependent potentials and solvation potentials, and it assigns a score between 0–1 to each residue. High values are indicative of a reliable model. In the model built for rCtD–Fra e 9, the global GMQE value was 0.78.

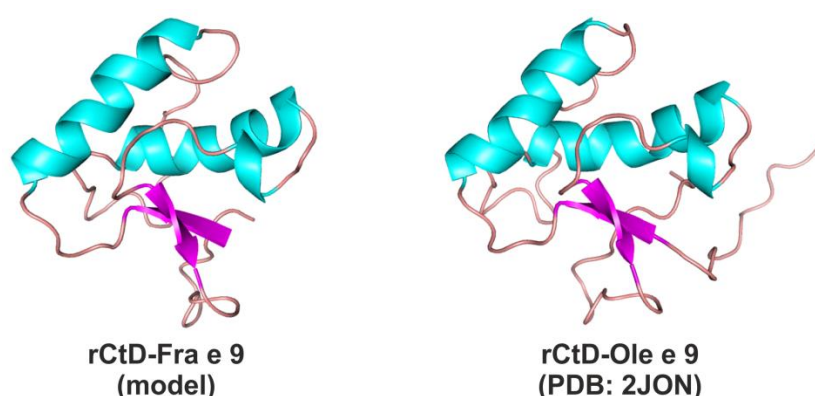


Figure 4.7. rCtD–Fra e 9 and rCtD–Ole e 9 structures. rCtD–Fra e 9 structure (**left**) is a model obtained with the SWISS-MODEL webserver, whereas rCtD–Ole e 9 (**right**) corresponds to the structure solved by NMR deposited in PDB. Secondary structure elements follow this colour code: magenta for β -strands, cyan for α -helices and salmon for loops.

3. CARBOHYDRATE BINDING ACTIVITY OF RCTD–FRA E 9

The ability of purified rCtD–Fra e 9 to bind soluble polysaccharides was assessed by quantitative AGE under non-denaturing conditions. Carbohydrates of different lengths and linkages were assayed: agarose (α 1 \rightarrow 3/ β 1 \rightarrow 4), CM-cellulose (β 1 \rightarrow 4), lichenan (β 1 \rightarrow 3 and β 1 \rightarrow 4), laminaritetraose (β 1 \rightarrow 3), laminarihexaose (β 1 \rightarrow 3) and laminarin (β 1 \rightarrow 3) (Table 4.5).

Table 4.5 Carbohydrates used in the binding assays.

Name	Backbone linkage	Branching [†]	Average DP
Agarose	α (1 \rightarrow 3)/ β (1 \rightarrow 4)	-	~800
CM-cellulose	β (1 \rightarrow 4)	-	~300–10000
Lichenan	β (1 \rightarrow 3)/ β (1 \rightarrow 4)	-	~400
Laminaritetraose	β (1 \rightarrow 3)	-	4 (exact)
Laminarihexaose	β (1 \rightarrow 3)	-	6 (exact)
Laminarin	β (1 \rightarrow 3)	β (1 \rightarrow 6) (7:1)	~25

[†]Branched:backbone bonds ratio is indicated in parentheses.

Relative to rCtD–Ole e 9, purified rCtD–Fra e 9 displayed a significant but lower specific binding to laminarin as demonstrated by the clear shift of mobility for this domain in AGE (Figure 4.8). The rCtD–Fra e 9 did not show significant binding capacity to any of the other polysaccharides assayed, except a weak mobility change in the presence of lichenan (β -1,3/1,4-glucan) (Figure 4.8).

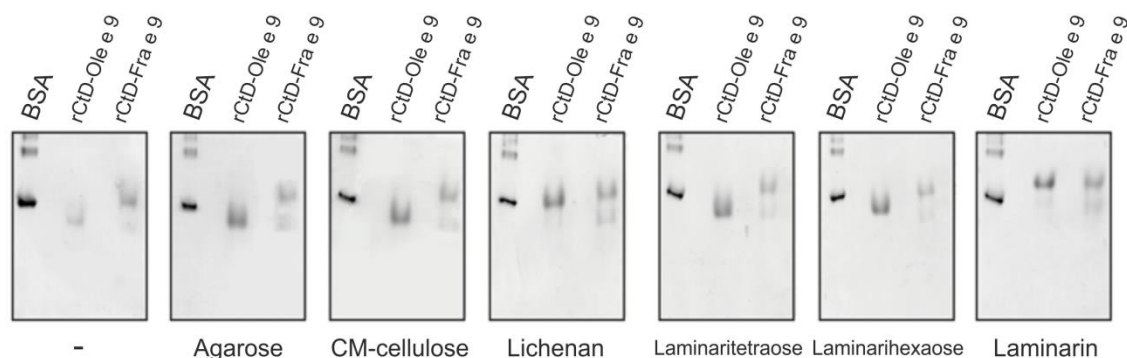


Figure 4.8. Carbohydrate binding assays with rCtD–Ole e 9 and rCtD–Fra e 9. C-terminal domains of Ole e 9 and Fra e 9 (2 μ g) and BSA (0.7 μ g) were electrophoresed in non-denaturing conditions, in polyacrylamide gels in the presence ($2.5 \text{ mg} \cdot \text{mL}^{-1}$) and the absence of polysaccharides (–).

Several assays at different concentration values of laminarin were performed to determine the dissociation constant, K_d , with rCtD–Ole e 9 and rCtD–Fra e 9 (Figure 4.9A). For this purpose, the reciprocal relative migration distance, $(1/(R-r))$, was plotted against the carbohydrate concentration, and K_d was calculated as the reciprocal of the value given by the intersection of the regression line with the abscissa axis (Figure 4.9B). rCtD–Fra e 9–laminarin complex possesses a $K_d = 1.1 \pm 0.4 \text{ mM}$, indicating a weaker interaction compared to that of the rCtD–Ole e 9–laminarin complex, whose K_d is $0.065 \pm 0.003 \text{ mM}$.

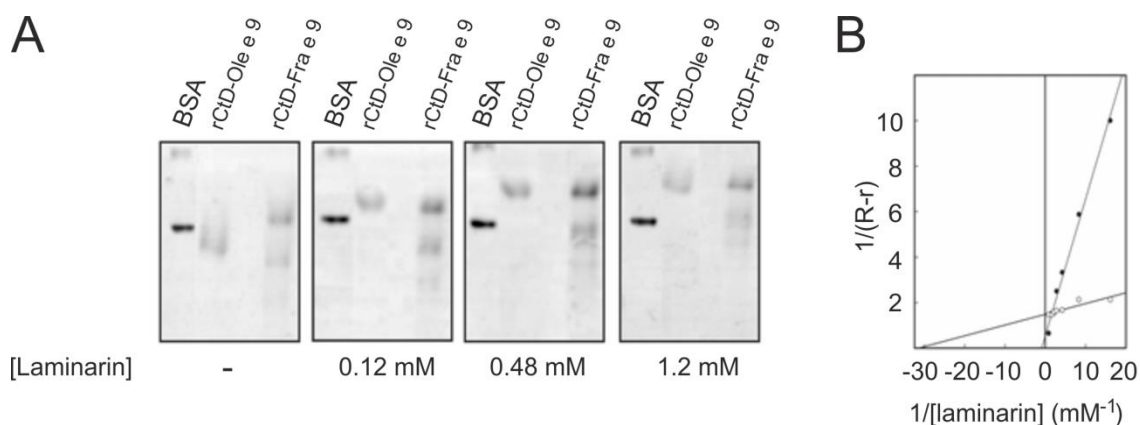


Figure 4.9. Calculation of K_d for laminarin complexes. **A.** Different concentrations of laminarin were assayed to determine the dissociation constant, K_d , of the binding to rCtD–Ole e 9 and rCtD–Fra e 9. **B.** Fitting of the reciprocal relative migration against the reciprocal ligand concentration for K_d calculation. Empty circles (○) correspond to rCtD–Fra e 9 and solid circles (●) to rCtD–Ole e 9.

4. TESTING THE INTERACTION OF rCtD–Ole e 9 AND rCtD–Fra e 9 WITH LAMINARIN BY NMR

The interaction of rCtD–Fra e 9 and rCtD–Ole e 9 with laminarin was first tested comparing the ^1H – ^{15}N –HSQC spectra recorded for free proteins and for the ($\sim 1:1$ and $1:2.5$) complexes in samples where only the protein moiety was labelled. As explained before, a higher excess of laminarin could not be used due to precipitation problems. This means that, in the case of rCtD–Fra e 9, a concentration of laminarin above the calculated K_d value was not reached ($\approx 1 \text{ mM} < 1.1 \text{ mM}$).

For rCtD–Fra e 9, upon the addition of laminarin a significant number of resonances shift, as seen in Figure 4.10, confirming the direct interaction in a fast exchange regime in the NMR timescale. To facilitate detection of the residues most affected by complex formation, the chemical shift perturbations were used (Figure 4.11), and mapped the protein interacting surface (Figure 4.12) into the protein's structural model. The rCtD–Fra e 9 residues most affected by the interaction ($\Delta\delta_{\text{avg}} > \text{mean value}$) are: K373, S389, I391–D392, V394–S396, G398–G399, V405, A407, N415, A419–A421, Y423–M425, W428–N435, G438, F441, G445 and S450.

For rCtD–Ole e 9, upon the addition of laminarin changes in the position of several signals in the spectra were observed, indicating that the region affected by the interaction includes residues W372, G386, N389–C392, I396, V420–M421, G429, D436–A441. Some signals become broader and even in some cases this broadening lead to the complete disappearing of the signals; this is the case of the signals from residues G404, H427–A428, S432, N434 and C435. This observation suggests a direct interaction between rCtD–Ole e 9 and laminarin in an intermediate–slow exchange regime in the NMR timescale (ms or slower).

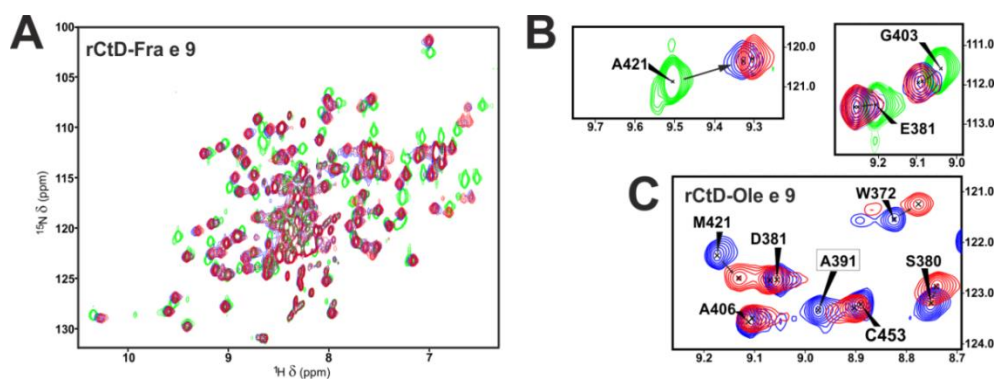


Figure 4.10. Titration of rCtD–Fra e 9 and rCtD–Ole e 9 with laminarin monitored by NMR. **A.** Superposition of the ^1H – ^{15}N –HSQC spectra of rCtD–Fra e 9 acquired during the titration with laminarin. Spectra at different relative laminarin:rCtD–Fra e 9 concentration are showed in different colours (green = 0:1; blue = 1:1; red = 2.5:1) **B.** Two examples of signals shifting upon the addition of laminarin in the rCtD–Fra e 9 sample. **C.** Superposition of a region of the ^1H – ^{15}N –HSQC spectra of rCtD–Ole e 9 at different relative laminarin:rCtD–Ole e 9 concentrations (blue = 0:1; red = 2.5:1). Signals that change their position are shown and also one which disappears (A391, framed label).

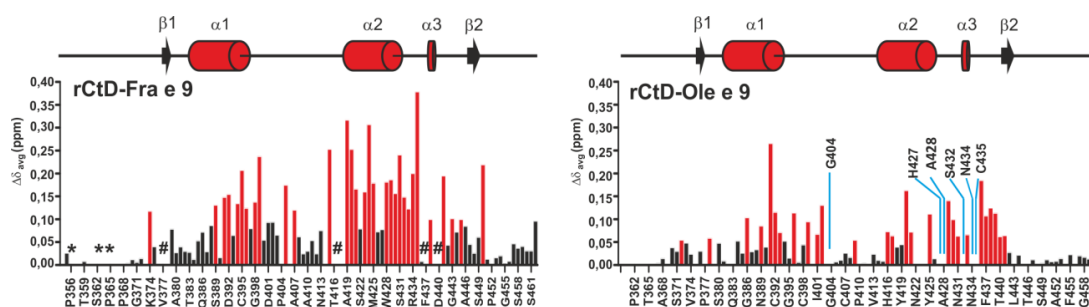


Figure 4.11. Chemical shift perturbations ($\Delta\delta_{\text{avg}}$) of rCtD–Fra e 9 and rCtD–Ole e 9 observed upon the addition of an excess of laminarin. The most significant perturbations ($>$ mean value) are highlighted in red and signals disappearing are indicated with a blue line. Non-assigned residues are marked with an asterisk (*) and those whose signals are highly overlapped are denoted by a hashtag (#).

In both domains, the mapping of the interaction region shows that the main perturbations are concentrated in one side of the molecule, around the helices α_1 , α_2 and α_3 . In addition, for rCtD–Fra e 9, the loop following helix α_3 and strand β_2 is also affected (Figure 4.11).

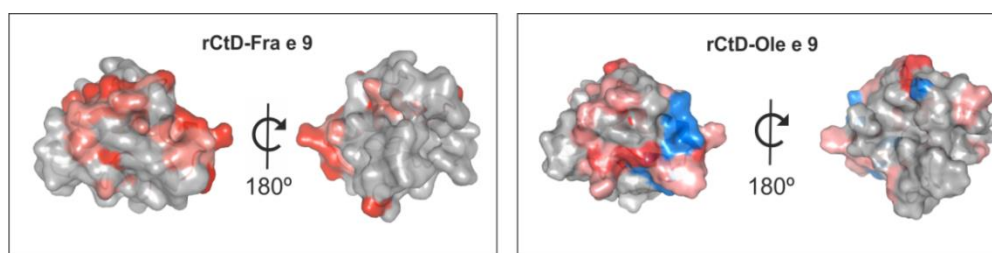


Figure 4.12. Mapping of the chemical shift perturbations on rCtD–Fra e 9 and rCtD–Ole e 9 surfaces. In the case of rCtD–Ole e 9, the N-terminal segment (residues A360–S371) is not displayed for a better comparison. Relative $\Delta\delta_{\text{avg}}$ values are represented with red colour gradient. Residues whose signals disappear upon the addition of laminarin are coloured in blue.

The interaction between the studied domains and laminarin was also tested using the NMR relaxation data obtained for both rCtD–Fra e 9 and rCtD–Ole e 9 (Table 4.6). As expected, globally T_1 values increases and T_2 values decreases as a consequence of molecular association. For rCtD–Fra e 9, mean values are $\langle T_1 \rangle = 0.64$ s and $\langle T_2 \rangle = 0.097$ s for the free domain, and $\langle T_1 \rangle = 0.96$ s and $\langle T_2 \rangle = 0.051$ s in the presence of laminarin. For rCtD–Ole e 9, mean values are $\langle T_1 \rangle = 0.83$ s and $\langle T_2 \rangle = 0.091$ s for the free domain, and $\langle T_1 \rangle = 1.26$ s and $\langle T_2 \rangle = 0.061$ s in the presence of laminarin.

5. HYDRODYNAMIC DATA OF FREE RCTD-OLE E 9 AND RCTD-FRA E 9 AND THEIR COMPLEXES WITH LAMINARIN

It is very well known that $\langle T_1/T_2 \rangle$ can be used to estimate the molecular correlation time (τ_c) in globular systems with isotropic tumbling, by simply applying equation 4.2. However, contributions of flexible tails (low NOE values) or regions affected by exchange

processes can introduce distortions in the estimations and it is good practice to exclude these residues from the calculation process. In these cases, a good approach to obtain τ_c is the use of theoretical hydrodynamic calculations (Pérez-Cañadillas et al., 2002), as described in the “Materials and methods” section. The calculated correlation times applying the two mentioned methods are in Table 4.6. Theoretical molecular masses of the complexes obtained by interpolation from known empiric values are 17.9 kDa (τ_c from hydrodynamic calculations) and 12.2 kDa (τ_c from equation) for rCtD–Ole e 9; and 16.9 kDa (τ_c from hydrodynamic calculations) and 11.9 kDa (τ_c from equation) for rCtD–Fra e 9. Calculated masses using the correlation times derived from the equation give bad results (stoichiometry under 1:1 in rCtD–Ole e 9, or even a molecular mass lower than the free protein in rCtD–Fra e 9). Therefore, τ_c calculated by the hydrodynamic approach are more reliable, and those values agree, within the error of this approach, with a 1:1 complex.

Table 4.6 NMR relaxation data for rCtD–Ole e 9 and rCtD–Fra e 9.

Sample		$\langle T_1 \rangle$ (s)	$\langle T_2 \rangle$ (s)	τ_c (ns) Equation 4.2	τ_c (ns) Hydrodynamic	Molecular masses (kDa)
rCtD–Ole e 9	Free	0.83	0.091	5.6	6.5	10.5 ^a
	Complex	1.26	0.061	7.5	10.9	17.9 ^b
rCtD–Fra e 9	Free	0.64	0.097	5.1	5.6	11.4 ^a
	Complex	0.96	0.051	7.3	10.3	16.9 ^b

^aTheoretical molecular masses. ^bMolecular masses calculated from hydrodynamic data.

Hydrodynamic behaviour of rCtD–Ole e 9 and rCtD–Fra e 9 in the presence of laminarin was also studied by analytical ultracentrifugation to determine the homogeneity of the protein solutions, the association state and the stoichiometry of the complex protein–oligosaccharide.

For both complexes, the sedimentation velocity profiles showed an apparent single boundary that could be described in terms of a single deposited species (Figure 4.13A), with experimental sedimentation coefficients of 2.15 S and 2.34 S for rCtD–Ole e 9 and rCtD–Fra e 9 (the corresponding standard values, $S_{20,w}$, were 2.24 S and 2.07 S, respectively). Therefore, the absence of any other deposited species indicates that rCtD–Fra e 9 and rCtD–Ole e 9 behave as homogeneous systems in the presence of laminarin at the experimental conditions (see Material and methods section).

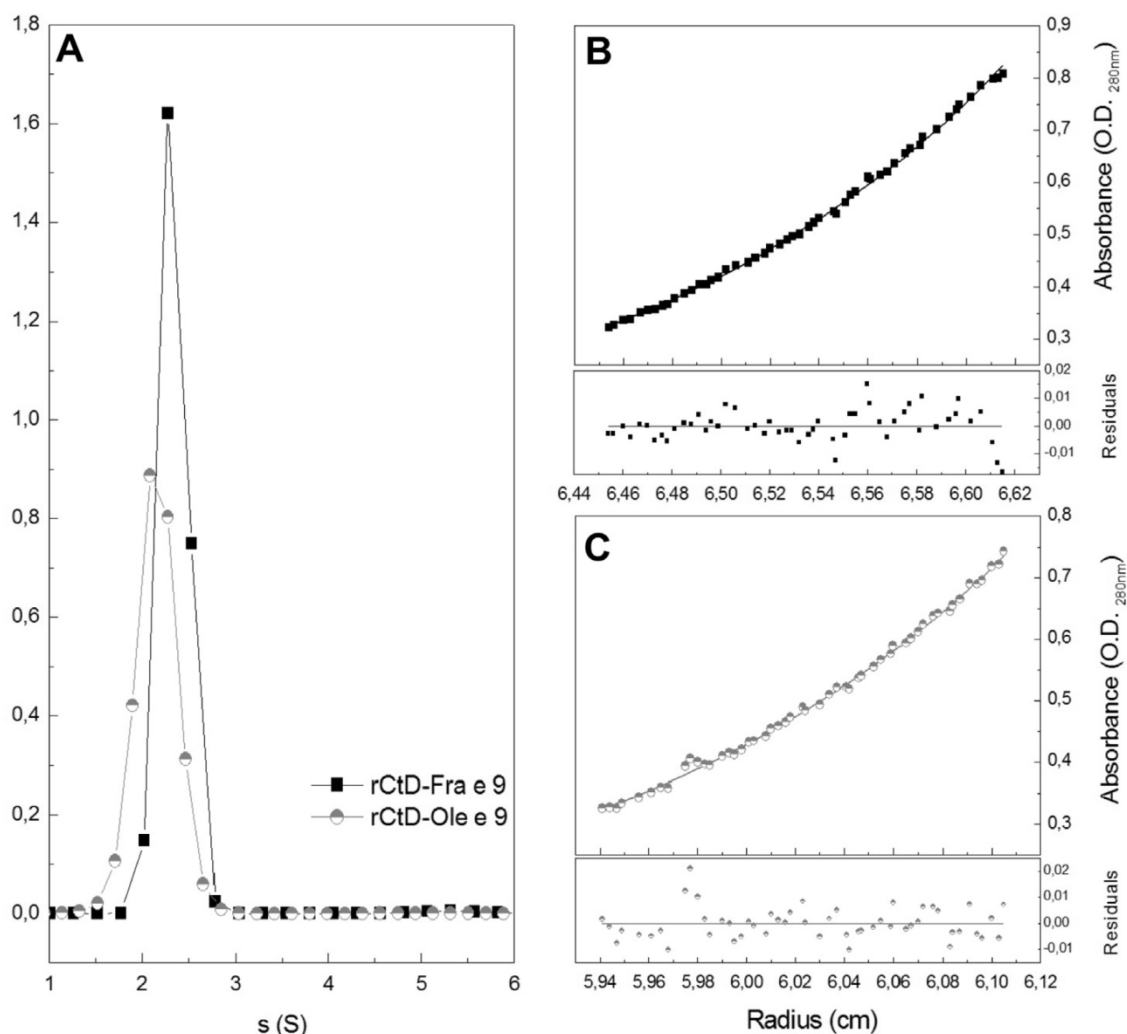


Figure 4.13. Velocity and equilibrium sedimentation experiments. **A.** Distribution of sedimentation coefficients for rCtD-Fra e 9 (■) and rCtD-Ole e 9 (●) measured at 45,000 rpm, 25 °C. **B** and **C.** Sedimentation equilibrium profile of rCtD-Fra e 9 (■) and rCtD-Ole e 9 (●) at 19,000 rpm and 25 °C. The continuous line corresponds to the fit of the experimental data to a single species model. Corresponding residual plots are below.

On the other hand, the association processes can be studied by equilibrium sedimentation experiments when the mass of the complex significantly differs from the mass of the isolated components. This analysis yields the buoyant molecular weight of the deposited species (M_b). To convert these values to absolute molecular masses (the average molecular mass, $M_{w,app}$) the partial specific volume (v) of the particle is required ($M_b = M_{w,app} (1 - \rho v)$, where ρ is the density of the solution). The partial specific volume of a protein can be easily calculated from its residue composition (Perkins, 1986), but this is not the case for the laminarin or for the protein-carbohydrate complex. As it has been described for glycoproteins (Ghirlando et al., 1995; Lewis & Junghans, 2000), in these situations the stoichiometry of the complex can be readily obtained from the buoyant molecular masses instead of the $M_{w,app}$. Therefore, the theoretical buoyant masses of the rCtD-Fra e 9 and rCtD-Ole e 9 proteins were calculated and compared with the experimental values (Table 4.7).

The experimental equilibrium profiles were fitted to an ideal model (single species model) yielding very similar results at different rotor speeds (19,000, 22,400 and 33,000 rpm). The distribution of the residuals and the mass conservation for each velocity confirmed the homogeneity of the samples (Figure 4.13B and C), in agreement with the velocity results. The average value of buoyant mass from all experiments performed was of $5901 \pm 67 \text{ g} \cdot \text{mol}^{-1}$ for rCtD–Fra e 9 and $5835 \pm 154 \text{ g} \cdot \text{mol}^{-1}$ for rCtD–Ole e 9 laminarin mixtures, respectively (Table 4.7). These values are higher than those calculated for the free proteins alone and indicate that correspond to a protein–carbohydrate complex.

Table 4.7 Characterization of the protein–laminarin complexes by sedimentation equilibrium.

Sample	Molecular mass* ($\text{g} \cdot \text{mol}^{-1}$)	v (20 °C) ($\text{cm}^3 \cdot \text{g}^{-1}$)	M_b^1 ($\text{g} \cdot \text{mol}^{-1}$)	M_b^2 ($\text{g} \cdot \text{mol}^{-1}$)	$M_b^1 - M_b^2$ ($\text{g} \cdot \text{mol}^{-1}$)
rCtD–Ole e 9	10,509.5	0.7074	2,960.6	$5,835 \pm 154$	2,875.1
rCtD–Fra e 9	11,364.4	0.7055	3,223.2	$5,901 \pm 67$	2,677.8

*Theoretical molecular mass

¹Theoretical buoyant mass for rCtD–Ole e 9 and rCtD–Fra e 9; solution density = $1.0124 \text{ g} \cdot \text{cm}^{-3}$.

²Average of the experimental buoyant masses of the proteins in the presence of laminarin 0.91 mM. Error is expressed as the mean standard deviation, based on measurements at three different rotor speeds (19,000, 22,400, and 33,000 rpm, at 25 °C).

The molecular mass of glycoproteins has been proposed to be the sum of the molecular mass of the protein and the molecular mass of the carbohydrate portion ($M(1 - \rho v) = M_p(1 - \rho v) + M_c(1 - \rho v)$ where p and c denote to the protein and the carbohydrate, respectively (Ghirlando et al., 1995; Lewis & Junghans, 2000). Here this was assumed as valid for the formation of the protein–laminarin complex. Considering this, when the theoretical buoyant mass of the protein (M_b^1 in Table 4.7) is subtracted from the experimental buoyant mass (i.e., the buoyant mass of the complex; M_b^2 in Table 4.7) results a resting buoyant masses of $2677.8 \text{ g} \cdot \text{mol}^{-1}$ and $2875.1 \text{ g} \cdot \text{mol}^{-1}$. These values should correspond to the molecular mass of the laminarin moiety in each complex. Taking into account that in the literature, the partial specific volume for carbohydrates is estimated to be in the range of $0.60\text{--}0.64 \text{ cm}^3 \text{ g}^{-1}$ (Durchschlag, 1986), we calculated a theoretical partial specific volume of $\sim 1936\text{--}2159 \text{ cm}^3 \text{ g}^{-1}$ for laminarin (Durchschlag, 1986). Comparing this result with those obtained from the subtraction of $M_b^1 - M_b^2$, and despite the significant experimental uncertainties, the results seem compatible with a 1:1 stoichiometry for the protein–laminarin complex, which is consistent with the NMR results.

6. MOLECULAR DOCKING FOR RCTD–OLE E 9 AND RCTD–FRA E 9 WITH LAMINARIN

In order to reach a better understanding of the molecular basis behind the interaction of C-terminal domains studied with laminarin, a molecular docking was executed. Results show that, in these complexes, the laminarin molecule is located between helices $\alpha 1$ and $\alpha 2$, in a region covered with aromatic residues (Figure 4.14A). Helix $\alpha 3$ and the loop between $\alpha 3$ and $\beta 2$ are also directly involved in the interaction. In rCtD–Ole e 9, these residues are Y390,

H416, Y424, H427, W433, and F437. In rCtD–Fra e 9, the aromatic residues covering that region are those occupying the same positions in the sequence when compared to rCtD–Ole e 9, this is: Y393, H420, W428, F437, and F441. The main observed difference is the presence/absence of one histidine residue in the binding side. The position of H427 in rCtD–Ole e 9 is occupied by a serine in rCtD–Fra e 9 (S431) (Figure 4.14B). Also, in the right panel of Figure 4.14A, the absence in rCtD–Fra e 9 of the additional His present in rCtD–Ole e 9 can be observed.

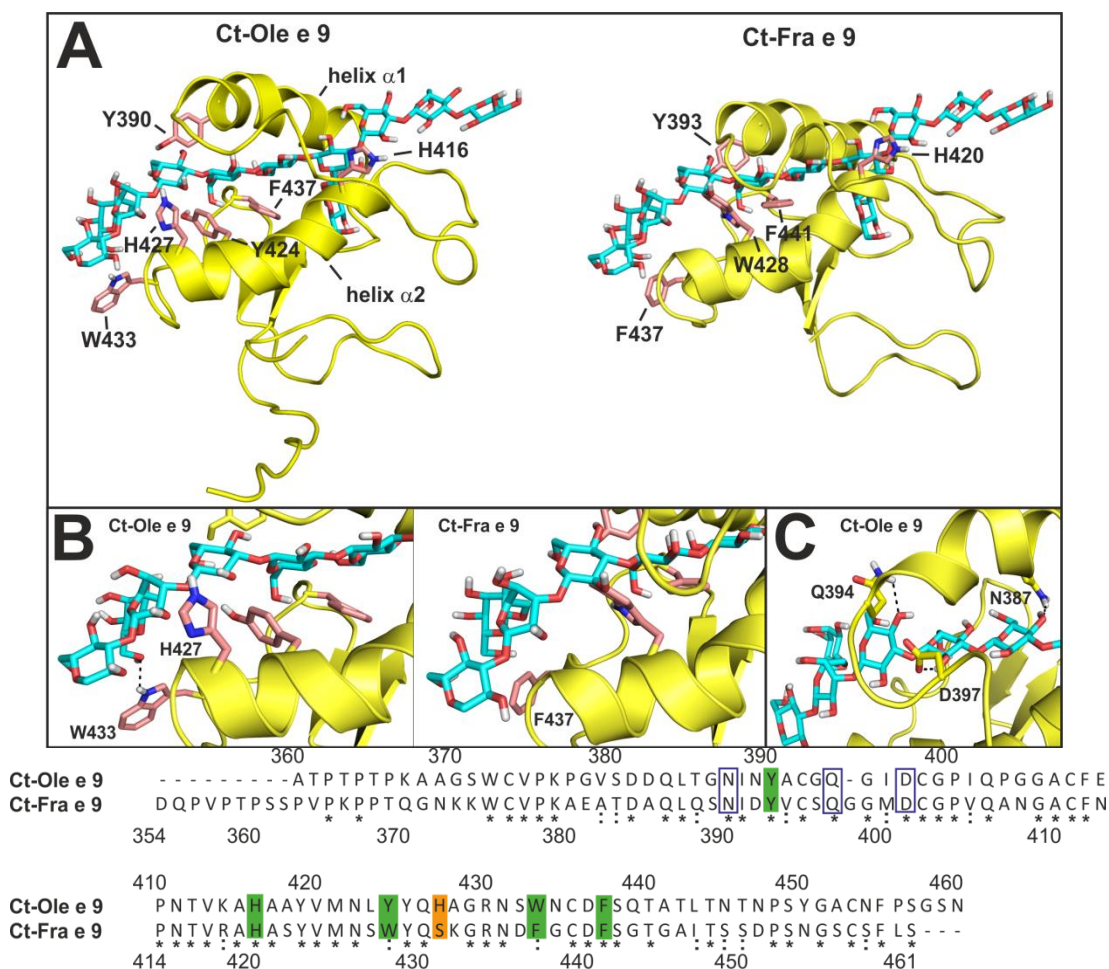


Figure 4.14. Driven docking models for rCtD–Ole e 9 and rCtD–Fra e 9 with laminarin. **A.** Comparison of the driven docking models obtained for rCtD–Ole e 9 (**left**) and rCtD–Fra e 9 (**right**). Proteins are represented in yellow cartoon, with the side chains of the aromatic residues (labelled) involved in the interaction drawn in stick representation and coloured in salmon. Laminarin molecules are drawn in blue sticks and only the fragment interacting with the modules is shown. **B.** Close-ups of the region corresponding to the C-terminus of helix α 2 for both protein domains. **C.** Close-up of the hydrogen bonds established by rCtD–Ole e 9 with laminarin. At **bottom**, sequence alignment for rCtD–Ole e 9 and rCtD–Fra e 9, performed using the SIM alignment tool (<http://web.expasy.org/sim/>) (Huang & Miller, 1991). Aromatic residues located in the region of interaction with laminarin are highlighted in green. The lack of the His residue in rCtD–Fra e 9 is coloured in orange. Non-aromatic residues involved in hydrogen formation with the ligand are boxed in dark blue. Identical residues are marked with an asterisk (*), whereas similar residues are indicated with a colon (:).

In addition, the rCtD–Ole e 9 model show the formation of a hydrogen bond between W433 and laminarin not present in the rCtD–Fra e 9 complex, since a Phe (F437) is

occupying the homologous position (Figure 4.14B). Interestingly, these models reveal additional atomic details that are the key for the complex stabilisation. Thus, some hydrogen bonds are established between rCtD–Ole e 9 and laminarin through the side chains of N386, Q394, and D397. These interactions are also present in rCtD–Fra e 9, through the corresponding residues (N390, Q397, and D401) (Figure 4.14C). Analysing thoroughly the docking structures, the differences at the atomic level in the interaction between rCtD–Ole e 9 and rCtD–Fra e 9 with the ligand can be summarised in the next table (Table 4.8).

Table 4.8. Summary of the interactions involved in the laminarin–CBM complexes analysed.

	rCtD–Ole e 9	rCtD–Fra e 9
Aromatic interacting residues		
Total number	6	5
	1 Trp, 2 Tyr, 1 Phe, 2 His	1 Trp, 1 Tyr, 2 Phe, 1 His
Hydrogen bonds		
Total number	4	3
Residues involved	N387, Q394, D397, W433	N309, Q397, D401
Affinity for laminarin		
K_d (mM)	0.065 ± 0.003	1.1 ± 0.4
Exchange regime ^a	Intermediate–slow	Fast

^aIn the NMR time scale.

7. DISCUSSION

CBMs have been widely investigated not only because of their biochemical and biological importance, but also for their potential applications in other fields. (Shoseyov et al., 2006) As mentioned before, CBMs are usually found forming part of large proteins and their function is to enhance the activity of the linked catalytic module by putting the substrate in close contact with the enzyme, or *vice versa*. (Alisdair B. Boraston, David N. Bolam, Harry J. Gilbert, & Gideon J. Davies, 2004)

Molecular interactions defining the ligand specificity of CBMs and the substrate recognition and selection mechanisms are not very well understood at the molecular level yet, despite recent advances. In this context, two similar CBMs, the C-terminal domains of two allergens with β –1,3–glucanase activity, rCtD–Ole e 9 and rCtD–Fra e 9 have been studied comparatively. It is well known that NMR spectroscopy can provide unique information about the binding of weak protein–carbohydrate complexes in solution (μ –mM range) (del Carmen Fernández-Alonso et al., 2012), such as the protein interaction surface and the groups involved in recognition. (García-Mayoral et al., 2012) NMR information, together with hydrodynamic data, have been useful to propose a model accounting for laminarin binding, aiding in the understanding of the enzymatic mechanism of large β –1,3–glucanases.

In agreement with their sequence similarity, NMR results presented here have shown that free rCtD–Fra e 9 has the same elements of secondary structure and an analogous fold

as that previously reported for free rCtD–Ole e 9. (Miguel Á Treviño et al., 2008) Dynamic properties, from relaxation and ultracentrifugation experiments, have shown that both Ct–domains behave as monomers, and that the calculated correlation times (6.5 ns and 5.6 ns for Ole e 9 and Fra e 9 domains, respectively) are in agreement with other proteins of similar size and shape. (de la Torre et al., 2000)

It has been probed that the structure of these proteins is maintained in the complexes with laminarin. NMR data confirm the presence of well-folded domains upon binding and no dramatic chemical shift changes are concomitant with complex formation. This is in agreement with known examples of transient complexes with K_d values of the same order (mM) as the ones determined here. (García-Mayoral et al., 2012) rCtD–Fra e 9 forms a lower affinity complex with laminarin than rCtD–Ole e 9 does. In fact, binding assays clearly shows that rCtD–Ole e 9 has higher affinity ($K_d = 0.032$ mM) for laminarin than rCtD–Fra e 9 ($K_d = 1.1$ mM). In this regard, the affinity of rCtD–Fra e 9 is closer to the one shown by Ole e 10 ($K_d = 4.9$ mM) an independent CBM of the same family (CBM43) (Barral et al., 2005), than to the homologous domain in olive tree pollen rCtD–Ole e 9.

These findings prompt the question of what are the structural bases for the different carbohydrate recognition properties of the homologous and phylogenetically related rCtD–Ole e 9 and rCtD–Fra e 9. To address this question, molecular docking has been performed and results have been examined thoroughly. As observed in Figure 4.14 (top), the spatial distribution of the molecules during the binding process is the same for both rCtD–Ole e 9 and rCtD–Fra e 9, as expected for two homologous CBMs. It is remarkable that almost every aromatic residue present in these domains is located in the interaction site. This agrees with the known fact that CBMs binding oligosaccharides make use of aromatic residues to establish specific interactions with sugar rings, as explained in the “Introduction” section. As rCtD–Ole e 9 and rCtD–Fra e 9 show an equivalent fold, with the same number and distribution of secondary structure elements (Figure 4.7), it is reasonable to think that the observed differences between rCtD–Ole e 9 and rCtD–Fra e 9 in the affinity for laminarin may arise from variations in the residues decorating the interaction site.

Binding sites of the CBMs are generally composed of aromatic residues and the orientation of these residues is a key determinant of specificity. (Alisdair B. Boraston et al., 2004) It has been previously mentioned that positions with aromatic residues are highly conserved between rCtD–Ole e 9 and rCtD–Fra e 9, except for H427, which is only present in rCtD–Ole e 9. The presence of an additional aromatic ring in the most perturbed region of the interacting site of rCtD–Ole e 9 could be one of the factors contributing to its higher affinity for laminarin compared to rCtD–Fra e 9. Furthermore, some variations in the type of aromatic residues have been detected between both protein domains. Residues Y424 and W433 from rCtD–Ole e 9 change to W428 and F437 in rCtD–Fra e 9. It is well-known that the nature of the aromatic ring affect the binding affinity in this kind of systems, since CH– π interactions depend on the electronic distribution of the molecule. It has been reported that the preferred order for CH– π interactions is Trp > Tyr > Phe > His. According to this, the presence of three amino acids with high preference for establishing CH– π interactions (one Trp and two Tyr) in rCtD–Ole e 9, instead of the two present (one Trp and one Tyr) in

rCtD–Fra e 9, may contribute to a more efficient interaction with the ligand. (Asensio, Ardá, Cañada, & Jiménez-Barbero, 2012; Hudson et al., 2015)

Hydrogen bonding also plays a significant role in the binding process in protein–sugar complexes. Examining the models obtained by molecular docking, some residues capable of forming hydrogen bonds with the ligand were identified. In particular, three conserved amino acids seem to interact with the laminarin molecule: an Asn, a Gln, and an Asp (N387, Q394, and D397 in rCtD–Ole e 9; N390, Q397, and D401 in rCtD–Fra e 9). Additionally, in rCtD–Ole e 9, W433 could also establish a hydrogen bond interaction with the laminarin molecule through its H_{e1} proton. This extra hydrogen bond may also contribute to the increased affinity for laminarin observed in rCtD–Ole e 9. Furthermore, the mentioned presence of the H427 residue in that region of the interacting site, together with the residue Y424, provides an “aromatic platform” that probably favours a firmer ligand attachment. In this context, the hydrogen bond formed by the close W433 residue can assist the accommodation of the ligand in this region of the CBM.

It is necessary to remark that high sequence variability has been observed in this kind of allergenic proteins. For the same species, and even for the same individual, different variants of the same protein can be found, characterised by slight dissimilarities in the primary structure. In the case of the full-length Ole e 9, several variants have been isolated and, precisely the place of H427 is occupied by Ser in some described variants. (Duffort et al., 2006) This is a very interesting question, as it means that the organism is able to express diverse versions of the protein with probably variable affinity towards the ligand.

To confirm the significance of the residues involved in carbohydrate recognition and binding in rCtD–Ole e 9 and rCtD–Fra e 9, a search for homologous sequences was performed using BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Ten representative sequences were selected for comparison, attending to the BLAST score (Table 4.9); and the alignment is shown in Figure 4.15.

Table 4.9. Homologous sequences selected for comparing with rCtD–Ole e 9 and rCtD–Fra e 9.*

Uniprot code	Protein	Organism	Residues
Q84V39	Major pollen allergen Ole e 10	<i>Olea europaea</i> (olive)	1–123
W8P8Q7–1	Fra e 10.01 allergen	<i>Fraxinus excelsior</i> (ash tree)	1–123
B5M9E5	β–glucosidase 08	<i>Solanum lycopersicum</i> (tomato)	358–459
Q53MA8	β–1,3–endoglucanase	<i>Oryza sativa ssp. Japonica</i> (rice)	334–431
K9MBJ7	β–1,3–glucanase 14	<i>Solanum tuberosum</i> (potato)	361–462
Q9MBB5	β–1,3–glucanase SgGN1	<i>Salix gilgiana</i> (Asian willow)	355–478
Q9SD84	Plasmodesmata callose-binding protein 2	<i>Arabidopsis thaliana</i>	1–106
Q9M4A9	β–1,3–glucanase gns2	<i>Pisum sativum</i> (pea)	334–453
P52409	Glucan endo–1,3–β–glucosidase GLC1	<i>Triticum aestivum</i> (wheat)	339–461
B9T3M9	Glucan endo–1,3–β–glucosidase (putative)	<i>Ricinus communis</i> (castorbean)	347–448

*Data from <http://www.uniprot.org/>

rCtD-Ole e 9	342	-----ATPT-----	PTPKAGSWGCPKPGVSDQLT		367
rCtD-Fra e 9	338	-----DQFVPFTPSSSPVK--	PPTQGNNKKWCPKAEATDAQLQ		372
Q84V39	1	-----MRGTAGVPDQPVPFTPSPSVTPSSSPVK--	PPTQGNNKKWCPKAEATDAQLQ		50
W8P8Q7-1	1	-----MRGTAGVPDQPVPFTPSPSVTPSSSPVK--	PPTQGNNKKWCPKAEATDAQLQ		50
B5M9E5	358	-----GGNGKGQPK---	TPAENKKFCMPKVEATDAQLQ		387
Q53MA8	334	-----NPSPKPA-----	PSGGGKWCPKAGASDTDLQ		360
K9MBJ7	361	-----GGNGKGQPK---	TPAENKKFCMPKVATDAQLQ		390
B9T3M9	347	-----GPTAPAPS-----	SSQGKKWCPKSDASEDALQ		374
Q9MBB5	355	---TEVFEIDLSGKT--P----	LSGYKKPLPLPTNNEPYKGLWCVVAKENRSVAVK		402
Q9SD84	1	MAPLVLYLLTLT-----	MAGHTSASVCCKTGSLDSVLQ		29
Q9M4A9	334	---SSVDAGILRNGQKPVA---	PVKGGKMPTPR----PVVGQKWCVPKADASPGALQ		383
P52409	339	---TPVYDLGLLRDGASVAPTSPNPSPNPSPKPA----	PSGGGWCKVAKDANGTDLQ		390
			:*: . : :		
rCtD-Ole e 9	368	GNINIVCSQ-GIDCGPIQPGGACFPNPNTVKFAAYVMNLITYQHAGRNSWCDFSQTATLT		426	
rCtD-Fra e 9	373	SNIDYVCSQGGMDCGPVQANGACFNPNPTVRAHASVYVMNSWYQSKGRNDFGCDFSGTGAIT		432	
Q84V39	51	SNIDYVCSQGGMDCGPIQANGACFNPNPTVRAHASVYVMNSWYQSKGRNDFGCDFSGTGAIT		110	
W8P8Q7-1	51	SNIDYVCSQGGMDCGPIQANGACFNPNPTARAHASVYVMNSWYQSKGRNDFGCFNSFGTGAIT		110	
B5M9E5	388	SNINIVCSQ-GVDCFTPIQVGGSCFKPNPTIRSHAAFAMNSYYQKEGRNNFNCDFAGTGVA		446	
Q53MA8	361	NNINIVCSQ--YVDCFKPIQSGGACFDPPNVQSHAAFMNAFYQANGRHNYDCDFKGTGAVT		418	
K9MBJ7	391	SNINIVCSQ-GVDCFTPIQVGGSCFKPNPTIRSHAAFAMNSYYQREGRRNNFNCDFAGTGVA		449	
B9T3M9	375	KNIDYVCSQ-GVDCRAIQAGGCFFDPNPTVRSHASVYMNAYYTFRQDYNCDFNHTRGLT		433	
Q9MBB5	403	DALAWACSQGNKTCDEIQPGKECYKPVSLFRHASYAFSSYWAEFKIGGVCSFNGLATT		462	
Q9SD84	30	KTLDIYACGN-GADCNPPTHKGSCEFNPDNVRAHCNNAVNSFFQKKQASESCNFGTATLT		93	
Q9M4A9	384	ANINIVCSQ-GIDCRPIQPGGVCIYAANNVKAIAIATYAMNAYYQANGKHNYNCDFSHSGVTT		442	
P52409	391	NNINIVCSQ--FVDCFKPIQSGGACFBPNSLOAHASYVMNAYYQANGHTDLACDFKGTGI		448	
		: . *. * : * : : *			
rCtD-Ole e 9	427	NTNPSYGACNPFPSGSN		442	
rCtD-Fra e 9	433	SSDPSNGSCSFSL---		445	
Q84V39	111	SSDPSNGSCSFSL---		123	
W8P8Q7-1	111	SSDPSNGSCSFSL---		123	
B5M9E5	447	FSDPSYGTCKFES---		459	
Q53MA8	419	SNDPSYSGCKYVS---		431	
K9MBJ7	450	ASDPSYGTCKFES---		462	
B9T3M9	434	NSDPSYEACSYPPFDG-		448	
Q9MBB5	463	FKDPSFQCCKFPSVTL		478	
Q9SD84	94	TTDPSYTGCAFPS---		106	
Q9M4A9	443	VSNPSHDNCRI----		453	
P52409	449	SSDPSYSGCKYVS---		461	
		* ** *			

Figure 4.15. Sequence alignment of rCtD–Ole e 9 and rCtD–Fra e 9 with ten homologous sequences found in the Uniprot database (see Table 4.9). Conserved Cys residues are in yellow boxes and conserved aromatic positions are in red boxes. Blue arrows indicate conserved residues that form hydrogen bonds with laminarin. Position of H427 in rCtD–Ole e 9 is indicated with a black arrow. Fully conserved residues are marked with an asterisk (*), whereas conservation of residues with strongly or weakly similar properties are denoted with a colon (:) or a period (.), respectively.

In relation with the sequence alignment shown in Figure 4.15 some conclusions can be drawn. Many conserved residues are located in the regions more affected by laminarin binding in rCtD–Ole e 9 and rCtD–Fra e 9. Aromatic amino acids involved in the interaction are conserved. Cysteines are also fully conserved. Since they play an essential structural role, this observation points to the fact that a similar fold may be preserved in all the considered proteins.

The conservation of aromatic residues and cysteines in these homologous sequences from different carbohydrate-active proteins, belonging to different species, are two significant details that support the idea of a possible similar binding mechanism for all of them. However, as observed for rCtD-Ole e 9 and rCtD-Fra e 9, slight variations at a residue level can lead to significant changes in affinity. According to this, it has been reported that CBMs may have very subtle specificity differences that result in significant functional consequences *in vivo*. (Davies & Williams, 2016) This can be the case in the present examples.

CONCLUSIONS

1. rCtD–Fra e 9 possesses the same number and type of secondary structure elements located in the same regions compared to rCtD–Ole e 9, suggesting an analogous fold.
2. rCtD–Fra e 9 showed a significant specific binding with laminarin, and the affinity is lower than that showed by rCtD–Ole e 9.
3. Interaction between rCtD–Fra e 9 and laminarin was verified by the CSP observed in the NMR spectra and by NMR relaxation measurements, and the main interaction region was identified between helices $\alpha 1$ and $\alpha 2$, the same region identified for rCtD–Ole e 9.
4. Effects on NMR spectra upon addition of laminarin to a rCtD–Fra e 9 sample are compatible with a quick exchange regime. In contrast, effects observed for a rCtD–Ole e 9 sample suggested an intermediate–slow exchange regime.
5. Molecular masses of the protein–laminarin complexes were estimated from τ_c values for rCtD–Ole e 9 and rCtD–Fra e 9, and they were compatible with a 1:1 stoichiometry. This was further confirmed by analytical ultracentrifugation experiments.
6. Molecular docking allowed the identification of a series of aromatic residues located in the interacting site in both rCtD–Ole e 9 and rCtD–Fra e 9 as responsible for establishing stacking interactions with the sugar rings of the ligand (Y390, H416, Y424, H427, W433, and F437 in rCtD–Ole e 9; Y393, H420, W428, F437, and F441 in rCtD–Fra e 9).
7. Some residues are capable of forming hydrogen bonds with the ligand in both rCtD–Ole e 9 and rCtD–Fra e 9 (N387, Q394, and D397 in rCtD–Ole e 9; N390, Q397, and D401 in rCtD–Fra e 9).
8. A mechanism has been proposed for both rCtD–Ole e 9 and rCtD–Fra e 9, in which the cooperative action of CH– π interactions established by aromatic residues, and hydrogen bonds formed by polar residues, is responsible for the binding of laminarin.
9. The different affinity towards laminarin observed between rCtD–Ole e 9 and rCtD–Fra e 9 has been explained by the presence of an additional His residue (H427) in rCtD–Ole e 9, an extra hydrogen bond formed by W433, and the slight variations in the nature of aromatic residues in some positions between both domains.
10. A superficial search for other homologous domains resulted in the discovery of a high conservation of the Cys and aromatic residues, pointing to the fact that the binding mechanism could be preserved in most members of the CBM family 43.

BIBLIOGRAPHY

- Aramini, J. (2010). NMR determined Rotational correlation time. Retrieved from http://www.nmr2.buffalo.edu/nesg.wiki/NMR_determined_Rotational_correlation_time
- Asensio, J. L., Ardá, A., Cañada, F. J., & Jiménez-Barbero, J. s. (2012). Carbohydrate–aromatic interactions. *Accounts of chemical research*, 46(4), 946-954.
- Bacic, A., Fincher, G. B., & Stone, B. A. (2009). *Chemistry, biochemistry, and biology of 1-3 beta glucans and related polysaccharides*: Academic Press.
- Barber, D., De la Torre, F., Feo, F., Florido, F., Guardia, P., Moreno, C., . . . Salcedo, G. (2008). Understanding patient sensitization profiles in complex pollen areas: a molecular epidemiological study. *Allergy*, 63(11), 1550-1558.
- Barral, P., Suárez, C., Alche, J., Rodriguez-Garcia, M., Rodriguez, R., Batanero, E., . . . Rivas, G. (2005). An olive pollen protein with allergenic activity, Ole e 10, defines a novel family of carbohydrate-binding modules and is potentially implicated in pollen germination. *Biochem. J*, 390, 77-84.
- Beattie, A., Hirst, E. L., & Percival, E. (1961). Studies on the metabolism of the Chrysophyceae. Comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from the brown algae. *Biochemical Journal*, 79(3), 531-537.
- Bielecki, S., & Galas, E. (1991). Microbial β -Glucanases Different from Cellulases. *Critical Reviews in Biotechnology*, 10(4), 275-304. doi:10.3109/07388559109038212
- Bluhm, T. L., Deslandes, Y., Marchessault, R. H., Pérez, S., & Rinaudo, M. (1982). Solid-state and solution conformation of scleroglucan. *Carbohydrate Research*, 100(1), 117-130. doi:[http://dx.doi.org/10.1016/S0008-6215\(00\)81030-7](http://dx.doi.org/10.1016/S0008-6215(00)81030-7)
- Bolam, D. N., Xie, H., Pell, G., Hogg, D., Galbraith, G., Henrissat, B., & Gilbert, H. J. (2004). X4 modules represent a new family of carbohydrate-binding modules that display novel properties. *Journal of Biological Chemistry*, 279(22), 22953-22963.
- Boraston, A., Bolam, D., Gilbert, H., & Davies, G. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J*, 382, 769-781.
- Boraston, Alisdair B., Bolam, David N., Gilbert, Harry J., & Davies, Gideon J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical Journal*, 382(Pt 3), 769-781. doi:10.1042/BJ20040892
- Carvalho, A. L., Goyal, A., Prates, J. A., Bolam, D. N., Gilbert, H. J., Pires, V. M., . . . Fontes, C. M. (2004). The family 11 carbohydrate-binding module of *Clostridium thermocellum* Lic26A-Cel5E accommodates β -1, 4-and β -1, 3–1, 4-mixed linked glucans at a single binding site. *Journal of Biological Chemistry*, 279(33), 34785-34793.
- Castrillo, I., Treviño, M. A., Palomares, O., Rico, M., Santoro, J., & Bruix, M. (2006). NMR assignment of the C-terminal domain of Ole e 9, a major allergen from the olive tree pollen. *Journal of Biomolecular NMR*, 36, 67-67.
- Ceasar, S. A., & Ignacimuthu, S. (2012). Genetic engineering of crop plants for fungal resistance: role of antifungal genes. *Biotechnology letters*, 34(6), 995-1002.
- Coutinho, P., & Henrissat, B. (1999). Recent advances in carbohydrate bioengineering.
- Chen, L., Garrett, T. P., Fincher, G. B., & Høj, P. B. (1995). A tetrad of ionizable amino acids is important for catalysis in barley β -glucanases. *Journal of Biological Chemistry*, 270(14), 8093-8101.
- Chi, Y. H., Koo, Y. D., Dai, S. Y., Ahn, J.-E., Yun, D.-J., Lee, S. Y., & Zhu-Salzman, K. (2010). N-glycosylation at non-canonical Asn-X-Cys sequence of an insect recombinant cathepsin B-like counter-defense protein. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 156(1), 40-47. doi:<http://dx.doi.org/10.1016/j.cbpb.2010.01.017>

- Davies, G. J., & Williams, S. J. (2016). Carbohydrate-active enzymes: sequences, shapes, contortions and cells. *Biochemical Society Transactions*, 44(1), 79-87.
- de la Torre, J. G., Huertas, M., & Carrasco, B. (2000). HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson*, 147(1), 138-146.
- de Vries, S. J., van Dijk, M., & Bonvin, A. M. J. J. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nat. Protocols*, 5(5), 883-897.
- del Carmen Fernández-Alonso, M., Díaz, D., Berbis, M. Á., Marcelo, F., Cañada, J., & Jiménez-Barbero, J. (2012). Protein-carbohydrate interactions studied by NMR: from molecular recognition to drug design. *Current protein & peptide science*, 13(8), 816.
- Dogra, V., & Sreenivasulu, Y. (2015). Cloning and functional characterisation of β -1, 3-glucanase gene from *Podophyllum hexandrum* — A high altitude Himalayan plant. *Gene*, 554(1), 25-31. doi:<http://dx.doi.org/10.1016/j.gene.2014.10.012>
- Duffort, O., Palomares, O., Lombardero, M., Villalba, M., Barber, D., Rodríguez, R., & Polo, F. (2006). Variability of Ole e 9 Allergen in Olive Pollen Extracts: Relevance of Minor Allergens in Immunotherapy Treatments. *International Archives of Allergy and Immunology*, 140(2), 131-138.
- Durchschlag, H. (1986). Thermodynamic Data for Biochemistry and Biotechnology, edited by HJ Hinz: Berlin: Springer-Verlag.
- Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., . . . Kay, L. E. (1994). Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, 33(19), 5984-6003.
- Fujimori, N., Enoki, S., Suzuki, A., Naznin, H. A., Shimizu, M., & Suzuki, S. (2016). Grape apoplasmic β -1,3-glucanase confers fungal disease resistance in Arabidopsis. *Scientia Horticulturae*, 200, 105-110. doi:<http://dx.doi.org/10.1016/j.scienta.2016.01.008>
- Gao, L., Wang, S., Zhang, Y., Li, X., Wang, H., & Liu, D. (2016). Identification and characterisation of a β -1, 3-glucanase gene, TcLr19Glu, involved in wheat resistance against *Puccinia tritici*. *Journal of Plant Biochemistry and Biotechnology*, 1-8. doi:10.1007/s13562-015-0344-4
- García-Mayoral, M. F., Canales, A. n., Díaz, D., López-Prados, J., Moussaoui, M., de Paz, J. L., . . . Boix, E. (2012). Insights into the glycosaminoglycan-mediated cytotoxic mechanism of eosinophil cationic protein revealed by NMR. *ACS chemical biology*, 8(1), 144-151.
- Gemmell, T. R., & Trimble, R. B. (1999). Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1426(2), 227-237.
- Ghirlando, R., Keown, M. B., Mackay, G. A., Lewis, M. S., Unkeless, J. C., & Gould, H. J. (1995). Stoichiometry and thermodynamics of the interaction between the Fc fragment of human IgG1 and its low-affinity receptor Fc. gamma. RIII. *Biochemistry*, 34(41), 13320-13327.
- Goddard, T., & Kneller, D. (2004). SPARKY 3. *University of California, San Francisco*, 14, 15.
- Guex, N., Peitsch, M. C., & Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, 30(S1), S162-S173.
- Gupta, R., Jung, E., & Brunak, S. (2004). Prediction of N-glycosylation sites in human proteins.
- Hashimoto, H. (2006). Recent structural studies of carbohydrate-binding modules. *Cellular and Molecular Life Sciences CMLS*, 63(24), 2954-2967.
- Him, J. L. K., Pelosi, L., Chanzy, H., Putaux, J.-L., & Bulone, V. (2001). Biosynthesis of (1 \rightarrow 3)- β -D-glucan (callose) by detergent extracts of a microsomal fraction from

- Arabidopsis thaliana. *European Journal of Biochemistry*, 268(17), 4628-4638. doi:10.1046/j.1432-1327.2001.02382.x
- Huang, X., & Miller, W. (1991). A time-efficient, linear-space local similarity algorithm. *Advances in Applied Mathematics*, 12(3), 337-357.
- Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L., & Woolfson, D. N. (2015). Carbohydrate-Aromatic Interactions in Proteins. *Journal of the American Chemical Society*, 137(48), 15152-15160. doi:10.1021/jacs.5b08424
- Huecas, S., Villalba, M., & Rodriguez, R. a. (2001). Ole e 9, a major olive pollen allergen is a 1, 3- β -glucanase isolation, characterisation, amino acid sequence, and tissue specificity. *Journal of Biological Chemistry*, 276(30), 27959-27966.
- Ishida, T., Fushinobu, S., Kawai, R., Kitaoka, M., Igarashi, K., & Samejima, M. (2009). Crystal Structure of Glycoside Hydrolase Family 55 β -1,3-Glucanase from the Basidiomycete Phanerochaete chrysosporium. *Journal of Biological Chemistry*, 284(15), 10100-10109. doi:10.1074/jbc.M808122200
- Jamal-Talabani, S., Boraston, A. B., Turkenburg, J. P., Tarbouriech, N., Ducros, V. M.-A., & Davies, G. J. (2004). Ab initio structure determination and functional characterisation of CBM36: a new family of calcium-dependent carbohydrate binding modules. *Structure*, 12(7), 1177-1187.
- Kay, L. E., Torchia, D. A., & Bax, A. (1989). Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, 28(23), 8972-8979.
- Kikuchi, T., Shibuya, H., & Jones, J. T. (2005). Molecular and biochemical characterisation of an endo-beta-1,3-glucanase from the pinewood nematode Bursaphelenchus xylophilus acquired by horizontal gene transfer from bacteria. *Biochemical Journal*, 389, 117-125. doi:10.1042/bj20042042
- Kirubakaran, S. I., & Sakthivel, N. (2007). Cloning and overexpression of antifungal barley chitinase gene in Escherichia coli. *Protein expression and purification*, 52(1), 159-166.
- Kister, A., Fokas, A., Papatheodorou, T., & Gelfand, I. (2006). Strict rules determine arrangements of strands in sandwich proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 103(11), 4107-4110.
- Lashbrook, C. C., Gonzalez-Bosch, C., & Bennett, A. B. (1994). Two divergent endo-beta-1,4-glucanase genes exhibit overlapping expression in ripening fruit and abscising flowers. *The Plant Cell*, 6(10), 1485-1493. doi:10.1105/tpc.6.10.1485
- Laue, T., Shah, B., Ridgeway, T., & Pelletier, S. (1992). In Analytical Ultracentrifugation in Biochemistry and Polymer Science (Harding, SE, Rowe, AJ & Horton, JC, eds). *Cambridge: Royal Society of Chemistry*.
- Leubner-Metzger, G., Frundt, C., Vogeli-Lange, R., & Meins, F. (1995). Class I [beta]-1,3-Glucanases in the Endosperm of Tobacco during Germination. *Plant Physiology*, 109(3), 751-759.
- Lewis, M. S., & Junghans, R. P. (2000). Ultracentrifugal analysis of molecular mass of glycoproteins of unknown or ill-defined carbohydrate composition. *Methods in enzymology*, 321, 136-149.
- Liu, D., He, X., Li, W., Chen, C., & Ge, F. (2012). A β -1,3-glucanase gene expressed in fruit of Pyrus pyrifolia enhances resistance to several pathogenic fungi in transgenic tobacco. *European Journal of Plant Pathology*, 135(2), 265-277. doi:10.1007/s10658-012-0083-5
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*, 42(Database issue), D490-D495. doi:10.1093/nar/gkt1178

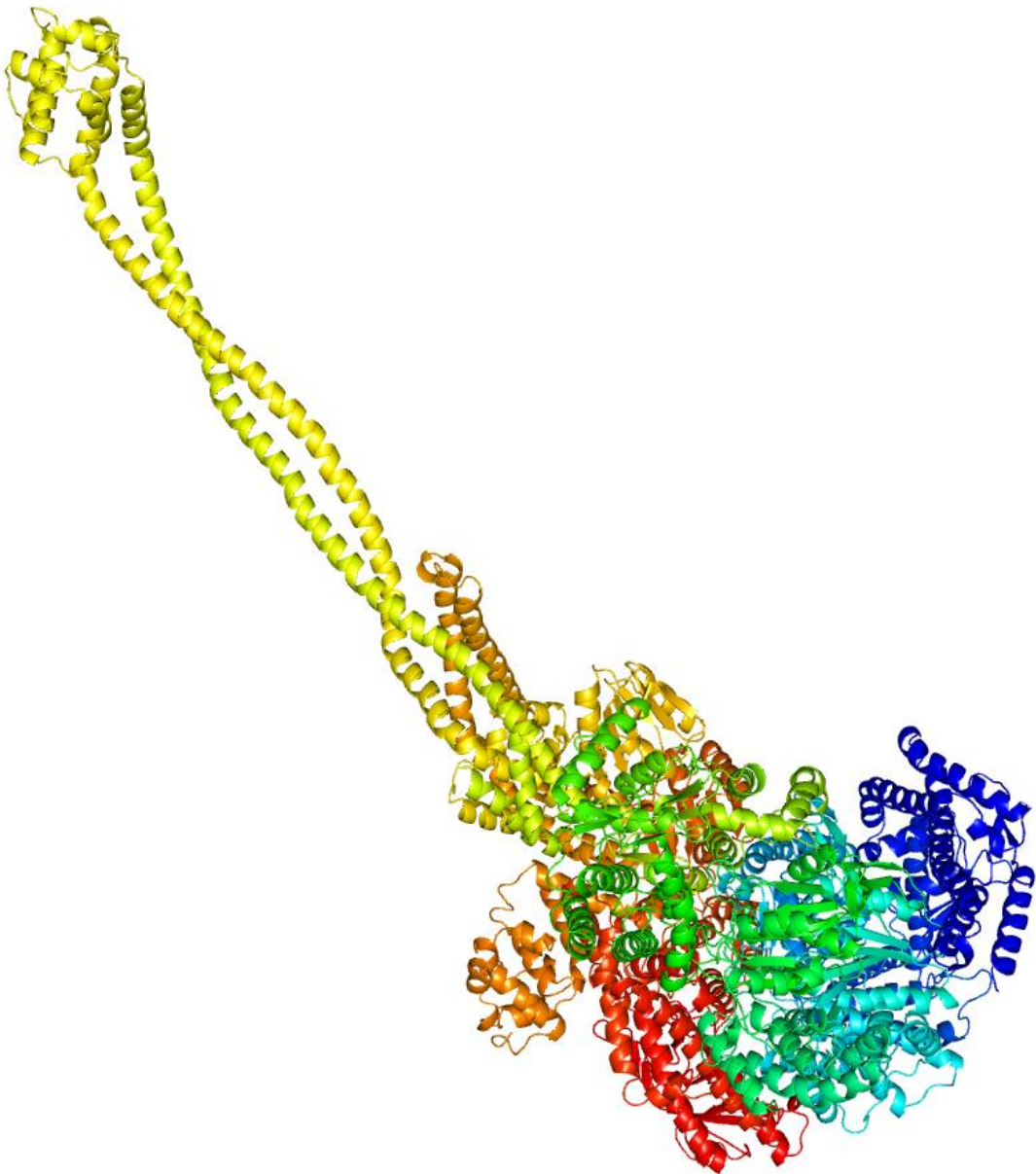
- Lowenthal, M. S., Davis, K. S., Formolo, T., Kilpatrick, L. E., & Phinney, K. W. (2016). Identification of novel N-glycosylation sites at non-canonical protein consensus motifs. *Journal of Proteome Research*. doi:10.1021/acs.jproteome.5b00733
- Martín-Cuadrado, A.-B., Fontaine, T., Esteban, P.-F., Dedo, J. E. d., Medina-Redondo, M. d., Rey, F. d., . . . Aldana, C. R. V. d. (2008). Characterisation of the endo- β -1,3-glucanase activity of *S. cerevisiae* Eng2 and other members of the GH81 family. *Fungal Genetics and Biology*, 45(4), 542-553. doi:<http://dx.doi.org/10.1016/j.fgb.2007.09.001>
- Martin, K., McDougall, B. M., McIlroy, S., Jayus, Chen, J., & Seviour, R. J. (2007). Biochemistry and molecular biology of exocellular fungal β -(1,3)- and β -(1,6)-glucanases. *FEMS Microbiology Reviews*, 31(2), 168-192. doi:10.1111/j.1574-6976.2006.00055.x
- Mohd Jamil, N. A., Rahmad, N., Mohd Nor Rashid, N., Mohd Yusoff, M. H. Y., Shaharuddin, N. S., & Mohd Saleh, N. (2013). LCMS-QTOF Determination of Lentinan-Like β -D-Glucan Content Isolated by Hot Water and Alkaline Solution from Tiger's Milk Mushroom, Termite Mushroom, and Selected Local Market Mushrooms. *Journal of Mycology*, 2013, 8. doi:10.1155/2013/718963
- Mouyna, I., Hartl, L., & Latgé, J.-P. (2013). β -1,3-glucan modifying enzymes in *Aspergillus fumigatus*. *Frontiers in Microbiology*, 4, 81. doi:10.3389/fmicb.2013.00081
- Palomares, O. (2005). *Proteínas de reserva de mostaza amarilla y 1,3-beta-glucanasa de polen de olivo como modelos de estudio de proteínas alergénicas*. Universidad de Complutense de Madrid.
- Palomares, O., Villalba, M., Quiralte, J., Polo, F., & Rodríguez, R. (2005). 1,3- β -glucanases as candidates in latex-pollen-vegetable food cross-reactivity. *Clinical & Experimental Allergy*, 35(3), 345-351. doi:10.1111/j.1365-2222.2004.02186.x
- Palomares, O., Villalba, M., & Rodríguez, R. (2003). The C-terminal segment of the 1, 3-beta-glucanase Ole e 9 from olive (*Olea europaea*) pollen is an independent domain with allergenic activity: expression in *Pichia pastoris* and characterisation. *Biochem. J*, 369, 593-601.
- Peat, S., Whelan, W., & Roberts, J. (1957). 778. The structure of lichenin. *Journal of the Chemical Society (Resumed)*, 3916-3924.
- Pérez-Cañadillas, J. M., Guenneugues, M., Campos-Olivas, R., Santoro, J., del Pozo, A. M., Gavilanes, J. G., . . . Bruix, M. (2002). Backbone dynamics of the cytotoxic ribonuclease α -sarcin by ^{15}N NMR relaxation methods. *Journal of Biomolecular NMR*, 24(4), 301-316.
- Perkins, S. J. (1986). Protein volumes and hydration effects. *European Journal of Biochemistry*, 157(1), 169-180.
- Qin, Z., Yan, Q., Yang, S., & Jiang, Z. (2015). Modulating the function of a β -1,3-glucanotransferase to that of an endo- β -1,3-glucanase by structure-based protein engineering. *Applied Microbiology and Biotechnology*, 100(4), 1765-1776. doi:10.1007/s00253-015-7057-4
- Quiocho, F. A. (1986). Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annual review of biochemistry*, 55(1), 287-315.
- Read, S. M., Currie, G., & Bacic, A. (1996). Analysis of the structural heterogeneity of laminarin by electrospray-ionisation-mass spectrometry. *Carbohydrate Research*, 281(2), 187-201. doi:[http://dx.doi.org/10.1016/0008-6215\(95\)00350-9](http://dx.doi.org/10.1016/0008-6215(95)00350-9)
- Rodríguez, R., Villalba, M., Batanero, E., Palomares, O., & Salamanca, G. (2007). Emerging pollen allergens. *Biomedicine & Pharmacotherapy*, 61(1), 1-7. doi:<http://dx.doi.org/10.1016/j.biopha.2006.09.014>
- Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal*, 78(3), 1606-1619.
- Schulz, B. L. (2012). *Beyond the sequon: sites of N-glycosylation*. INTECH Open Access Publisher.

- Shoseyov, O., Shani, Z., & Levy, I. (2006). Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews*, 70(2), 283-295.
- Sillitoe, I., Lewis, T. E., Cuff, A., Das, S., Ashford, P., Dawson, N. L., . . . Orengo, C. A. (2015). CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res*, 43(D1), D376-D381. doi:10.1093/nar/gku947
- Singh, D., Ambroise, A., Haicour, R., Sihachakr, D., & Rajam, M. V. (2014). Increased resistance to fungal wilts in transgenic eggplant expressing alfalfa glucanase gene. *Physiology and Molecular Biology of Plants*, 20(2), 143-150. doi:10.1007/s12298-014-0225-7
- Steinbüchel, A. (2002). *Biopolymers*: Wiley-Blackwell.
- Stone, B. A., & Clarke, A. E. (1992). *Chemistry and Biology of 1, 3-[beta]-Glucans*: La Trobe University Press Melbourne.
- Sun, L., Gurnon, J. R., Adams, B. J., Graves, M. V., & Van Etten, J. L. (2000). Characterisation of a β -1,3-Glucanase Encoded by Chlorella Virus PBCV-1. *Virology*, 276(1), 27-36. doi:<http://dx.doi.org/10.1006/viro.2000.0500>
- Sunderasan, E., Hamzah, S., Hamid, S., Ward, M., Yeang, H., & Cardoso, M. (1995). Latex B-serum-1, 3-Glucanase (Hev b II) and a component of the microhelix (Hev b IV) are major latex allergens.
- Torres, M. (2014). *Taumatinas y beta-1,3-glucanasas: dos familias de proteínas de defensa de plantas implicadas en procesos alérgicos*. Universidad Complutense de Madrid.
- Torres, M., Palomares, O., Quiralte, J., Pauli, G., Rodríguez, R., & Villalba, M. (2015). An Enzymatically Active β -1,3-Glucanase from Ash Pollen with Allergenic Properties: A Particular Member in the Oleaceae Family. *PLoS One*, 10(7), e0133066. doi:10.1371/journal.pone.0133066
- Treviño, M. Á., García-Mayoral, M. F., Barral, P., Villalba, M., Santoro, J., Rico, M., . . . Bruix, M. (2004). NMR solution structure of Ole e 6, a major allergen from olive tree pollen. *Journal of Biological Chemistry*, 279(37), 39035-39041.
- Treviño, M. Á., Palomares, O., Castrillo, I., Villalba, M., Rodríguez, R., Rico, M., . . . Bruix, M. (2008). Solution structure of the C-terminal domain of Ole e 9, a major allergen of olive pollen. *Protein Science*, 17(2), 371-376.
- Valliere-Douglass, J. F., Eakin, C. M., Wallace, A., Ketchem, R. R., Wang, W., Treuheit, M. J., & Balland, A. (2010). Glutamine-linked and Non-consensus Asparagine-linked Oligosaccharides Present in Human Recombinant Antibodies Define Novel Protein Glycosylation Motifs. *Journal of Biological Chemistry*, 285(21), 16012-16022. doi:10.1074/jbc.M109.096412
- Varea, J., Saiz, J. L., López-Zumel, C., Monterroso, B., Medrano, F. J., Arrondo, J. L. R., . . . Menéndez, M. (2000). Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? *Journal of Biological Chemistry*, 275(35), 26842-26855.
- Villalba, M., Rodríguez, R., & Batanero, E. (2014). The spectrum of olive pollen allergens. From structures to diagnosis and treatment. *Methods*, 66(1), 44-54.
- Wagner, S., Radauer, C., Hafner, C., Fuchs, H., Jensen-Jarolim, E., Wüthrich, B., . . . Breiteneder, H. (2004). Characterisation of cross-reactive bell pepper allergens involved in the latex-fruit syndrome. *Clinical & Experimental Allergy*, 34(11), 1739-1746.
- Williamson, M. P. (2013). Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 73, 1-16. doi:<http://dx.doi.org/10.1016/j.pnmrs.2013.02.001>

- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., . . . Sykes, B. D. (1995). ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *Journal of Biomolecular NMR*, 6(2), 135-140.
- Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol*, 222(2), 311-333.
- Wojtkowiak, A., Witek, K., Hennig, J., & Jaskolski, M. (2012). Two high-resolution structures of potato endo-1,3-[beta]-glucanase reveal subdomain flexibility with implications for substrate binding. *Acta Crystallographica Section D*, 68(6), 713-723. doi:doi:10.1107/S090744491200995X
- Woolfson, D. N., Evans, P. A., Hutchinson, E. G., & Thornton, J. M. (1993). Topological and stereochemical restrictions in β -sandwich protein structures. *Protein engineering*, 6(5), 461-470.
- Zamora-Carreras, H., Torres, M., Bustamante, N., Macedo, A. L., Rodríguez, R., Villalba, M., & Bruix, M. (2015). The C-terminal domains of two homologous Oleaceae β -1,3-glucanases recognise carbohydrates differently: Laminarin binding by NMR. *Archives of Biochemistry and Biophysics*, 580, 93-101. doi:<http://dx.doi.org/10.1016/j.abb.2015.07.004>

CHAPTER 5

CYTOSKELETAL MOTORS



The work included in this chapter has been developed in collaboration with the group of Dr. I. Rodríguez-Crespo, belonging to the Department of Biochemistry and Molecular Biology from Universidad Complutense de Madrid (Madrid, Spain). Biochemical and molecular biology experimentation have been carried out by Dr. J. Merino-Gracia and Dr. I. Rodríguez-Crespo.

INTRODUCTION

1. MOLECULAR MOTORS

Molecular motors are amongst the most enthralling biological devices that can be found in nature. Their name comes from the fact that these molecules are capable of transforming energy in motion or mechanical work. Considering this definition, a wide range of molecules can be catalogued as molecular motors:

- Cytoskeletal or cytoplasmic motors: proteins which are able to move along linear tracks with a certain direction, and transporting an associated cargo. (Alberts B, Johnson A, Lewis J, & al., 2002)
- Polymerization motors: the polymerization process of some proteins (such as actin, dynamin, or the microtubules) generates forces that can originate motion and propulsion. (McNiven, 1998; Meiri et al., 2014; Roux, Uyhazi, Frost, & De Camilli, 2006)
- Rotary motors: some protein complexes (such as the F_0-F_1 -ATP synthase family) during the performance of their activities transform energy in a rotational motion which couples different steps of the process. (Meiri et al., 2014; Wilkens, 2005)
- Nucleic acid motors: protein complexes (such as helicases, topoisomerases, and polymerases) that execute an activity on nucleic acids which implies directional motions. (Singleton, Dillingham, & Wigley, 2007)

2. CYTOSKELETAL MOTORS

Cytoskeletal motors are proteins that are capable of binding to polarized cytoskeletal filaments and use the energy derived from repeated cycles of ATP hydrolysis to move steadily along it. Many cytoskeletal motors can be differentiated on the bases of the type of filament they bind to (either actin or microtubules), the direction in which they move along the filament, and the kind of cargoes they carry. (Alberts B et al., 2002)

2.1 COMPONENTS OF THE CYTOSKELETON

Cytoskeleton is composed by three types of filaments: actin filaments (microfilaments), intermediate filaments, and microtubules. All these components possess structural functions to maintain or transform the cell shape and protect it from mechanical stress. They are organized into dynamical networks and play a key role in the arrangement and maintenance of the integrity of intracellular compartments. In the case of the actin

filaments and microtubules, they also act as tracks for cytoskeletal motors, and thus, they are crucial for the organization of the cellular components. (Fletcher & Mullins, 2010)

Filament network architecture is controlled by the regulation of the polymerization processes through a set of regulatory proteins, such as nucleation–promoting factors (they initiate filament formation), capping proteins (they terminate filament growth), polymerases (they promote faster or more sustained filament growth), depolymerizing factors and severing factors (they disassemble filaments), and cross–linkers and stabilizing proteins (they organize and reinforce higher–order network structures). (Fletcher & Mullins, 2010)

Actin filaments, intermediate filaments and microtubules present important differences in some of their properties, such as their mechanical stiffness, the dynamics of their assembly, their polarity, and the type of cytoskeletal motors which they can bind. Intermediate filaments, unlike actin filaments and microtubules, are not polarized and, therefore, directional movement of cytoskeletal motors along them is not possible. The main characteristics of the actin filaments and microtubules are: (Fletcher & Mullins, 2010)

- Actin filaments

Actin filaments or microfilaments are the thinnest cytoskeletal structures with a diameter of around 6 nm. They are dynamic flexible polymers of actin, a protein that can be found in two forms: the monomeric globular G–actin and the filament–forming polymeric F–actin. Actin is capable of binding and hydrolysing ATP units. Actin assembly process passes off in three steps (Figure 5.1): first, a nucleation phase takes place when active G–actin monomers (G–actin bound to ATP) aggregate into short, unstable oligomers, until they reach a certain length and they can act as stable seeds. In the second step, elongation into a filament occurs by the addition of actin monomers to both of its ends. In the last step, a steady state is reached, as the equilibrium between G–actin and F–actin is established. The total filament mass remains constant as G–actin monomers exchange with subunits at the filament ends. After the incorporation into a filament, actin subunits gradually hydrolyse ATP and become stable ADP–F actin. Actin filaments show polarity, since the monomers all orient with their structural ATP–binding cleft towards the same end of the filament (minus–end). Filament growth is faster in the plus–end. (Fletcher & Mullins, 2010; Lodish et al., 2008)

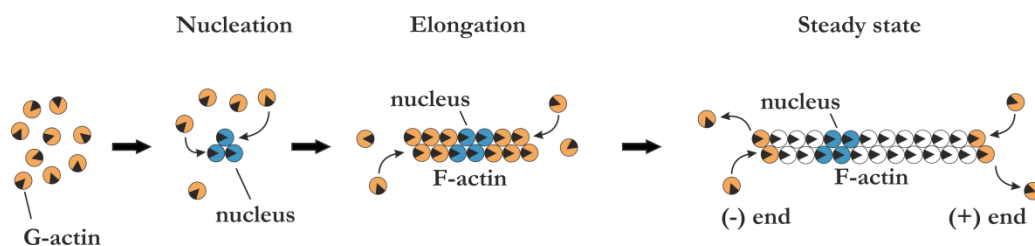


Figure 5.1 Actin assembly scheme. ATP–G–actin monomers are shown in orange), stable oligomers in blue, stable ADP–F–actin in white, ATP–binding clefts are shown as black triangles. (Adapted from (Lodish et al., 2008)).

■ Microtubules

Microtubules are the stiffest cytoskeletal structures and show a complex assembly/disassembly dynamics. They are polymers of globular tubulin subunits arranged in a cylindrical hollow tube of 24 nm in diameter. Their building blocks are very stable heterodimers of α - and β -tubulin (Figure 5.2) that can bind one GTP molecule in each subunit. The GTP-binding site in α -tubulin binds GTP irreversibly, whereas in β -tubulin GTP binding is reversible and hydrolysis to yield GDP occurs. In this second site, GDP can be displaced by GTP and it is called “exchangeable site”. Longitudinal association of $\alpha\beta$ -tubulin units leads to the formation of protofilaments, and this protofilaments can establish interactions laterally between them to form the cylindrical microtubules. Typically, each microtubule is composed by 13 protofilaments. (Figure 5.2) Microtubules are polar structures, and this polarity arises from the head-to-tail arrangement of the $\alpha\beta$ -tubulin dimers in the protofilaments. All the protofilaments have the same orientations, and thus, one end of the microtubule is ringed by α -tubulin (minus-end) and the other one, by β -tubulin (plus-end). (Figure 5.2) In most cases, the minus-end is adjacent to a microtubule-organizing centre (such as centrosomes or basal bodies), from which it assembles, and the plus-end is distal and it is where assembly and disassembly occurs preferentially. (Fletcher & Mullins, 2010; Lodish et al., 2008)

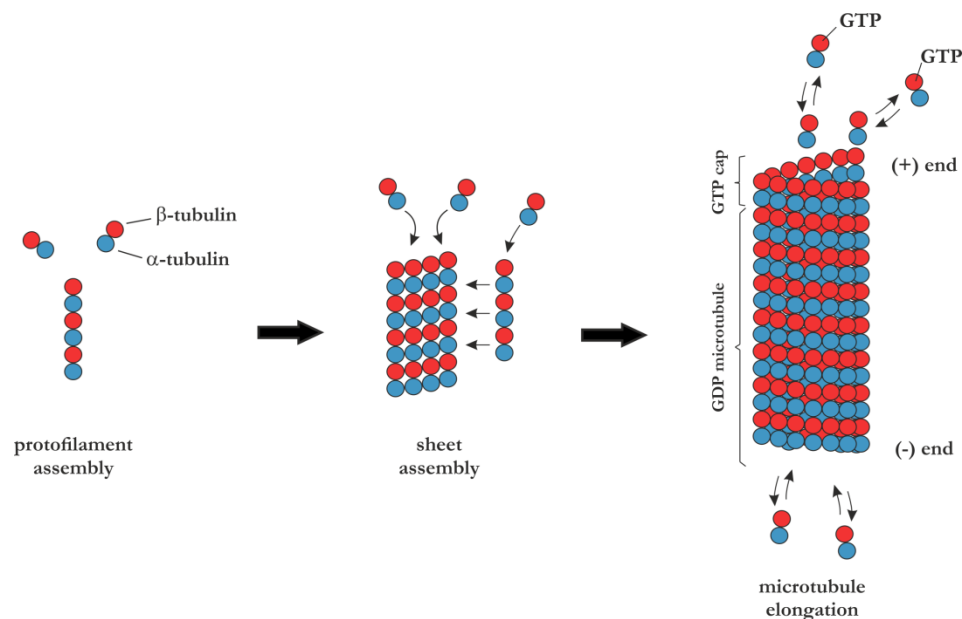


Figure 5.2. Microtubule assembly scheme. Free $\alpha\beta$ -tubulin dimers associate longitudinally to form short protofilaments. These protofilaments, which are unstable, quickly associate laterally into more stable curved sheets. Eventually, a sheet wraps around into a microtubule with thirteen protofilaments. The microtubule then grows by the addition of subunits to the ends of protofilaments composing the microtubule wall. Free tubulin dimers have GTP molecules bound and, after the incorporation to a microtubule, GTP associated to the β -tubulin is hydrolysed to yield GDP. Frequently, a cap of GTP-bound subunits is formed at the plus-end, where the rate of polymerization is faster than in the minus-end.

2.2 CLASSES OF CYTOSKELETAL MOTORS

Three main groups of cytoskeletal motors are known: myosins, kinesins and dyneins:

■ Myosins

This was the earliest group of motor proteins identified and they were first described in the skeletal muscle cells, where they are responsible for the generation of the force necessary for muscle contraction, as they move along actin filaments. Myosins form a superfamily containing at least 18 different classes (although recent classifications based on phylogenetic analyses include up to 31 classes, (Sebé-Pedrós, Grau-Bové, Richards, & Ruiz-Trillo, 2014)), and they can be found in a wide range of organisms, including protozoa, fungi, algae, plants, molluscs, amphibians, birds, and mammals. General myosin structure consists of an N-terminal myosin motor domain that interacts with actin filaments, and a variable C-terminal tail domain. Myosin tail domains have followed different evolutionary pathways in order to allow the proteins the interaction with other subunits, the dimerization and the association with different cargoes. (Figure 5.3) (Alberts B et al., 2002; Schliwa & Woehlke, 2003)

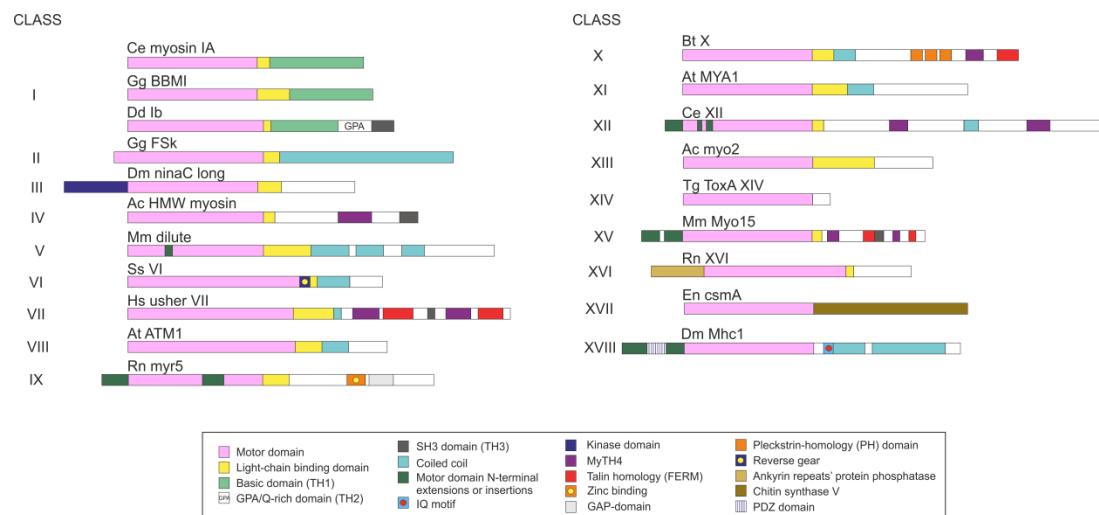


Figure 5.3. Classes of myosins. Representative myosins belonging to the first XVIII identified classes. Their modular components are explained. The two-letter code before protein name is referred to the species (Ce = *Caenorhabditis elegans*; Gg = *Gallus gallus*; Dd = *Dictyostelium discoideum*; Dm = *Drosophila melanogaster*; Ac = *Acanthamoeba castellanii*; Mm = *Mus musculus*; Ss = *Sus scrofa domestica*; Hs = *Homo sapiens*; At = *Arabidopsis thaliana*; Rn = *Rattus norvegicus*; Bt = *Bos Taurus*; Tg = *Toxoplasma gondii*; En = *Emiricella nidulans*). Adapted from (Kendrick-Jones, Hodge, Lister, Roberts, & Buss, 2001) and (Tzolovsky, Millo, Pathirana, Wood, & Bownes, 2002).

■ Kinesins

Kinesins are cytoskeletal motor proteins that move along the microtubules. They form a large protein superfamily with an only common element: the motor domain. Usually, the motor domain, which interacts with microtubules, is located in the N-terminal region of the heavy chain and generates a motion towards the plus-end of the microtubule.

However, a particular family of the kinesins has the motor domain at their C-terminal region and the motion generated is oriented to the minus-end of the microtubule. Different kinesins can act as monomers, homodimers or heterodimers. The majority of kinesins possess a binding site in the tail that can bind either membrane-enclosed organelles or another microtubule. These motor proteins are involved in cellular processes such as mitotic and meiotic spindle formation, or chromosome segregation during cell division. (Alberts B et al., 2002)

- Dyneins

Dyneins move along microtubules, like kinesins do, but they are not related. The motion of these motor proteins is directed towards the minus-end of the microtubules. Their structure comprises two or three heavy chains, including the motor domain, and variable large number of associated light chains. Two major classes of dyneins can be distinguished: axonemal dyneins and cytoplasmic dyneins. Axonemal dyneins are typically heterodimers or heterotrimers with two or three heads, respectively, which contain the motor domains. They are related to the beating motion of cilia and flagella. Cytoplasmic dyneins are typically homodimers with two heads containing the motor domains. These dyneins are responsible for the organization of some organelles and the vesicle traffic inside the cell. (Alberts B et al., 2002; Diamant & Carter, 2013; K. K. Pfister et al., 2006; Schliwa & Woehlke, 2003; Wickstead & Gull, 2011)

2.3 MECHANISMS OF CYTOSKELETAL MOTORS

Cytoskeletal motors are proteins capable of binding and hydrolyse ATP molecules. When an ATP molecule is hydrolysed in the active site of a cytoskeletal motor, the loss of the γ -phosphate group generates an empty space of approximately 0.5 nm, which probably triggers the rearrangement of the structural elements surrounding the ATP-binding site. This rearrangement is amplified through a coordinated communication of the conformational change from the ATP-binding site to the track-binding regions, producing the release from the cytoskeletal components and the propulsion in a unique direction along the track to reach a new binding site on the filament or microtubule (Figure 5.4). The consequence of the cyclic repetition of this process leads to a unidirectional motion of the motor along the cytoskeletal track. (Alberts B et al., 2002; Schliwa & Woehlke, 2003)

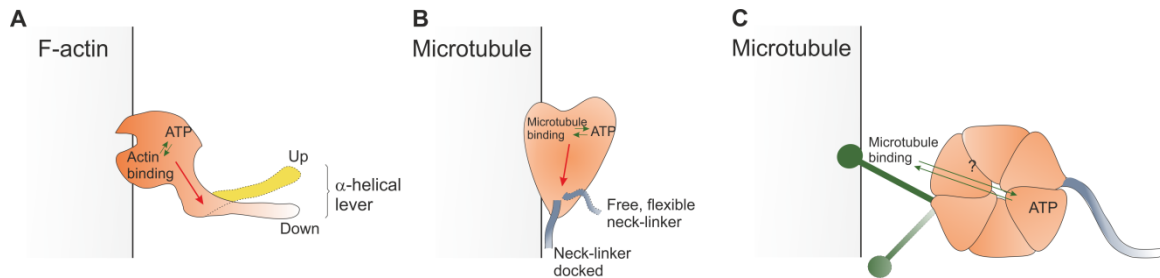


Figure 5.4. Scheme of the intramolecular communication mechanism in cytoskeletal motors. ATP hydrolysis causes a conformational change near the ATP-binding site, and this change is communicated to the track-binding site (green arrows). **A.** Myosin is attached to the F-actin filament, and the conformational change is amplified by a helix stabilized by light chains (not shown), which acts as swinging lever. **B.** Kinesin is bound to a microtubule, and the amplification is realized by a flexible element, the neck linker, that connects the motor domain with the neck helix. **C.** In the dynein motors, the pathways of intramolecular communication are still unknown.

2.4 FUNCTIONS OF CYTOSKELETAL MOTORS

The three main functions performed by cytoskeletal motors are the transport and positioning of membrane-enclosed organelles, the generation of contraction movements and the ciliary and flagellar beating.

▪ Organelle transport and positioning

Typically, kinesins and dyneins are responsible for this function. The best example is the organization of the endoplasmic reticulum (ER) and the Golgi apparatus. The ER structures are aligned with microtubules and spread to the edge of the cell. Kinesins can attach the ER membranes to the microtubule cell network and drag them out, moving towards microtubule plus-ends. Conversely, dyneins interact with the membrane of the Golgi apparatus and displace this organelle near the cell centre, walking along the microtubules towards their minus-ends (Figure 5.5). (Alberts B et al., 2002; Schliwa & Woehlke, 2003)

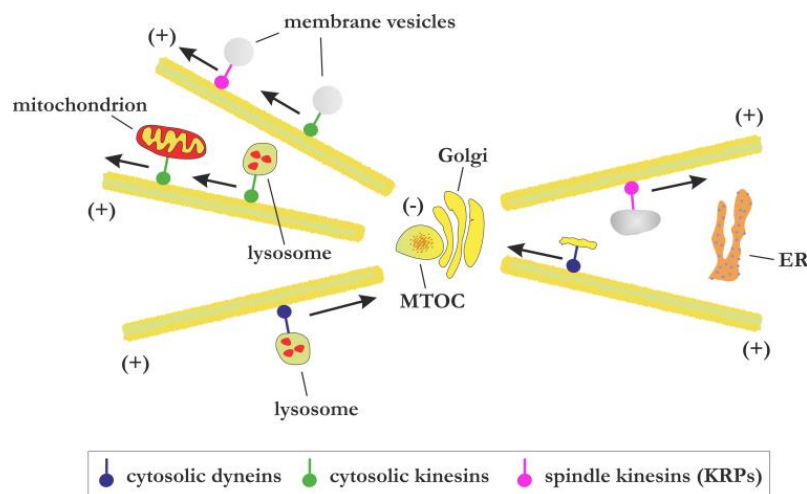


Figure 5.5. General scheme of the organelle transport mediated by kinesins and dyneins in the cell. MTOC: Microtubule organizing centre. ER: endoplasmic reticulum. Adapted from (Lodish et al., 2008).

- Contraction

Myosins are usually related to contraction generation. This process takes place through the sliding of highly organized arrays of actin filaments against arrays of myosin filaments, within the sarcomere, the basic unit of the striated muscle tissue. Myosin filaments, each of them composed by around 300 motor heads, walk towards the plus-ends of actin filaments (Figure 5.6). (Alberts B et al., 2002; Schliwa & Woehlke, 2003)

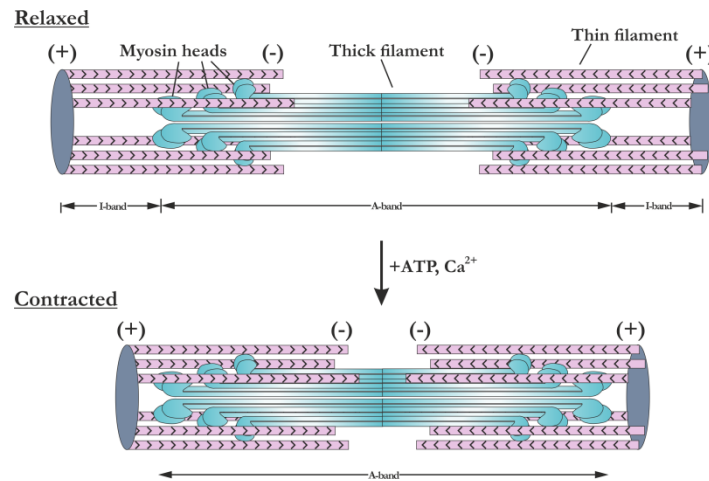


Figure 5.6. Sliding-filament model for contraction in striated muscle. A sarcomere is represented in relaxed state (**top**), and in the contracted state (**bottom**). The presence of ATP and Ca²⁺ causes myosin heads pivot, pulling the actin thin filaments towards the centre of the sarcomere. Since thin filaments are anchored at the Z-discs (grey), the movement shortens the sarcomere length during contraction. Adapted from (Lodish et al., 2008).

- Ciliary and flagellar beating

Cilia and flagella are highly specialized structures built from microtubules and dynein. Basically, these structures are thin cellular appendages with a bundle of microtubules in their core. Their function is to provide the cell with a way of propelling in a fluid or move the surrounding fluid over the cell surface. The movement of cilia and flagella comes from the bending of its core, called axoneme (Figure 5.7). This bending is produced by the movement of dynein motors along the microtubules making up the axoneme, causing a relative slide between adjacent microtubules. (Alberts B et al., 2002; Schliwa & Woehlke, 2003)

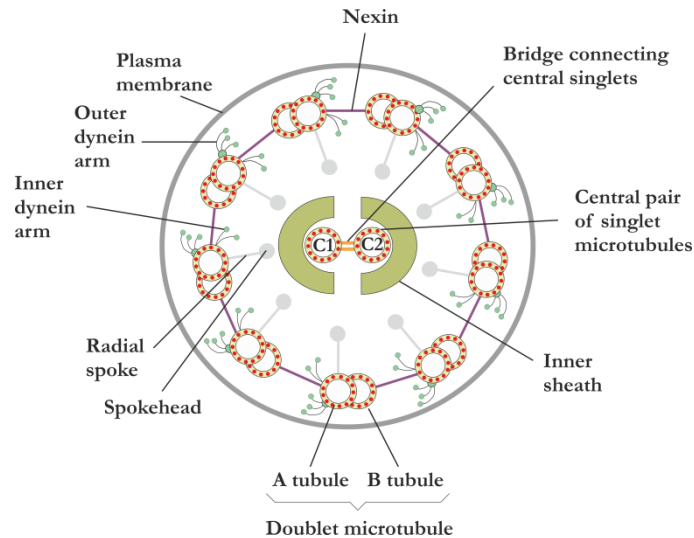


Figure 5.7. Cross-section scheme of a typical flagellum's axoneme. Its major structures are indicated. Adapted from (Lodish et al., 2008).

2.5 MEMBRANE ASSOCIATION OF CYTOSKELETAL MOTORS

There is a variety of ways in which cytoskeletal motors can bind their membrane-enclosed cargoes, but usually the linkage involves the formation of a complex, with protein components mediating the interaction between the motor and the cargo.

Some cytoskeletal motors interact directly with the lipid components of the membranes. This is the case of myosin I, whose tail domain is capable of binding acidic phospholipids present in the organelle membranes (Figure 5.8H). Direct interaction with phospholipids has been also observed in the kinesin-like protein Unc104, through its pleckstrin homology domain (Figure 5.8D). (Schliwa & Woehlke, 2003)

Other motors establish direct interactions with integral membrane proteins contained in the organelle membrane. One example is the Tctex-1 light chain of dynein, which binds to the membrane-embedded rhodopsin (Figure 5.8F). In some cases, the interacting protein is a transmembrane receptor, as observed in the kinesin light chains (Figure 5.8A). (Schliwa & Woehlke, 2003)

Indirect interactions with transmembrane receptors are also possible, and they can be mediated by linker proteins (kinesin heavy chains) or linker complexes (kinesin light chains) (Figure 5.8B and 5.8C). Indirect interactions with other integral membrane proteins have been observed mediated by carrier proteins (the case of myosin V interacting with membrane-anchored rab27a via melanophilin, Figure 5.8G), or by big complexes (as in the interaction of cytoplasmic dyneins and integral membrane proteins via dynactin complex and spectrin, Figure 5.8E). (Schliwa & Woehlke, 2003)

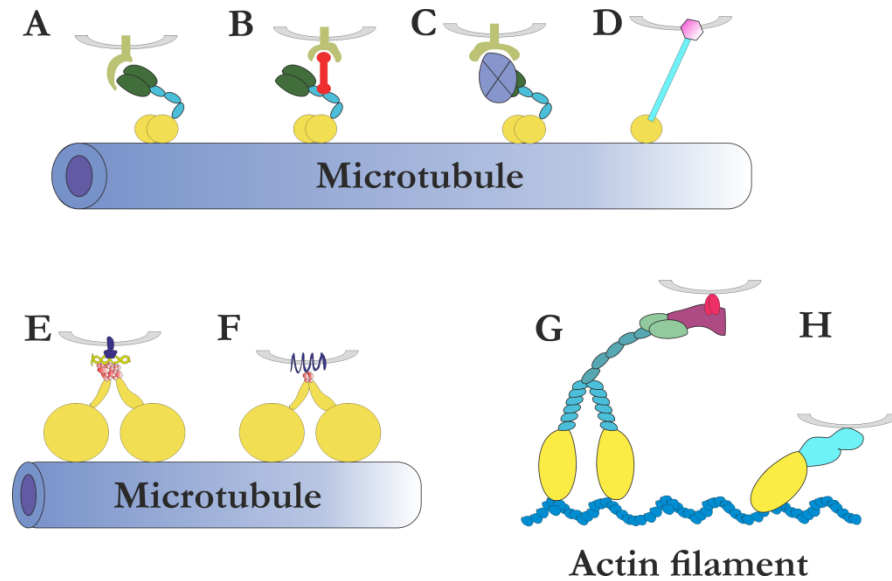


Figure 5.8. Types of cargo-linkage for kinesins (A–D), dyneins (E–F), and myosins (G–H). For description of each type of linkage, see the text above. Plasma membrane is represented in grey. Cellular motors are in yellow. Adapted from (Schliwa & Woehlke, 2003).

3. COMPONENTS OF THE DYNEIN COMPLEX

Dyneins are motor proteins forming big complexes with several subunits and accessory proteins. They constitute a big family and can be classified following different criteria, such as phylogenetical evolution, type of heavy chains, or subunit composition. As explained before, dyneins move towards the minus-end of microtubules using the energy obtained from the hydrolysis of ATP molecules. Several dynein subunits can be distinguished: (Diamant & Carter, 2013; Merino-Gracia, 2016; K. K. Pfister et al., 2006; Tynan, Gee, & Vallee, 2000)

- Dynein heavy chains (DHCs)

They possess ATPase activity and generate the driving force of the protein motor. They are huge proteins comprising over 4,000 amino acids and present a modular structure divided into two domains: an N-terminal tail that acts as a scaffold for the assembly of the dynein complex, and a C-terminal head containing the enzymatic and mechanic activity. Structural and functional features of the heavy chains are fairly conserved in axonemal and cytoplasmic dyneins.

- Dynein intermediate chains (DICs)

DICs are only present in dimeric dyneins. DICs associate to DHCs through WD40 repeats located in the C-terminal region. The N-terminal segment is an extended region with variable length where other subunits, such as dynactin complex or the light chains of cytoplasmic dyneins, can interact. (Vaughan, Leszyk, & Vaughan, 2001)

- Dynein light–intermediate chains (DLICs)

They are polypeptides of around 50–60 kDa, exclusively found in cytoplasmic dyneins. They are GTPases and are directly involved in cellular migration and mitotic spindle formation through their interaction with centrosomal proteins.

- Dynein light chains (DLCs)

DLCs are small globular proteins that participate in the dynein complex assembly. Three families are described in cytoplasmic dyneins: DYNLL, DYNLT, and DYNLRB. The DLCs belonging to those families associate to the N–terminal domain of dimeric DICs, in consecutive and non–overlapping positions. It was proposed that DLCs act as adaptors to associate proteins and complexes to the dyneins in order to be transported along the microtubules. However, some structural studies have not been able to verify that hypothesis, and thus, the exact function of the DLCs remain to be unveiled. (S. J. King, Bonilla, Rodgers, & Schroer, 2002; Williams et al., 2007)

3.1 DYNEIN INTERMEDIATE CHAINS (DICs)

Dynein intermediate chains (DICs) are essential components of the dynein complex, and can be found both in axonemal and cytoplasmic dyneins. The molecular mass of DICs is around 70 kDa and their function is essentially structural, serving as scaffold for the assembly of the complex. A DIC dimer stabilized by the DLCs acts as a platform for the association of the DHCs and accessory complexes, such as dynactin. There is a variety of DIC isoforms which define different populations with distinct functionalities, and all of them can combine to form diverse homo– and hetero–dimers capable of binding to the DLCs. (Lo, Kan, & Pfister, 2006; Lo, Kogoy, Rasoul, King, & Pfister, 2007; Merino-Gracia, 2016)

Two differentiated domains can be distinguished in DICs:

- N–terminal domain: it is a domain with a high ratio of charged/hydrophobic residues that minimizes the tendency for a hydrophobic–collapse–driven folding, favouring an intrinsically disordered state. The initial segment of the N–terminal domain possesses a highly unstructured Ser– and Pro–rich region, which contains phosphorylation sites. DLCs binding sites are located in this intrinsically disordered region, and the association of the light chains induces a certain order with the appearance of a coiled coil and the subsequent dimerization. (Makokha, Hare, Li, Hays, & Barbar, 2002; Nyarko, Hare, Hays, & Barbar, 2004)
- C–terminal domain: it comprises seven WD40 repeats (repeats of around 40 aa ending in a WD dipeptide) that fold in a coiled–coil conformation (β –propeller). This region is directly responsible for the DHCs association. (Tynan et al., 2000)

3.2 DYNEIN LIGHT CHAINS (DLCs)

Dynein light chains (DLCs) are small globular homodimeric proteins that bind to the N-terminal domain of DICs, as well as to other proteins. It was initially thought that DLCs act as adaptors by associating proteins to dyneins for their intracellular transport. However, many thermodynamical and structural studies defied that hypothesis, since cell proteins associate to DLCs using the same binding sites as DICs do. (Merino-Gracia, 2016)

DLCs are dimeric and thus they possess two identical interaction surfaces, leading to a high propensity to interact with proteins which are also dimeric. At present, DLCs are considered as molecular staples that play a structural role maintaining dimeric complexes. There are three DLC families described (DYNLL, DYNLRB, and DYNLT), each of them comprising two members. DLCs always interact with the disordered region of the DICs, where they function as a scaffold for the sequential assembly of the dynein complex. (Barbar, 2008; Merino-Gracia, 2016)

3.2.1 DYNLL FAMILY

DYNLL family consists of two proteins, DYNLL1 and DYNLL2 (also known as LC8a and LC8b, respectively). They are small globular proteins composed by 89 amino acids and with a molecular mass of 10.5 kDa (Figure 5.9). They are dimeric, highly conserved and expressed ubiquitously. Both proteins are almost identical, except for a few residues which determine the specificity for some of their ligands. (Day et al., 2004; Merino-Gracia, 2016)

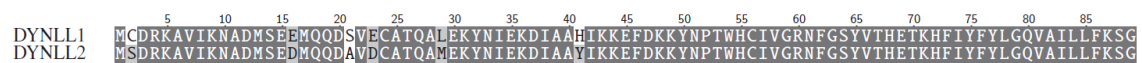


Figure 5.9. Sequence alignment of human DYNLL1 and DYNLL2. The minor differences seems to be related to changes in binding specificity, and redox behaviour (Merino-Gracia, 2016). Colour code: identical residues in dark grey, similar residues in light grey.

DYNLL1 folds into a unique structural unit in which two antiparallel α -helices from the N-terminal region lean against a five-strand antiparallel β -sheet (Figures 5.10 and 5.11). The only structural change between the monomeric and the dimeric forms takes place on the β 3 loop, which is unordered with a slight helical tendency in the monomer. When DYNLL1 dimerizes, the two β -sheets face each other; the β 3 loop of one monomer adopts a β -strand structure and incorporates to the β -sheet of the other one, extending it. β 1 strand, located in one of the edges of the β -sheet, has a structural distortion (β -bulge) which hinders the incorporation of a new chain, whereas β 3' strand (situated in the other edge of the β -sheet) with a hydrophobic environment, favours the insertion of an additional strand (belonging to the target protein) (Figure 5.11). (Liang, Jaffrey, Guo, Snyder, & Clardy, 1999; Makokha, Huang, Montelione, Edison, & Barbar, 2004; Merino-Gracia, 2016)

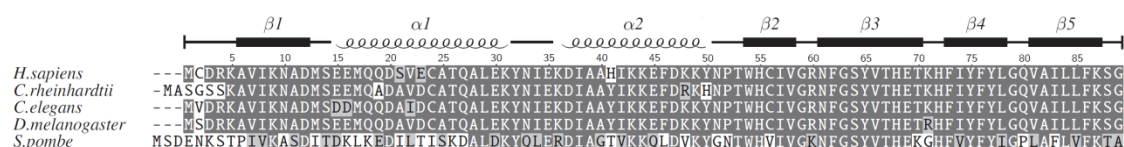


Figure 5.10. Sequence alignment of human DYNLL1 and homologues from different representative eukaryotes. Secondary structure elements are indicated. Colour code: identical residues in dark grey, similar residues in light grey.

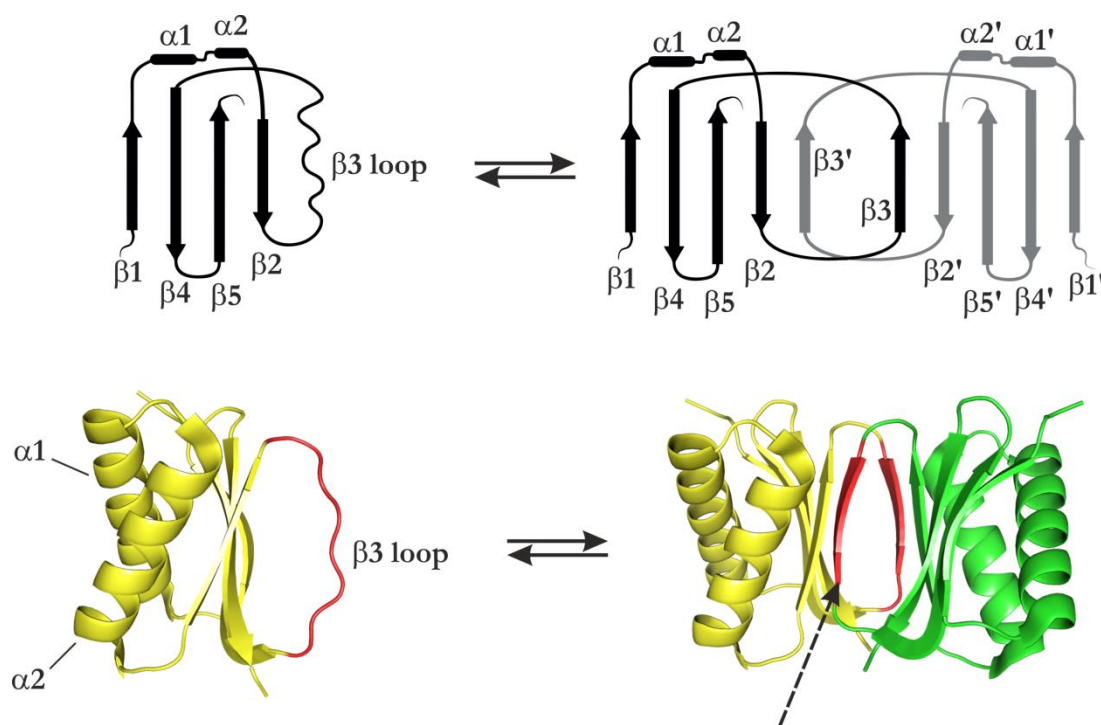


Figure 5.11. DYNLL1 structure and dimer formation. The 3D structure displayed corresponds to the X-ray diffraction structure of the apo form of DYNLL1 from *D. melanogaster* deposited in the Protein Data Bank with accession code 3BRI. (Benison, Karplus, & Barbar, 2008) $\beta 3$ loop, which forms a β -strand when dimer is formed, is coloured in red. Dashed arrow indicates one of the symmetrical binding sites.

DYNLL1 dimer interacts with cell proteins through two binding sites located in the interface between both monomers. The binding sites are two symmetrical grooves upholstered with residues from both monomers. The deepest part of the groove shows a highly hydrophobic character, and the incoming protein inserts in it adding an extra β -strand to the preformed β -sheet (Figure 5.11). The interaction of the cell proteins with DYNLL1 takes place through interacting motifs which typically display two types of consensus sequences: KxTQTX and xGIQVD. The central glutamine of these sequences seems to be decisive for the interaction. (Fan, Zhang, Tochio, & Zhang, 2002; Merino-Gracia, 2016; Rodríguez-Crespo et al., 2001)

DYNLL1 is an integral component of both cytoplasmic and flagellar dyneins. At present, it is believed that DYNLL1 acts as a molecules staple, favouring and stabilizing the DLC–DIC assembly to form a pre-complex which leads to the union with the DHCs and

the global stabilization of the dynein complex. It has been described the role of DYNLL1 in the regulation of apoptosis, through the interaction with BimL, a proapoptotic regulator. (Barbar, 2008; S. M. King, 2000; Merino-Gracia, 2016)

3.2.2 DYNLRB FAMILY

DYNLRB/Roadblock family (also known as LC7 or Km23 proteins) comprehends two members: DYNLRB1 and DYNLRB2 (Figure 5.12). Both form homodimers and bind to DICs near to the end of the N-terminal domain, close to the WD repeats. They fold in a unique way, dissimilar to those described for the other known DLCs. Their structure is characterised by a large β -sheet surrounding a helical bundle made up of the $\alpha 2$ helices from both monomers. (Merino-Gracia, 2016; Susalka et al., 2002)

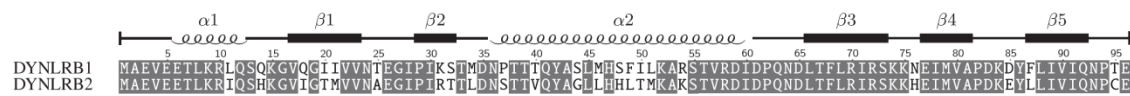


Figure 5.12. Sequence alignment of human DYNLRB1 and DYNLRB2. Secondary structure elements are indicated. Colour code: identical residues in dark grey, similar residues in light grey.

A few proteins have been reported as ligands for DYNLRB, including DIC, transforming growth factor β (TGF β), folate transporter, or Rab6 GTPase. The association of DYNLRB with DIC occurs through the residues 182–219, which fold in a split α -helix and get inserted on the β -sheet in an antiparallel orientation. In this interaction, amino acid side chains play a key role, making the recognition more selective than in the case of DYNLL1 or DYNLT1. (Ashokkumar, Nabokina, Ma, & Said, 2009; Jin, Gao, & Mulder, 2009; Merino-Gracia, 2016; Wanschers et al., 2008)

DYNLRB acts as a structural element in the dynein complex, and its interaction, which depends on the phosphorylation, is favoured by the formation of a previous complex with the other DLCs and DIC. DYNLRB is also implied in the signalling of TGF β complex, helping to its assembly and operating as an intermediate in the TGF β -signalling-dependent dynein activation. (Hall, Song, Karplus, & Barbar, 2010; Jin, Gao, & Mulder, 2013; Merino-Gracia, 2016; Nyarko & Barbar, 2011)

4. DYNLT FAMILY

DYNLT family proteins (DYNLT1/Tctex1 and DYNLT3/RP3) are structurally homologous to DYNLL1, although sequence homology is not observed. They are integral components of both flagellar and cytoplasmic dyneins, and are capable of associate to numerous cell proteins. (Merino-Gracia, 2016)

Besides DYNLT1/Tctex1 and DYNLT3/RP3, many proteins containing DYNLT1-like domains have been characterised (such as TC1D3 or TC1D4). Their functions are not totally understood. (DiBella et al., 2001; Merino-Gracia, 2016; Patel-King, Benashski, Harrison, & King, 1997)

4.1 STRUCTURE OF DYNLT PROTEINS

DYNLT1/Tctex1 and DYNLT3/RP3 (Figure 5.13) are both small globular proteins whose folded structure is only stable when they form dimers. In fact, the association of the monomers is very strong and dissociation constant has been reported to be in the femtomolar order. Thus, the homodimer is indivisible in physiological conditions. (Merino-Gracia, 2016; Talbott, Hare, Nyarko, Hays, & Barbar, 2006)

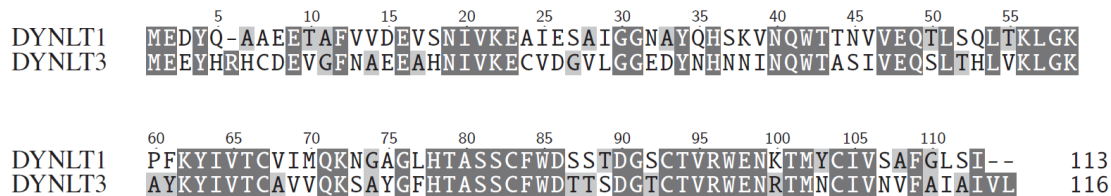


Figure 5.13. Sequence alignment of human DYNLT1 and DYNLT3. Colour code: identical residues in dark grey, similar residues in light grey.

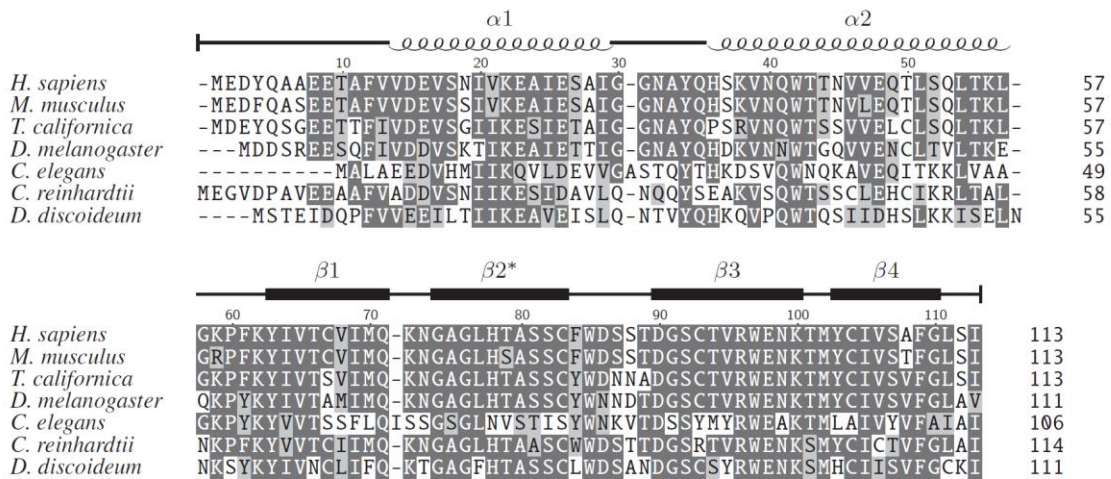


Figure 5.14. Sequence alignment of human DYNLT1 and homologues from different representative eukaryotes. Secondary structure elements are indicated. Colour code: identical residues in dark grey, similar residues in light grey.

DYNLT1 structure is essentially equal to that of DYNLL1, with a similar distribution of the secondary structure elements (Figure 5.14), including the presence of a swapped β -strand (β_2). The initial β -strand found in the N-terminal region of DYNLL1 is not present in DYNLT1 (Figure 5.15). The dimerization interface, which is larger than that from DYNLL1, is composed exclusively of face-to-face packed β -sheets and the swapped β -strands. The antiparallel α -helices of the N-terminal region lean on the opposite side of the β -sheets. In DYNLT1, the β -sheet is formed by four β -strands (one of them swapped), and gets extended with an additional strand from the ligand when it is bound to the protein (Figure 5.15). (Hall, Karplus, & Barbar, 2009; Merino-Gracia, 2016; Williams et al., 2007; Wu, Maciejewski, Takebe, & King, 2005)

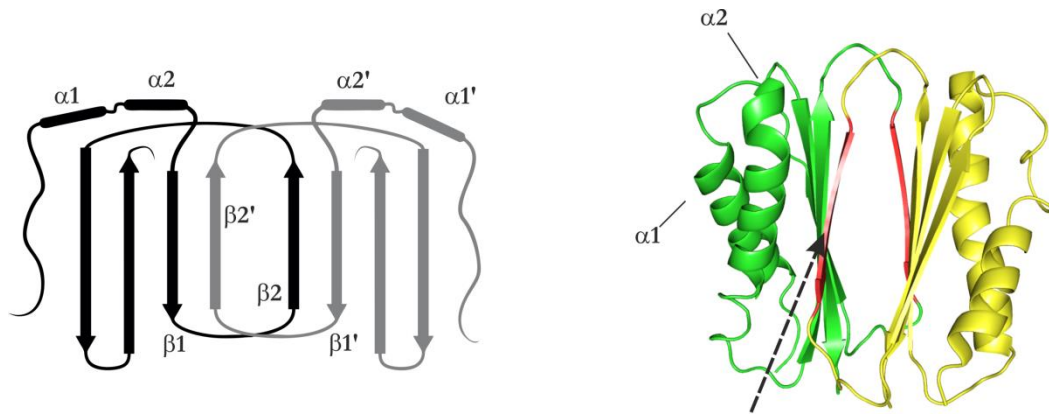


Figure 5.15. DYNLT1 structure and dimer formation. The 3D structure displayed corresponds to the X-ray diffraction structure of the apo form of DYNLT1 from *D. melanogaster* deposited in the Protein Data Bank with accession code 3FM7. (Hall et al., 2009) $\beta 2$ loop, which forms a β -strand when dimer is formed, is coloured in red. Dashed arrow indicates one of the symmetrical binding sites.

The formation of the dimer leads to a compact structure with two symmetric hydrophobic grooves, with a certain electrostatic polarity in the case of the human DYNLT1. The ligands interact through those grooves and they associate by forming an extended structure on the edge of the β -sheet, and making it larger by adding an additional β -strand (Figure 5.15). Interactions take place through hydrophobic contacts and hydrogen bonding, and this determines the specificity of the association. (Merino-Gracia, 2016; Williams et al., 2007; Williams, Xie, & Hendrickson, 2005)

The interaction between DYNLT1 and DIC is conditioned by the presence of several hydrophobic positions in the DIC sequence. The β -strand formed by the DIC chain upon the interaction with DYNLT1 contains a small distortion (β -bulge) which alters the hydrogen-bonding pattern and avoids the extension of the β -sheet by the addition of new strands, blocking the aggregation. (Hall et al., 2009; Merino-Gracia, 2016; Williams et al., 2007)

4.2 FUNCTIONS OF DYNLT PROTEINS

During the cell interphase, both DYNLT1 and DYNLT3 can be found dispersed in the cytosol, and tend to accumulate in a juxtanuclear localization, corresponding to the centrosome and Golgi apparatus. They can also be forming a perinuclear ring or, in some cells, they can be highly concentrated in the nucleus. As the cell cycle advances, DYNLT proteins redistribute to specific cell compartments. For instance, they can be found as components of the mitotic spindle where they can participate in its assembly and orientation. DYNLT1 and DYNLT3 present in the cleavage furrow or the kinetochores help to the correct alignment of chromosomes during cell division. Inside the nucleus, DYNLT3 interacts with SATB1, play an important role in the gene expression regulation, mitosis, and meiosis. (Campbell, Cooper, Dessing, Yates, & Buder, 1998; J.-r. Huang, Craggs, Christodoulou, & Jackson, 2007; X. Huang et al., 2011; Liu, Chuang, Sung, & Mao, 2015; Lo,

Kogoy, & Pfister, 2007; Merino-Gracia, 2016; Sarma & Yaseen, 2013; Tai, Chuang, & Sung, 1998; Yeh, Chuang, & Sung, 2005)

DLCs are well known to function as adaptors which take part in the anchoring of cargoes to dynein motors. It has been proposed that some kinds of viruses are able to associate to DYNLT proteins in order to exploit the cell motor machinery to get to the cell nucleus. (Merino-Gracia, 2016; Merino-Gracia, García-Mayoral, & Rodríguez-Crespo, 2011; Tai, Chuang, Bode, Wolfrum, & Sung, 1999)

Microtubule network is crucial for the internal organization of neurons, and various motor-dependent and motor-independent functions have been associated to DYNLT1 in these cells. DYNLT1 is involved in the GPCR-mediated signalling, through its interaction with G β subunits, which also participate in the mitotic spindle orientation and in dynein-independent neurogenesis. DYNLT1 can be related to neuronal differentiation regulation processes by interacting with neurotrophin receptors Trk. This DIC is also involved in the elongation of neurites and growth of neuronal prolongations. (Conde et al., 2010; Chuang et al., 2005; Merino-Gracia, 2016; Sachdev et al., 2007)

4.3 LIGAND BINDING TO DYNLT

Several works have demonstrated that DYNLT proteins can bind to a very heterogeneous set of ligands. These ligands are also proteins which are implied in numerous and diverse cellular processes, and they do not share any homology. Many attempts to determine a consensus sequence for the binding to DYNLT1 have been carried out. At present, only structures of DYNLT1 interacting with DIC have been published, as DIC is considered the canonical ligand. DIC segment inserts as an extended strand in the hydrophobic groove, prolonging the β -sheet and interacting directly with the swapped β -strand (β 2) (Figure 5.16). Interestingly, the DIC β chain shows a distortion due to de modification of the normal hydrogen bonding pattern. (Hall et al., 2009; Merino-Gracia, 2016; Williams et al., 2007)

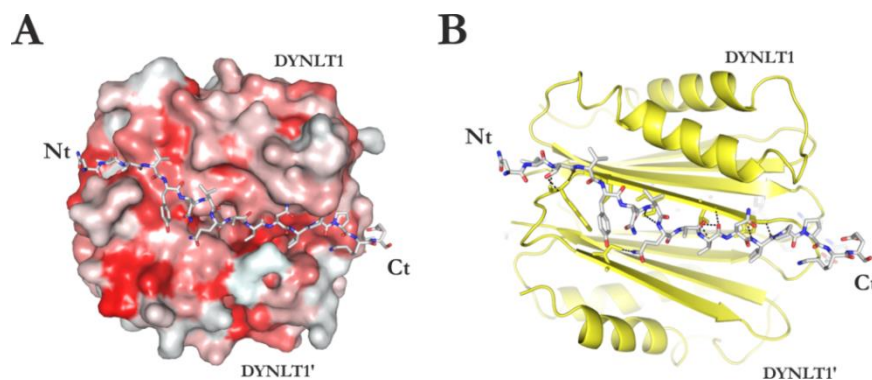


Figure 5.16. Interaction of DYNLT1 with DIC. A. Surface representation of DYNLT1 dimer coloured in red scale by hydrophobicity. DIC interacting portion, which adopts an extended, distorted β -strand conformation, is shown in sticks. Interaction takes place in the hydrophobic groove. B. Cartoon representation showing the contacts between the DIC interacting segment and DYNLT1. The prolongation of the β -sheet can be observed. Structure corresponds to DYNLT1 and DIC from *D. melanogaster* (PDB ID: 3FM7). (Hall et al., 2009)

In this thesis, some proteins capable of interacting with DYNLT1 were studied. These proteins can be divided into two types:

Small molecular GTPases

GTPases are molecular switches involved in cell signalling. These enzymes possess an active form, with bound GTP, and an inactive form, when GTP is hydrolysed and they keep the association with GDP. Their intrinsic hydrolytic ability is very slow, and therefore they need the participation of a GTPase activating protein (GAP), which triggers the hydrolysis of GTP to GDP, inactivating the enzyme and closing the pathway. To reactivate the GTPase, GDP must be replaced with GTP, a process facilitated by guanine nucleotide exchange factors (GEFs). There are two small GTPases which are known to interact with DYNLT1: RagA (also called FIP-1), and Rab3D. (Lukashok, Tarassishin, Li, & Horwitz, 2000; Merino-Gracia, 2016; Pavlos et al., 2011)

- RagA

Rag proteins are dimeric GTPases. In humans, there are four types of Rag proteins described and they form heterodimers. They are involved in the control of cell growth mediated by mTORC1/2 complexes and regulated by amino acids availability. It shares the usual structural features of its family, but possesses an additional dimerization domain, and lacks lipidization motifs to attach to the membrane. Interaction between RagA and DYNLT1 was described in the context of the infection with adenovirus, since RagA acts as a connector among viral proteins and DYNLT1 forming a ternary complex which can serve as a linker between the virus and the transport cell machinery. (Lukashok et al., 2000; Merino-Gracia, 2016)

- Rab3D

Rab3D is a protein belonging to the Rab GTPase family, which are involved in vesicle traffic through the vesicle positioning and membrane fusion. Rab3 proteins participate in the exocytotic pathway in cells with a high excretion rate, and particularly, Rab3D plays a role in the regulation of some steps after the *trans* Golgi network (TGN) in non-neuronal cells. Rab3D GTPase has a typical structure with an isoprenylation motif in the C-terminal region; switch I and II regions, which change their conformation depending on the nucleotide site occupation; and particular sequences for its specific effects. Rab3D is active in its GTP form and inactive in the GDP form. When inactive, this protein is associated to a membrane, where it is activated by a GEF. Upon the activation, vesicle formation is triggered and it is transported to the target compartment. There, a GAP stimulates the catalytic activity of Rab3D, which leads to its inactivation. In the inactive state, a dissociation inhibitor (GDI) dissociates the Rab3D from the membrane and keeps it in the cytosol, waiting for a new cycle. Interaction between Rab3D and DYNLT1 was described in osteoclastic cells, where Rab3D participates in the vesicle traffic between the Golgi apparatus and the ruffled border membrane. This

interaction has not known function, but it is suspected to help with the dynein recruitment for vesicle transport in bone resorption. (Merino-Gracia, 2016; Millar, Pavlos, Xu, & Zheng, 2002; Pavlos et al., 2011)

Lfc

Lfc protein is a GEF of the Rho family of GTPases, participating in the control of cytoskeleton reassembly for different processes, such as cell mobility and adhesion, or vesicle traffic. Lfc has several domains, but lacks the characteristic RH (regulator of G protein signalling homology) domain. DH (Dbl homology domain) and PH (pleckstrin homology) domains constitute the catalytic core, which binds inactive GTPases to favour the GDP/GTP exchange, and trigger their activation. This way, actin cytoskeleton is stimulated in response to microtubule depolymerisation. Interaction between Lfc and DYNLT1 was reported as the mechanism that drives the coupling of Lfc and the microtubules, and their subsequent inhibition, demonstrating that Lfc stays associated to microtubules in a dynein-dependent way through the formation of a ternary complex among Lfc, DYNLT1, and DIC. External signals, such as lysophosphatidic acid (LPA), through its GPCR (G protein-coupled receptor), induce an intracellular signalling cascade by means of the coupled G proteins which can activate Lfc. G α subunits can interact with Lfc by its N-terminal region, probably competing against DYNLT1, and release Lfc from the microtubular complex. On the other hand, G $\beta\gamma$ subunits can compete against DYNLT1 in its binding to DIC, leading to the same releasing effect. (Guilluy, Garcia-Mata, & Burrridge, 2011; Meiri et al., 2012; Meiri et al., 2014; Merino-Gracia, 2016; Sachdev et al., 2007)

OBJECTIVES

In this chapter, the specificity of the DYNLT1/Tctex1 canonical binding groove has been explored. DYNLT1/Tctex1 is a dynein light chain which has not only been described as a part of the scaffold of the dynein complex, but also as a molecular clamp to stabilize dimeric structures in the cytosol. DYNLT1/Tctex1 dimer is known to interact with other proteins through a groove located in the monomer/monomer interface, which is considered the canonical binding site. However, ligand binding through a secondary interaction site has been reported.

In this work, a thorough study of the interactions between DYNLT1/Tctex1 and some known ligands has been done, and the characterization of the secondary binding site of the dimer has been addressed by biochemical and biophysical approaches, including solution NMR.

The specific objectives proposed in this chapter are:

- To obtain the first refined 3D structure of a mammalian DYNLT1 saturated with a DIC-derived peptide, by using multinuclear solution NMR.
- To characterise the details of the interaction between DYNLT1 and DIC, and identify sequential or structural patterns to define its interaction with ligands.
- To investigate the interaction between DYNLT1 and other different binding partners.

MATERIALS AND METHODS

1. MATERIALS

¹⁵N-labelled (NH₄)Cl and ¹³C-labelled glucose were purchased from Cambridge Isotope Laboratories. Buffers, chemicals, oligonucleotides and common laboratory reagents were bought from Sigma–Aldrich, if not otherwise indicated. Pfu polymerase, T4 DNA ligase, restriction endonucleases and molecular mass markers were obtained from Fermentas. Synthetic peptides were ordered from Thermo Scientific and were at least 90 % pure. Sepharose 4B was from Pharmacia Fine Chemicals. D(+)-lactose monohydrate was from Scharlau. Glutamine, antibiotics, cell culture medium (Dulbecco’s modified Eagle’s medium) and X–Gal were purchased from Sigma Aldrich. Trypsin–EDTA and foetal bovine serum were from BioWhittaker Europe.

2. CONSTRUCTS

Mouse DYNLT1 cDNA (UniProtKB P51807), kindly provided by Dr. Thomas Sakmar (The Rockefeller University, New York) (Sachdev et al., 2007), was amplified and cloned into the yeast two–hybrid vectors pGAD and pGBT9 as well as in the recombinant expression vector pKLSLt (Angulo et al., 2011). Mammalian expression vectors were also created using the LSLt lectin (Angulo et al., 2011) in frame with DYNLT1 or with DYNLT1–DIC2 (Merino-Gracia et al., 2015). The human DYNLT1 cDNA (UniProtKB P63172) was a generous gift from Dr. Laszlo Nyitray (Eötvös Loránd University, Budapest), provided in the recombinant expression vector pBH4 with a His₆–tag and a TEV protease cleavage sequence. The full–length mouse dynein intermediate chain DYNC1I2 (UniProtKB O88485) was provided by Dr. Mingjie Zhang (The Hong Kong University of Science and Technology) (Fan, Zhang, Tochio, Li, & Zhang, 2001). The sequence of DIC known to bind to DYNLT1 and DYNLL1 (amino acids 130–160) was amplified and cloned in the yeast two–hybrid and recombinant expression vectors, as described above. The self–saturated DYNLT1–DIC2 chimeric protein was created by fusing the DYNLT1 binding region of the human DIC2 (UniProtKB Q13409) to the C–terminus of human DYNLT1 in frame after a short linker. The sequence of the resulting fusion protein is shown in page 299 (Figure 5.22) of the “Results and discussion” section. DYNLT1–DIC2 was cloned into the pBH4 vector for recombinant expression (Merino-Gracia et al., 2015). Mouse Activin Receptor IIB (UniProtKB P27040–2) and p3TP–lux (a reporter of TGFβ signalling) plasmids were provided by Dr. Carmelo Bernabeu (Centro de Investigaciones Biológicas, Madrid). The C–terminal deletion construct of ActRIIB was obtained by PCR and included residues 1–490, whereas full–length wild–type ActRIIB was 512 amino acids long. GFP–tagged constructs of both full–length and ΔCt ActRIIB were also obtained using a commercial pEGFP–N vector. Full–length canine Lfc cDNA (UniProtKB Q865S3) was a kind contribution of Dr. Karl Matter (Univ. College London, London).

3. YEAST TWO-HYBRID SCREEN AND β -GALACTOSIDASE ASSAY

Plasmids containing GAL4 binding domain were used, and they were confronted with plasmids containing the GAL4 activation domain as previously described (Navarro-Lérida, Martínez-Moreno, Ventoso, Álvarez-Barrientos, & Rodríguez-Crespo, 2007; Sánchez-Ruiloba et al., 2014). Double transformants were plated in Leu⁻/Trp⁻/His⁻ synthetic defined plates in the presence of 12 mM 3-amino triazole (triple dropout plates) as well as in Leu⁻/Trp⁻/His⁺. Interacting proteins expressed within the same yeast resulted in colonies that could rescue growth in the absence of His. These colonies were subsequently screened in the X-Gal assay. Blue colonies corresponded to a positive interaction whereas white colonies corresponded to absence of interaction.

4. RECOMBINANT PROTEIN EXPRESSION AND PURIFICATION

The pBH4-DYNLT1 and pBH4-DYNLT1-DIC2 plasmids were used to transform BL21 DE3 *Escherichia coli*. 2 l of bacterial culture in 2X yeast extract-tryptone were routinely used for recombinant expression. When the protein was used for NMR experiments, the bacteria transformed with pBH4-DYNLT1 were grown in M9 minimal medium supplemented with ¹⁵N-labelled (NH₄)Cl or with ¹⁵N-labelled (NH₄)Cl plus ¹³C-labelled glucose. Protein expression was induced by addition of 0.5 mM IPTG and incubation overnight at 20 °C with 150 rpm aeration rate. Bacterial cells were pelleted and frozen at -20 °C until used for protein purification.

Briefly, the bacterial cell lysis was performed in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 3 mM NaN₃, pH 8) on ice, with continuous stirring in the presence of protease inhibitors (1 µg·mL⁻¹ aprotinin, 1 µg·mL⁻¹ leupeptin, and 200 µM PMSF), 200 mg·L⁻¹ lysozyme, and 5 mM β -mercaptoethanol), followed by 4 cycles of sonication on ice. The cell lysate was clarified by centrifugation at 10,000 g and filtered through porous paper. The recombinant 6xHis-tagged proteins were sequentially purified using a Ni-NTA affinity column, followed by a digestion with TEV protease and a final separation using a HiTRAP Q HP anionic exchange column. The protein purity was confirmed by SDS-PAGE and Coomassie blue staining. Pure protein solutions were concentrated by centrifugation in Centricon® tubes (Millipore) to a final concentration of ~1mM. When DYNLT1 was expressed in mammalian cells, vector pKLSLt-DYNLT1 was used to transform COS7 cells. Protein purification was performed 36 h post-transfection using the protocol described above. DYNLT1 expressed in mammalian cells was used in the pepscan assays.

5. ISOTHERMAL TITRATION CALORIMETRY (ITC)

Interaction between DYNLT1 and various peptides was measured using a VP-ITC MicroCalorimeter (MicroCal, Northampton, MA) in 20 mM Hepes buffer, pH 7.4, containing 0.15 M NaCl at 25 °C as previously reported (Merino-Gracia, Costas-Insua, Canales, & Rodríguez-Crespo, 2016). Commercial peptides were dissolved in water at 2 mM concentration and were subsequently diluted to 0.5–1 mM in the aforementioned buffer

when put in the syringe. Protein solutions were introduced into the sample cells, whereas peptides were put in the syringe. All samples were degassed for at least 5 min in a ThermoVac (MicroCal). Control experiments consisting in the titration of peptide and DYNLT1 with buffer were performed. Each experiment started with an initial injection of 2.5 μL , followed by thirty 7.5 μL injections. The heat released in each injection was calculated from the raw data by integration of the peaks after subtraction of the baseline. All data were analysed using the Origin®5 program.

6. CELL TRANSFECTION AND IMMUNOFLUORESCENCE

Procedures described previously in other works were followed (Navarro-Lérida, Álvarez-Barrientos, & Rodríguez-Crespo, 2006; Navarro-Lérida et al., 2007). COS7 cells were from the European Collection of Cell Cultures.

7. FLUORESCENCE POLARIZATION ASSAYS

FP was performed in a Perkin Elmer MPF 44-E spectrofluorimeter. Saturation binding experiments were done for measuring binding affinity (K_d) between FITC-labelled peptides and DYNLT1 by applying an increasing amount of recombinant protein (typically 0–150 μM) to a fixed, low concentration of probe (5 to 100 nM). Incubation time was 10–15 minutes (at room temperature), and the assay was carried out in 20 mM Hepes, 150 mM NaCl, pH 7.4, in a final volume of 0.5 mL. Polarization of the FITC-labelled peptides was measured at excitation/emission wavelengths of 488/530 nm (bandwidth = 10 nm). The fluorescence anisotropy (r) values were obtained using the fluorescence polarization (FP) values with the equation:

$$r = \frac{2 \cdot FP}{3 - FP} \quad \text{Eq. 5.1}$$

The initial anisotropy (r_0) in the absence of added protein was measured. The fluorescence polarization (FP) values were fitted to the equation:

$$(FP - FP_0) = \frac{(FP_{max} - FP_0) \cdot [PDZ \text{ domain}]}{K_d + [PDZ \text{ domain}]} \quad \text{Eq. 5.2}$$

With FP being the measured fluorescence polarization, FP_{max} the maximal fluorescence polarization value, FP_0 the fluorescence polarization in the absence of added PDZ domain, and K_d the dissociation constant. As long as the concentration of labelled peptide is well below the true K_d during the assay, the K_d can be directly derived from this saturation curve.

8. NMR EXPERIMENTS

Pure unlabelled, ^{15}N -labelled, and doubly ^{13}C , ^{15}N -labelled human DYNLT-DIC construct in the range of 50–200 μM was prepared in water with 10 % of D_2O in 100 mM

KH₂PO₄ buffer, 1 mM DTT, pH 7.0. NMR samples contained 50 μ M DSS as an internal reference.

NMR spectra were acquired at 298 K on a Bruker AV 800 NMR spectrometer (Bruker, Rheinstetten, Germany) equipped with a triple-resonance z-gradient cryoprobe. ¹⁵N and ¹³C chemical shifts were referenced indirectly using the gyromagnetic ratios of ¹⁵N:¹H and ¹³C:¹H. (Wishart et al., 1995) The following standard pulse sequences, corresponding to the heteronuclear multidimensional approach for NMR assignments and structure calculation (Whitehead, Craven, & Waltho, 1997) were acquired and analysed: 2D ¹H–¹H NOESY (80 ms), ¹⁵N–HSQC and ¹³C–HSQC; and 3D spectra CBCA(CO)NH, CBCANH, HNCA, HN(CO)CA, HC(C)H–TOCSY, (H)CCH–TOCSY, ¹H–¹⁵N HSQC–NOESY (50 ms), ¹H–¹³C HSQC–NOESY (50 ms, aliphatic region), and ¹H–¹³C HSQC–NOESY (50 ms, aromatic region).

Spectra were processed with Bruker Topspin 2.1 (Bruker, Germany) and spectral analysis was performed with SPARKY3. (Goddard & Kneller, 2004) Backbone and side chain ¹H, ¹³C and ¹⁵N chemical shifts were assigned following conventional strategies. The resonance list is practically complete and has been deposited in the BioMagResBank under the code number 30074.

9. STRUCTURE CALCULATION

Structure calculation of the DYNLT–DIC construct was performed with CYANA (Güntert, 2004) using the symmetric homodimer protocol and the automatic NOE assignment facility combined with lists of manually assigned NOEs. In total, there were 3247 upper distance constraints, 143 of which were manually assigned. 514 backbone dihedral angle constraints were determined from chemical shift values using TALOS+ (Shen, Delaglio, Cornilescu, & Bax, 2009). Initially, 100 conformers were generated that were forced to satisfy experimental data using a standard automatic CYANA protocol. The 20 conformers with the lowest final CYANA target function values were selected for further refinement and, finally, minimized with Amber9 software (Case et al., 2005) using the Gibbs–Boltzmann continuum solvation model. Final structure quality was checked with PROCHECK–NMR (Laskowski, Rullmann, MacArthur, Kaptein, & Thornton, 1996) and the coordinates have been deposited in the Protein Data Bank under the accession number 5JPW. MOLMOL (Koradi, Billeter, & Wüthrich, 1996) and PyMOL (Schrödinger, LLC) were used for molecular analysis and display.

10. NMR DYNAMICS

¹H–¹⁵N heteronuclear NOE data were determined on the bases of conventional NOE measurements with a ¹⁵N-labelled sample. Experiments with and without proton saturation were acquired simultaneously in an interleaved manner with a recycling delay of 10 s, and were split during processing into separate spectra for analysis. The values for the heteronuclear NOEs were obtained from the ratio of the resonance intensities, according to:

I_{sat}/I_{ref} . Intensities and peak volumes were determined by using tools included in SPARKY3. (Goddard & Kneller, 2004) The uncertainty was estimated to be about 5 %.

11. CIRCULAR DICHROISM MEASUREMENTS

CD spectra were recorded on a Jasco J-715 spectropolarimeter using a 0.1-cm path length cell for the far UV measurements at 25 °C as previously reported. (Cook, Galve-Roperh, del Pozo, & Rodríguez-Crespo, 2002) Data analysis was performed using Microcal Origin software (Originlab, Northampton, MA, USA).

12. LUCIFERASE REPORTER ASSAY

Luciferase activity was measured using the Dual-Glo® Luciferase Assay System kit (Promega, Madison, WI) as specified by the manufacturer, in an EG&G Berthold Lumat LB 9507 luminometer. Briefly, HEK cells were transfected with the p3TP-lux plasmid in the presence or absence of full-length ActRIIB or an ActRIIB lacking the DYNLT1 binding site present at its cytoplasmic C-terminus (referred to as ActRIIB-ΔCt). Subsequently, cells were washed once with phosphate-buffered saline. After addition of 200 μL of lysis buffer, cells were scraped and centrifuged (4 °C, 12,000 g, 5 min). Measurement was carried out in 1:1 dilutions of the cell extract with the Dual-Glo® Luciferase Reagent, followed by an incubation of 10 min, within 2 hours. Each sample was measured for 20 s. All assays were performed in duplicate.

13. CELL TRANSFECTION AND PULLDOWN ASSAYS

Procedures described previously by the group of Dr. Rodríguez-Crespo were followed. (Merino-Gracia et al., 2016; Merino-Gracia et al., 2015) The association between LSLt-DYNLT1 and either GFP-tagged full-length ActRIIB-GFP (residues 1–512) or GFP-tagged ActRIIB-ΔCt (residues 1–490) in transfected COS7 cells was analysed. The pulldown assays were performed using Sepharose 4B in the absence of added antibodies. The cell extract was incubated with protease inhibitors plus alkaline phosphatase, clarified, and the RIPA-solubilized proteins were allowed to bind to Sepharose 4B beads due to the presence of the LSLt lectin. The beads were subsequently extensively washed to avoid nonspecific interactions of the cellular extract and processed as in a conventional immunoprecipitation.

14. PEPSCAN EXPERIMENTS

Automated peptide spot synthesis attached to a cellulose membrane (Abimed, Langenfeld, Germany) and incubation with the desired recombinant protein was done using the pepscan methodology as reported previously (Martínez-Moreno et al., 2003; Rodríguez-Crespo et al., 2001). The membrane was obtained with the synthesized pentadecapeptides immobilized by their C-termini via a polyethylene glycol spacer and N-terminal acetylated.

The cellulose membranes were coated with 1 % non-fat dried milk in TBS (50 mM Tris, pH 7.0, 137 mM NaCl, 2.7 mM KCl) for 4 h at room temperature. Incubation with the recombinant DYNLT1 (0.1 μ M) was done overnight at room temperature. Three washes (25 mL each) were performed with TBS/Tween-20 (0.05 %). Subsequently, the membrane was incubated for 2 h at room temperature with an antibody against DYNLT1. Three additional 10 min washes were performed with TBS/Tween-20 (0.05 %), followed by three more 10 min washes with TBS alone. Development of the membrane was performed by ECL following the instructions of the manufacturer. The intensity of each spot was subsequently quantified using the Fuji software.

RESULTS AND DISCUSSION

1. THE POLYBASIC SEQUENCE MOTIFS OF MAMMALIAN DYNEIN INTERMEDIATE CHAINS ARE NOT REQUIRED FOR DYNLT1 BINDING

In mammals, there are two intermediate chains per dynein complex with a molecular mass of ~74,000 Da, known as DYNC1IC1 and DYNC1IC2 (K. Pfister, Salata, Dillman, Torre, & Lye, 1996; K. K. Pfister et al., 2005; K. K. Pfister et al., 1996) (for simplicity's sake, they will be referred as DIC1 and DIC2). Alternative splicing and phosphorylation produce multiple intermediate chain isoforms (K. Pfister et al., 1996; K. K. Pfister et al., 1996). The binding sites for DYNLT1 and DYNLL1 light chains were found to be just C-terminal of the second alternative splice region in the intermediate chain gene (Hall et al., 2009; Mok, Lo, & Zhang, 2001; Williams et al., 2007). In addition, the binding sites for both light chains on the intermediate chain were found to be sequential (Figure 5.17).

		polybasic motif	DYNLT1 binding	DYNLL1 binding	
<i>Homo sapiens (DYNC1IC1)</i>	143	RRLHKL	GVSKVTQVDFLPREVVSYSKETQTPLATHQ	178	
<i>Homo sapiens (DYNC1IC2)</i>	133	RGPIKL	GMAKITQVDFPPREIVTYTKEQTTPVMAQP	168	
<i>Rattus norvegicus (DYNC1IC1)</i>	143	RRLNKL	GVSKVTQVDFLPREVVSYSKETQTPLATHQ	178	
<i>Rattus norvegicus (DYNC1IC2)</i>	133	RGPIKL	GMAKITQVDFPPREIVTYTKEQTTPVTAQP	168	
<i>Gallus gallus</i>	143	RRLHKL	GVSKITQVDFLPREVVSYSKETQTPLATHQ	178	
<i>Danio rerio</i>	126	CRMQRL	GASKIVQIDFMPKELVTYSKETQTTPVASNQ	161	
<i>Xenopus laevis</i>	104	RGPMKL	GMSKVTQVDFAPRETVSYTKEQTTPILAQN	139	
<i>Caenorhabditis elegans</i>	109	LLSVDLEFSSVQTACTDNKDSAIYSKGTQTDDERIS	144		
<i>Drosophila melanogaster</i>	105	KQPLNLSVYNVQATNIPPKETLVYTKQTQTSTGGG	140		

Figure 5.17. Sequence alignment of several mammalian DIC isoforms and those of selected model organisms known to bind DYNLT1 and DYNLL1 consecutively. The polybasic motif, and the DYNLT1 and DYNLL1 binding sites are shown on top. The characteristic TQT motif present in many DYNLL1-binding partners is highlighted in yellow.

The first question to be addressed was to find the minimal part of DIC capable of interacting with DYNLT1. Using a yeast two-hybrid assay, it was demonstrated that a short polypeptide of either isoform 1 or isoform 2 lacking part of the polybasic stretch was able to interact with DYNLT1 (Figure 5.18).

	pGAD	pGBT9	His ⁻	X-Gal
(a)	DYNLT1	DIC1(1-145)	—	—
(b)	DYNLT1	DIC1(1-180)	++	++
(c)	DYNLT1	DIC1(145-165)	++	++
(d)	DYNLT1	DIC2(135-155)	++	++

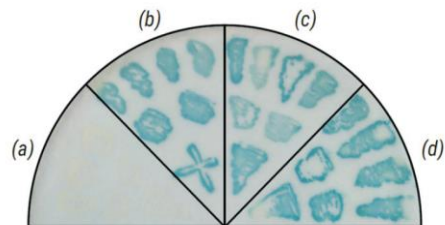


Figure 5.18. Yeast-two hybrid assay using DYNLT1 in the bait plasmid and various dynein intermediate chain constructs. Growth in the absence of histidine, in the presence of 3-amino triazole, or X-Gal activity is shown. The right panel shows a representative result of an X-Gal assay in a plate.

Then, recombinant DYNLT1 was expressed and purified, as well as a self-saturated DYNLT1 with DIC, DYNLT1–DIC2. Their fold was analysed using circular dichroism (Figure 5.19). Deconvolution of CD data revealed that both DYNLT1 in complex with a DIC1 peptide (KLGVSQVTQVDFLPREV), as well as its self-saturated DYNLT1 counterpart showed very similar spectra, with calculated secondary structure elements of 25.2 % α -helix, 10.8 % antiparallel β -strand, 10.9 % parallel β -strand, 18.5 % β -turns, 34.6 % random coil for the former, and 24.9 % α -helix, 10.9 % antiparallel β -strand, 11.9 % parallel β -strand, 18.5 % β -turns, 33.8 % random coil for the latter (Figure 5.19). DYNLT1 CD spectrum is in agreement with previously reported data (Makokha et al., 2002) and with the atomic information available.

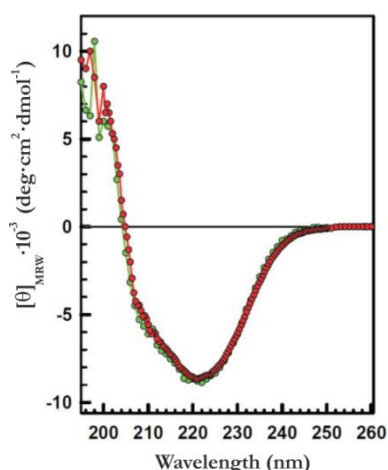


Figure 5.19. Far-UV CD spectra of purified recombinant DYNLT1 saturated with a DIC1 peptide (KLGVSQVTQVDFLPREV) (in red) and its self-saturated DYNLT1–DIC2 counterpart (shown in green).

A thermodynamic analysis of the interaction of DYNLT1 with a DIC1 peptide was performed using ITC resulting in a K_d of $6 \pm 1 \mu\text{M}$ (Figure 5.20). To further analyse in detail the binding of this DIC1 short peptide to DYNLT1 fluorescence polarization measurements were realized. Saturation curves were then created by applying increasing concentrations of the purified DYNLT1 domain and fixed concentration of the probe. The K_d value between probe and DYNLT1 was determined to be $4.8 \pm 0.5 \mu\text{M}$ (Figure 5.21). Taken together, these results indicate that a short peptide corresponding to residues 147–162 of dynein intermediate chain binds to DYNLT1 with high affinity, and that the polybasic stretch that precedes this segment is dispensable for binding. This finding contrasts with previous reports that indicated that the interaction required this polybasic region (Mok et al., 2001).

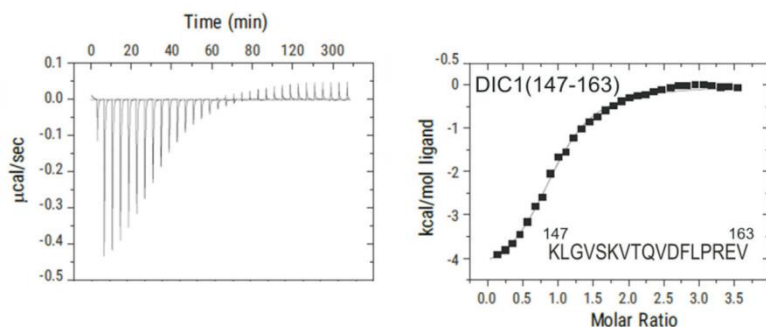


Figure 5.20. ITC analysis of the binding of a dynein intermediate chain peptide (KLGVSQVTQVDFLPREV) to purified DYNLT1. Thermogram is shown in the left panel, whereas the binding isotherm is shown in the right panel. Curve fitting rendered a value of $K_d = 6 \pm 1 \mu\text{M}$.

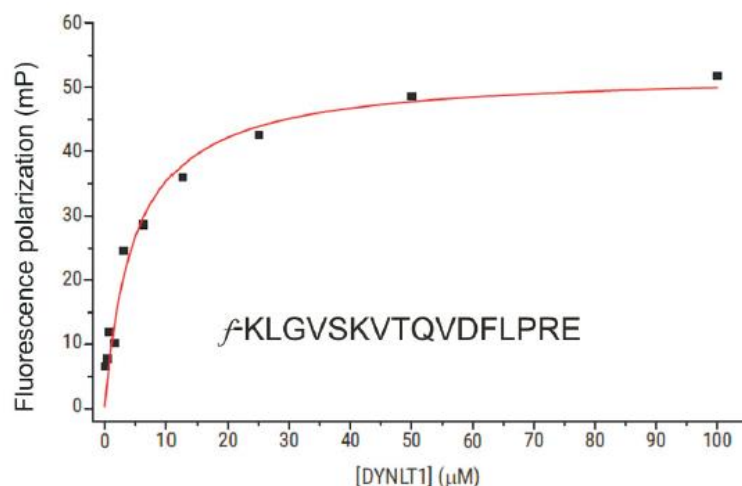


Figure 5.21. Representative binding curve for a FITC-labelled dynein intermediate chain peptide (f-KLGVSKVTQVDFLPRE, *f* stands for fluorescent) to DYNLT1, measured by fluorescence polarization. A calculated $K_d = 4.8 \pm 0.5 \mu\text{M}$ could be obtained. Results are representative of three independent experiments.

2. NMR SOLUTION STRUCTURE OF HUMAN DYNLT1 IN COMPLEX WITH DIC

To obtain three-dimensional information about the interaction of a human dynein intermediate chain segment to human DYNLT1, a recombinant DYNLT1 self-saturated construct with a dynein intermediate chain sequence was expressed in bacteria (Figure 5.22). This construction was designed to infer structural detailed data about the binding mode of interacting peptides that might occupy the canonical groove.

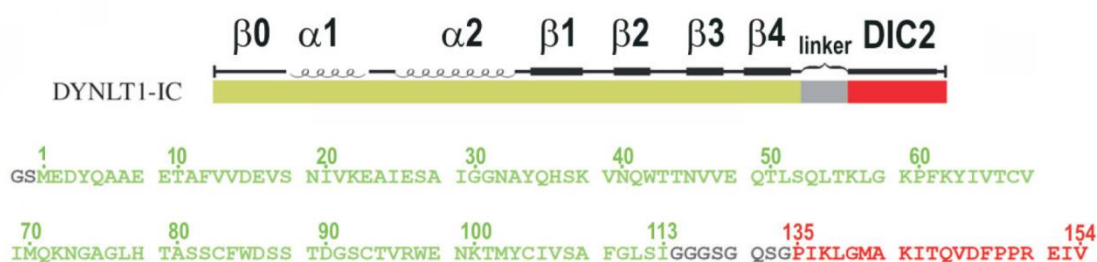


Figure 5.22. Molecular features of the DYNLT1–DIC2 chimera. A scheme of the designed chimeric construct in which a linker (coloured in grey) connects the C-terminus of DYNLT1 (coloured in green) with the N-terminus of DIC (coloured in red). Secondary structure elements are indicated, and amino acid sequence is shown. Numbering corresponds to Uniprot P63172 (full-length human DYNLT1, amino acids 1 to 113), and Uniprot Q13409 (human DIC2, amino acids 135 to 154). The numbering showed here will be used throughout the text.

This DYNLT1–DIC construct was previously characterised in yeast two hybrid and pull-down assays and its behaviour was proved to be similar to that of DYNLT1 when associated to dynein intermediate chain (Merino-Gracia et al., 2015). A DIC2 stretch fused to DYNLT1 was used, since it resulted in much better protein yield than a DIC1 construct

(data not shown). First, a test was performed to check whether the added C-terminal DIC2 portion acquired the dynamic regime of the DYNLT1 moiety. For that purpose, ^1H – ^{15}N heteronuclear NOE data were analysed (Figure 5.23). In fact, the DIC2 GMAKITQVDF region has the same mean NOE value as the core of the DYNLT1 part. This is a clear evidence of the interaction of the DIC peptide with the globular DYNLT1 moiety, as suggested by the conformational chemical shift analysis. On the contrary, residues corresponding to the linker and to the N-terminus of the DIC moiety have low NOE values, indicating that these regions behave as flexible segments.

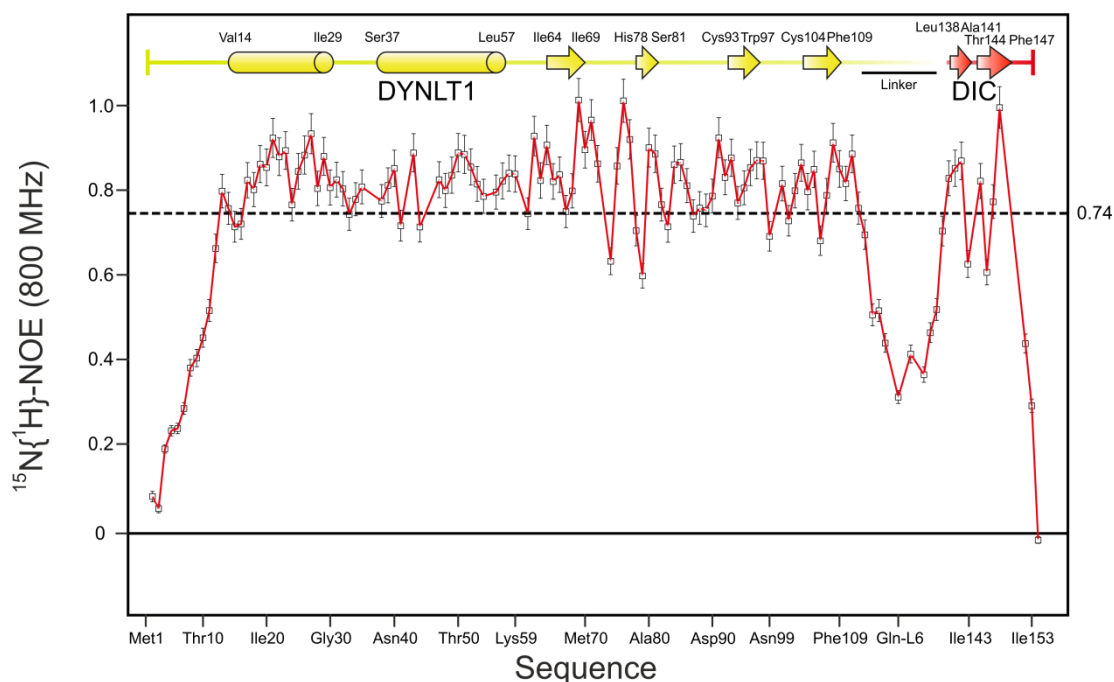


Figure 5.23. ^1H – ^{15}N heteronuclear NOE data of the purified DYNLT1–DIC chimera. Secondary structure elements are indicated on top. Dashed line represents the average value, which is indicated on the right. An error of 5 % is represented with error bars.

Next, NMR was used to confirm the correct folding of the recombinant DYNLT1–DIC2 construct by comparing its 1D ^1H spectrum with that of DYNLT1 in the same conditions (Figure 5.24). In all cases there was good signal dispersion and the conservation of the deshielded signals at about -0.3 and -0.8 ppm (corresponding to the methyl signals of Ile64 and Ile25 of wild-type DYNLT1, respectively), clearly indicates that 3D structure of DYNLT1 is conserved in the construct. NMR spectral assignment was performed by following standard heteronuclear protocols. Secondary structure elements determined from the conformational shifts ($\Delta\delta$) data (Figure 5.25) indicate that they are maintained in the DYNLT1 part of the construct. Low $\Delta\delta$ values observed in the linker section denote a mostly unstructured region, whereas significant $\Delta\delta$ values characteristic of a β -strand are observed in the DIC segment. In addition, the assigned ^1H – ^{15}N -HSQC spectrum of DYNLT1–DIC2 and the amino acid assignment is shown in Figure 5.26.

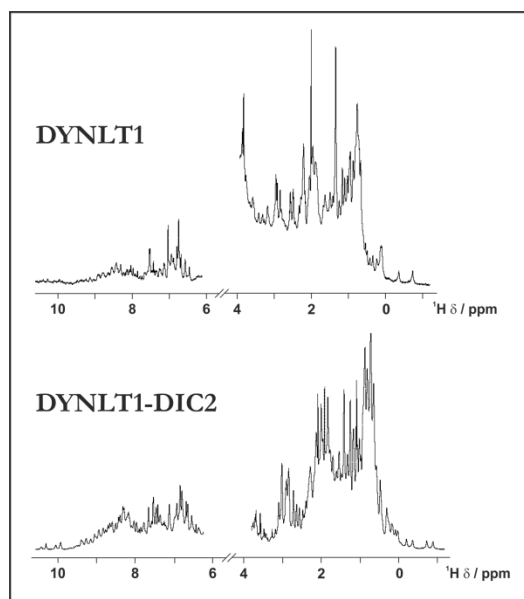


Figure 5.24. 1D ^1H spectra of DYNLT1 and DYNLT1–DIC2 construct. Region corresponding to the water signal (~ 4.7 ppm) has been removed for a better visualisation.

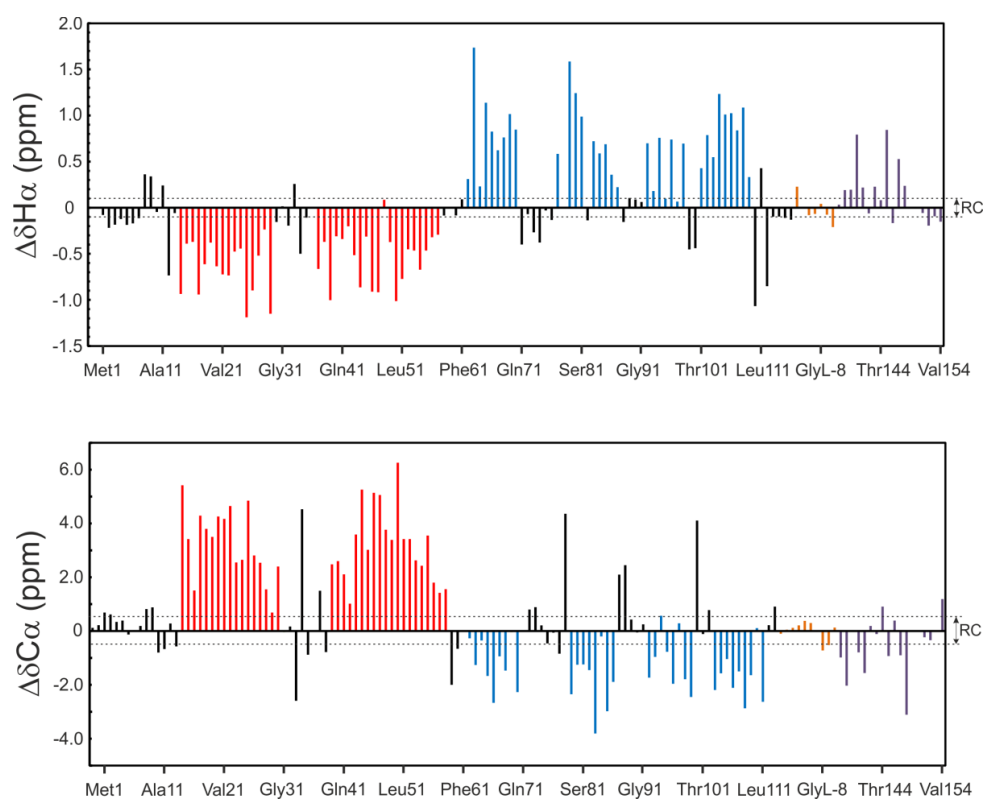


Figure 5.25. H_α and C_α conformational shifts ($\Delta\delta$) plot. Elements of the structure are highlighted in colours (red for α -helices, blue for β -strands, orange for the linker, and purple for the DIC segment). Random coil range is indicated by dashed lines. Values obtained at 25 °C.

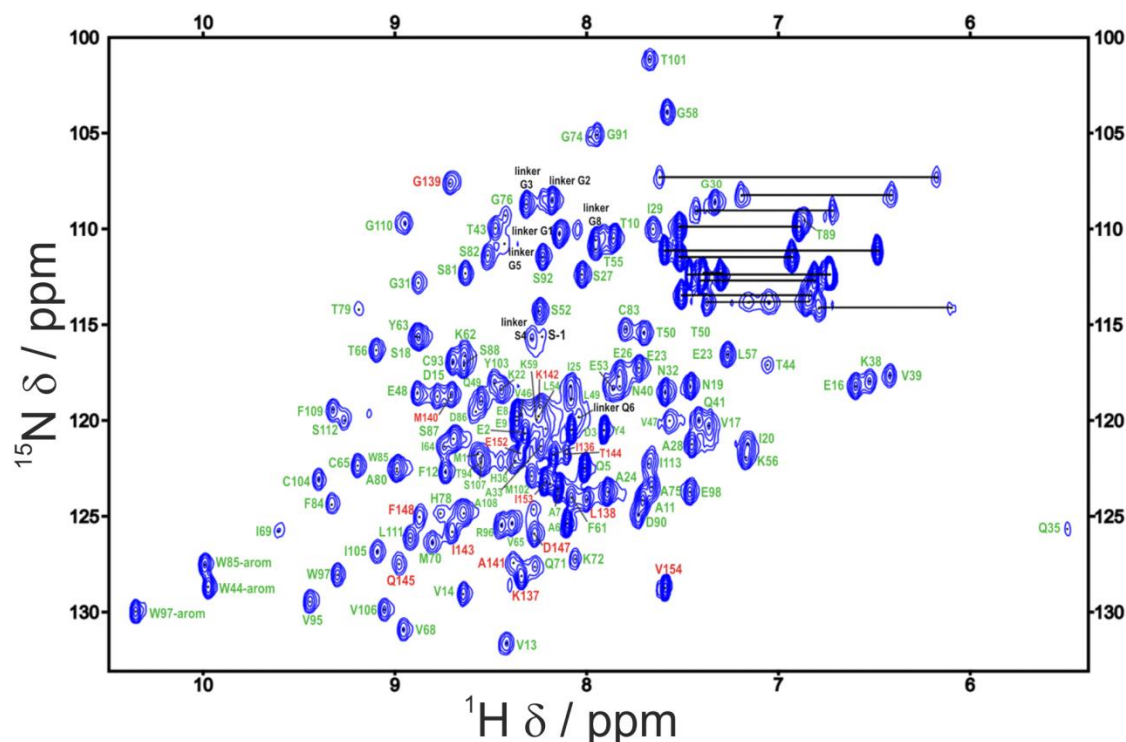


Figure 5.26. ^1H – ^{15}N –HSQC spectrum of DYNLT1–DIC2. Amino acid assignment is shown. Horizontal black bars indicate signals corresponding to geminal protons attached to amino groups of Asn and Gln residues. Labels in green correspond to residues of DYNLT1 sequence, those in red to DIC2 sequence, and in black those corresponding to the linker between both protein sequences.

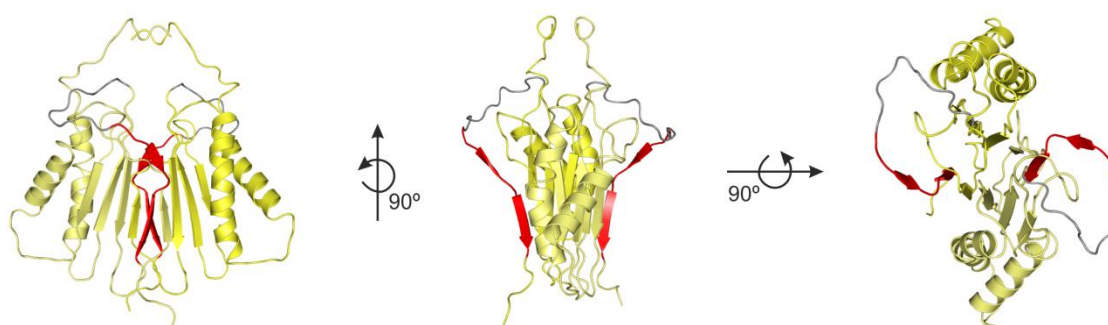
3. 3D STRUCTURE BY SOLUTION NMR: ANALYSIS OF THE DIMER INTERFACE AND CANONICAL BINDING GROOVE

Statistics of the calculated structures are summarized in Table 5.1. A representative conformer of the DYNLT1–DIC2 structure, determined by NMR is shown in Figure 5.27. The structure is a well-defined β –strand swapped dimer and maintains the same tertiary structure shown by other members of the DYNLT family. Each monomer has two α –helices: α 1, 14–29 and α 2, 37–57. The β –strands are defined by residues: β 1, 64–69; β 2, 78–81; β 3, 93–97; β 4, 104–109, β 5; 140–142/144–148 (numbering according to Figure 5.22). The DIC sequences (one of each monomer) lie in the hydrophobic grooves on the structure surface extending the central β –sheet, and are oriented in the same N to C terminus direction. This β –structure is better defined in two fragments of the DIC sequence: Met–Ala–Lys and Thr–Gln–Val–Asp–Phe. Position Gly139 (DIC2 segment) shows φ and ψ values that are far from an ideal β –strand structure and some structural distortion is also observed around Ile143 (DIC2 segment). The formation of a bulge in the DIC stretch was also present in previously reported structures (Balan, 2013; Hall et al., 2009; Williams et al., 2007), in which the total extension of the β –strand was impeded by structural alterations around the central Ile position.

Table 5.1. NMR structural calculations statistics.

Restrains used in the structure calculation			
Total distance restraints			3247
Total angle restraints			514
Total restraints			3761
Number of restraints/residue			11.1
Type of distance and angular restraints			
Short range ($ i-j \leq 1$)			2061
Medium range ($1 < i-j < 5$)			592
Long range ($ i-j \geq 5$)			594
PHI angles			260
PSI angles			254
CYANA statistics (20 structures)			
	Mean	Min.	Max.
Target function (\AA^2)	8.59	3.20	11.24
Maximal distance violation (\AA)	0.29	0.13	1.26
Maximal torsion angle violation ($^\circ$)	-	-	-
AMBER statistics (20 structures)			
Number of violations (<0.4)			3
	Min.	Mean	Max.
Amber energy ($\text{kcal}\cdot\text{mol}^{-1}$)*	-8749	-8854 \pm 52	-8944
Averaged pairwise RMSD (\AA)			
	Backbone	Heavy atoms	
Global dimer	4.00 \pm 1.27	4.53 \pm 1.20	
Secondary structure	0.97 \pm 0.22	1.61 \pm 0.25	
Ramachandran, PROCHECK analysis			
Favorable		81.8 %	
Additional		17.8 %	
Generous		0.5 %	
Non-favorable		0 %	

*Uncertainty is given as the SD.

**Figure 5.27.** Solution structure of human DYNLT1 self-saturated with a DIC polypeptide occupying the canonical binding groove. Three different views of the calculated 3D structure of DYNLT1–DIC in a ribbon representation are shown. DIC segments are highlighted in red. The preformed β -sheets of the monomer–monomer interaction surface become extended by swapped antiparallel β -strands hence creating the binding surface for the incoming, kinked antiparallel β -strand of DIC.

Three hydrophobic contact areas in which the side chains of three DIC2 residues are accommodated have been previously observed in the *D. melanogaster* DYNLT1 crystal structure (Williams et al., 2007). Likewise, analysis of these data indicates that DIC Leu, Met, and Ile do, in fact, interact with DYNLT1 in hydrophobic binding cavities. On the other hand, the more C-terminal DIC2 Val and Phe side chains seem to be in a solvent exposed zone in which only minor contacts with DYNLT1 are expected (only two DYNLT1 hydrophobic residues, one Phe and one Trp, can be found in their vicinity, within a 8–10 Å range) (Figure 5.28).

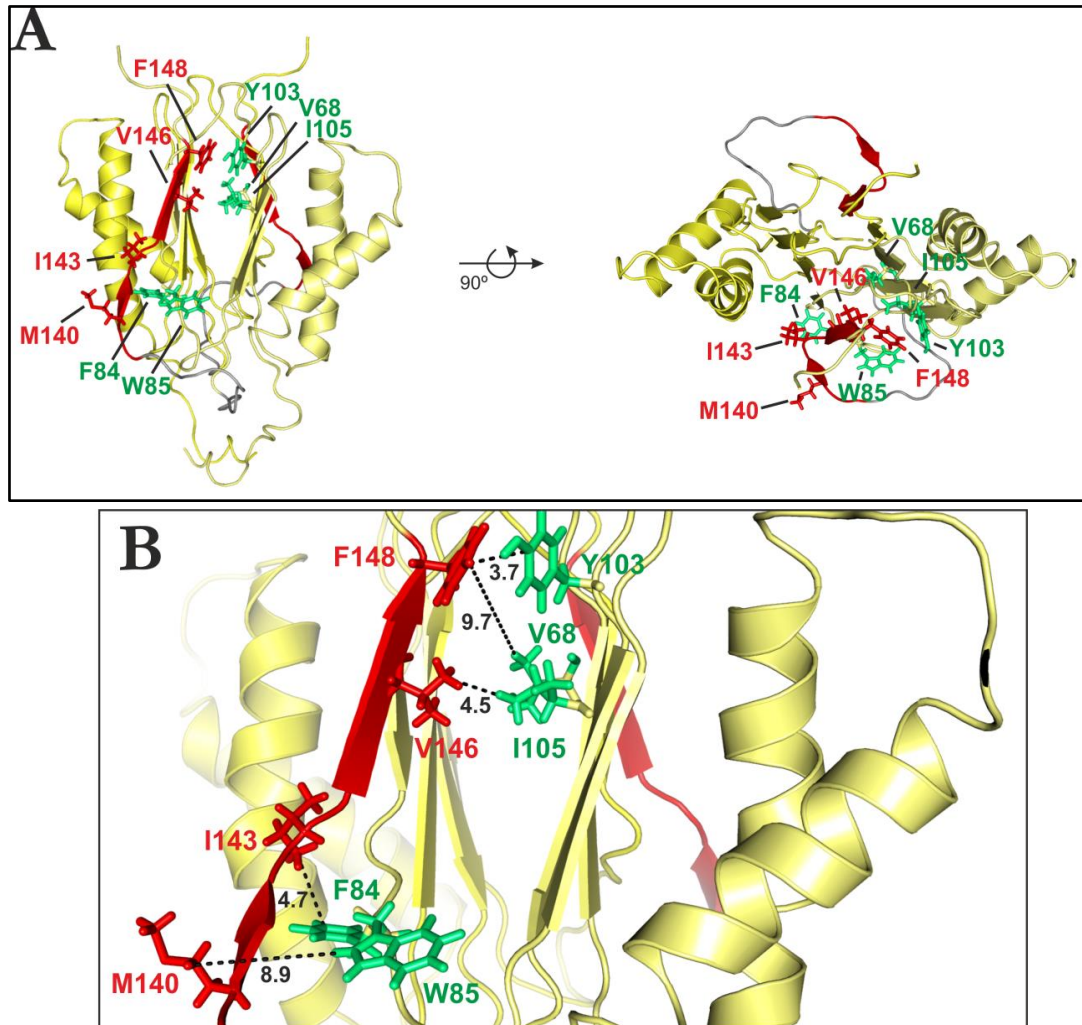


Figure 5.28. Solution structure of human DYNLT1 self-saturated with a DIC2 polypeptide occupying the canonical binding groove. **A.** Two different views rotated 90° of the calculated 3D structure of DYNLT1–DIC2 in a ribbon representation. The two chains of each DYNLT1 monomer are coloured in yellow, whereas the DIC2 segment is highlighted in red. The preformed β -sheets of the dimer interaction surface become extended by swapped antiparallel β -strands hence creating the binding surface for the incoming, kinked antiparallel β -strand of DIC2. The hydrophobic contacts between the DIC2 segment (KLGMAKITQVDFP) and the DYNLT1 structure are also shown in a stick format. **B.** Close-up of the interacting Val146, Phe148, Met140, and Ile143 residues, as well as the observed distances in Å (DIC2 segment in red, potentially interacting hydrophobic residues in green).

Since the two binding sites of the DYNLT1 dimer molecule are identical, it is interesting to mention that the length of the artificial linker designed in the chimera allows the possible coexistence of two different conformational structures that are perfectly compatible with the observed restrictions used in the structural calculation. In one, the sequence of DIC2 extends the β -strand of its own monomer while in the other conformation the chimeric end is inserted into the opposite side of the dimer, extending the β -sheet of the other monomer. This can be a possible explanation for some broad NMR signals observed in the spectra, as this would imply the presence of two dimeric populations in solution. Unfortunately both types of conformers are indistinguishable by NMR in the current conditions.

4. IDENTIFICATION OF NOVEL DYNLT1-INTERACTING PROTEINS CAPABLE OF OCCUPYING THE CANONICAL HYDROPHOBIC GROOVE

As discussed earlier, no structural data is available for DYNLT1 in complex with any polypeptide chain besides dynein chain. To explore the sequence requirements that lead to cellular and viral peptides to bind to the hydrophobic groove of DYNLT1 a binding assay using the pepscan technique was performed. A matrix of pentadecapeptides attached to a cellulose membrane through a spacer was created using the DIC1 sequence and each position was substituted for each of the 20 natural amino acids. Recombinant DYNLT1 was incubated with a cellulose membrane bearing 300 spots with synthetic pentadecapeptides that expand the binding possibilities. These incubation assays revealed that residues at the N-terminus (Lys) or C-terminus (Arg) lied outside the binding groove and essentially, any amino acid could occupy those positions (Figure 5.29). Remarkably, there are five positions in human DIC, those occupied by Leu148, Val150, Val153, Val156 and Phe158 where only hydrophobic amino acids can be accepted. These five hydrophobic binding pockets were also identified in the crystal structure (Williams et al., 2007) and the results presented here clearly demonstrate that they dictate the interaction of associated polypeptides.

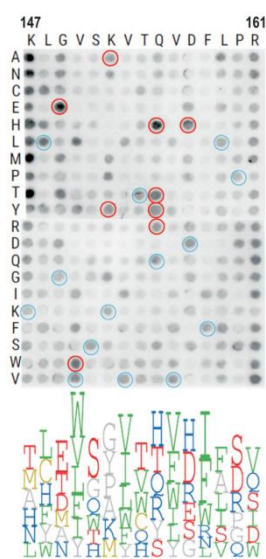


Figure 5.29. Analysis of the binding plasticity of DYNLT1 canonical groove. Binding of DYNLT1 to a matrix of pentadecapeptides corresponding to human DIC in which each position was substituted by every natural amino acid. A logo distribution is shown at the bottom obtained after densitometric quantification of the spots. Only the six residues with higher signal are shown for clarity. Residues below average signal were not included in the logo representation. Blue circles indicate identical substitutions that serve as positive controls. Red circles indicate substitutions that rendered positive binding and were subsequently confirmed in the yeast two-hybrid assay.

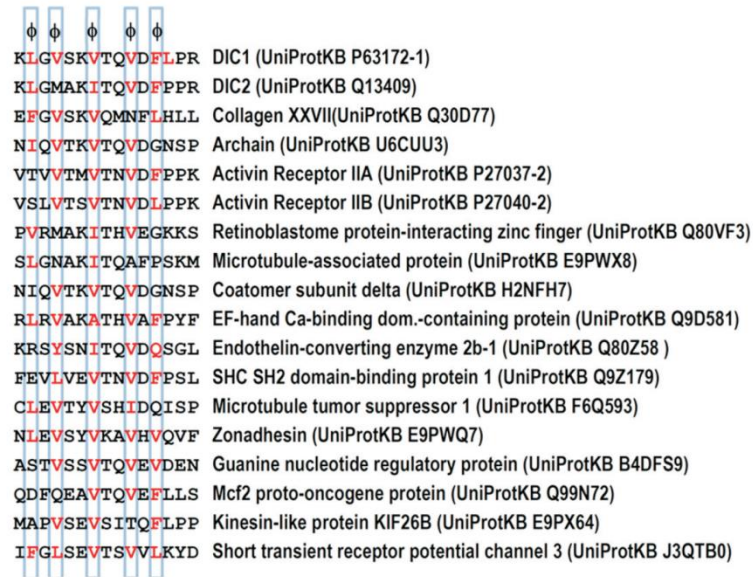
Although the DYNLT1-binding sites within DIC1 and DIC2 almost share the central K(I/V)TQVDF heptapeptide, it is true that the pepscan results show that the flanking residues found in DIC2 (Met, Ala, and Pro) are not pointedly favoured over other amino acids. This might indicate that this conserved heptapeptide significantly contributes to the binding energy.

Using a yeast two-hybrid assay in which each of these five hydrophobic amino acids was eliminated, it was confirmed that the binding was completely lost or significantly reduced (Figure 5.30). On the other hand, certain positions of DIC1 seemed to accept a significant diversity of residues, such as Gly149, Gln155, and even Leu159, maintaining the binding to DYNLT1. Interestingly, a few substitutions rendered peptides with increased binding to DYNLT1 and, although some of them were between conserved amino acids, the appearance of certain changes was hard to predict beforehand considering the nature of the amino acids side chains. This was the case of the Gly149Glu, Gly149Thr, Lys152Ala, Lys152Pro, Lys152Tyr, Gln155His, Gln155Thr, Gln155Tyr, Gln155Arg and Asp157His substitutions. In order to confirm that these mutations were indeed tolerated, DIC mutants were created to be used in the yeast two-hybrid assay and confirm that all these unexpected substitutions did indeed result in dynein intermediate chain sequences that interacted with DYNLT1 tightly (Figure 5.30).

Construct	His ⁺	X-Gal
DIC wt	++	++
DIC Leu148Gly	--	--
DIC Val150Gly	--	--
DIC Val153Gly	+/-	+/-
DIC Val156Gly	+/-	+/-
DIC Phe158Gly	--	--
○ DIC Gly149Glu	++	++
○ DIC Val150Trp	++	++
○ DIC Lys152Ala	++	++
○ DIC Lys152Tyr	++	++
○ DIC Gln155His	++	++
○ DIC Gln155Thr	++	++
○ DIC Gln155Tyr	++	++
○ DIC Gln155Arg	++	++
○ DIC Asp157His	++	++

Figure 5.30. Yeast two-hybrid assay using DYNLT1 in the bait plasmid and various DIC constructs. Growth in the absence of the amino acid His, in the presence of 3-amino triazole or X-Gal activity is shown. The upper part shows the binding assay for constructs in which hydrophobic amino acids have been mutated into Gly (Leu148Gly, Val150Gly, Val153Gly, Val156Gly and Phe158Gly). The bottom part shows dynein intermediate chain mutants with increased binding to DYNLT1 in the pepscan experiment. In several cases, these mutations comprise non-conservative substitutions, hence difficult to predict.

In accordance with these results, sequence databases were inspected looking for cellular proteins with putative DYNLT1 binding motifs based on qualitative inspection. 15 selected proteins matched with the required five hydrophobic residues at the specified positions, and displayed residues that were compatible with our pepscan results (Figure 5.31). Future experiments will certainly determine if these polypeptide stretches of proposed DYNLT1-interacting proteins are indeed in an exposed protein surface and if these proteins do, in fact, bind to DYNLT1. For now, the interaction of DYNLT1 with one of those proposed target proteins, Activin Receptor IIB, has been demonstrated in this work, as detailed below.



KLGVSKVTQVDFLPR	DIC1 (UniProtKB P63172-1)
KLGMAKITQVDFPPR	DIC2 (UniProtKB Q13409)
EFGVSKVQMNFLLHL	Collagen XXVII(UniProtKB Q30D77)
NIQVTKVTQVDGNSP	Archain (UniProtKB U6CUU3)
VTVTMTVTNVDFPPK	Activin Receptor IIA (UniProtKB P27037-2)
VSLVTSTVTNVDFPPK	Activin Receptor IIB (UniProtKB P27040-2)
PVRMAKITHVEGKKS	Retinoblastome protein-interacting zinc finger (UniProtKB Q80VF3)
SLGNAKITQAFPSKM	Microtubule-associated protein (UniProtKB E9PWX8)
NIQVTKVTQVDGNSP	Coatomer subunit delta (UniProtKB H2NFH7)
RLRVAKATHVAFPYF	EF-hand Ca-binding dom.-containing protein (UniProtKB Q9D581)
KRSYSNITQVDQSG	Endothelin-converting enzyme 2b-1 (UniProtKB Q80Z58)
FEVLVEVTNVDFPSL	SHC SH2 domain-binding protein 1 (UniProtKB Q9Z179)
CLEVTYVSHIDQISP	Microtubule tumor suppressor 1 (UniProtKB F6Q593)
NLEVSIVKAVHVQVF	Zonadhesin (UniProtKB E9PWQ7)
ASTVSSVTQVEVDEN	Guanine nucleotide regulatory protein (UniProtKB B4DFS9)
QDFQEAVTQVEFLLS	Mcf2 proto-oncogene protein (UniProtKB Q99N72)
MAPVSEVSITQFLPP	Kinesin-like protein KIF26B (UniProtKB E9PX64)
IFGLSEVTSVVLKYD	Short transient receptor potential channel 3 (UniProtKB J3QTB0)

Figure 5.31. List of proposed cellular DYNLT1–interacting partners according to the results obtained in the pepscan analysis.

5. DYNLT1 INTERACTS WITH THE INTRACELLULAR DOMAIN OF ACTIVIN RECEPTOR IIB

In cells, a significant part of DYNLT1 is not attached to the dynein complex (Tai et al., 1998) raising the possibility that DYNLT1 might play roles unrelated to retrograde microtubule–associated movement. The homodimeric structure of DYNLT1 and the positioning of the interacting partners within the canonical binding grooves indicate that DYNLT1 might function as a dimerization clamp, in a similar fashion as that described for DYNLL1 (Barbar, 2008). Among the proposed DYNLT1–associated proteins Activin Receptor II (ActRIIB) was selected as a likely candidate to associate to DYNLT1. In fact, a potential DYNLT1–binding motif was identified exactly at the very C–terminus end of ActRIIB, with a very similar sequence to that shown by DIC (Figure 5.32).

ActRIIB is a type II member of the family of serine/threonine kinase receptors that bind and are activated by the activin and myostatin signalling pathways (De Caestecker, 2004). The N–terminal, extracellular part of ActRIIB binds to activin, whereas the C–terminal, intracellular part displays the serine–threonine kinase domain. Activins are hormonal regulators belonging to the TGF β family and they are involved in cell growth and differentiation processes. (Attisano, Wrana, Montalvo, & Massague, 1996) It has been reported in the literature the implication of DYNLT1 in the TGF β –receptor–mediated signalling through the association to the C–terminus of the bone morphogenetic protein (BMP). (Machado et al., 2003) Likewise, other dynein components, such as proteins from DYNLRB family, also participate in the mentioned signalling pathway, via association with type II TGF β receptors. (Jin et al., 2009; Tang et al., 2002) Moreover, Tctex1D4 (\equiv Tctex2 β), another member of DYNLT family, is also implied in this system through the binding to endoglin, a co–receptor of TGF β receptors. (Meng et al., 2006) Also these data point to a possible biological relevance of the DYNLT1–ActRIIB interaction.

ActRIIB 491 **SDCLVSLVTSVTNVDLPPESSI** 512
 Φ Φ Φ Φ Φ

	His ⁻
DIC1(1-180)	++
ActRIIB (428-512)	++
ActRIIB (428-482)	--

Figure 5.32. Sequence of ActRIIB at the binding site within the canonical groove of DYNLT1. The positions where hydrophobic amino acids are expected are shown at the bottom using the Φ symbol. Interaction between the C-terminus of ActRIIB (428–512) and DYNLT1 in a yeast two-hybrid assay. DIC1 was included as a positive control.

A yeast two-hybrid assay confirmed that a stretch comprising residues 428–512 of ActRIIB associates to DYNLT1 *in vitro*, hence establishing that both proteins do interact (Figure 5.32). Then, an ActRIIB construct in which the final 22 residues were deleted was created, resulting in a polypeptide that was unable to associate to DYNLT1. HEK293 cells were transfected with a C-terminally GFP-tagged full-length ActRIIB (residues 1–512), or GFP-tagged ActRIIB-ΔCt (residues 1–490). Confocal microscopy studies revealed that both full-length and the C-terminally deleted ActRIIB construct localized to the plasma membrane and endomembranes. On the other hand, as expected (Merino-Gracia et al., 2015; Tai et al., 1998), a mCherry-tagged DYNLT1 displayed a nuclear and cytosolic staining that extended essentially throughout the entire cell (Figure 5.33). Hence, due to the essentially identical subcellular distribution of full-length ActRIIB when compared with its GFP-tagged ActRIIB-ΔCt counterpart, it was not possible to determine if the association of ActRIIB and DYNLT1 took place at the plasma membrane or in intracellular membranes during the sorting pathway.

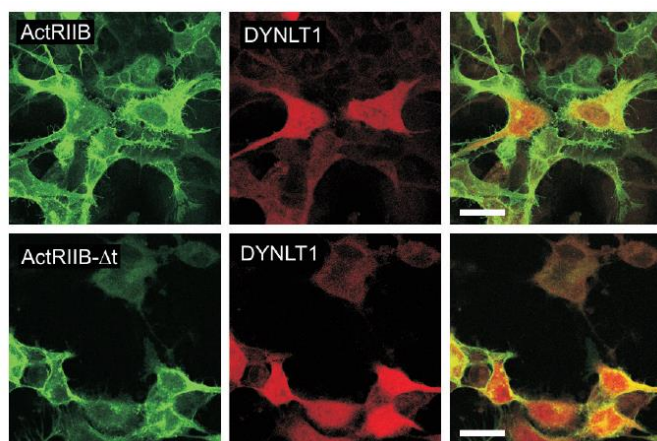


Figure 5.33. Confocal microscopy immunofluorescence of HEK293 cells transfected with mCherry-tagged DYNLT1 together with GFP-tagged ActRIIB constructs. The upper panels show full-length GFP-tagged ActRIIB whereas the bottom panels show C-terminally-deleted GFP-tagged ActRIIB. Merged panels are shown to the right in both cases. Bars, 25 μ m.

Nevertheless, DYNLT1 associated to full-length ActRIIB, but not to its ActRIIB- Δ Ct counterpart when transfected in COS7 cells (Figure 5.34), hence reinforcing the notion that they associate in cells.

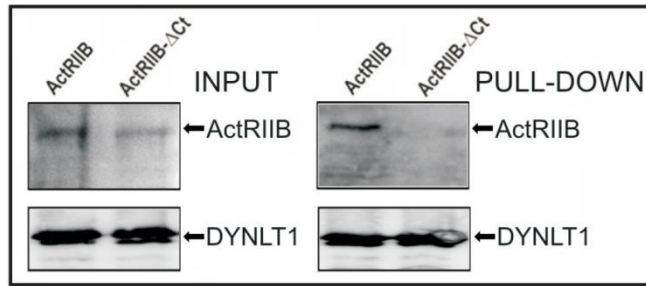


Figure 5.34. Lectin-tagged DYNLT1 and full-length ActRIIB or its ActRIIB-ΔCt counterpart in transfected COS7 cells were allowed to associate and the DYNLT1 moiety was sedimented with Sepharose beads. The appearance of each protein in the pellet fraction was determined with the appropriate antibodies.

The question of whether DYNLT1 was also involved in the signalling pathway of TGFβ was next investigated using a luciferase reporter assay. Overexpression of both full-length ActRIIB and ActRIIB-ΔCt in the presence of the p3TP-lux plasmid in transfected HEK293 cells resulted in a significant increase of the reporter activity (Figure 5.35). Interestingly, deletion of the DYNLT1-binding site in the context of a functional ActRIIB resulted in a marked increase of the reporter activity. Likewise, overexpression of a FLAG-tagged DYNLT1 construct that could bind to transfected full-length ActRIIB resulted in a diminished signalling response. Taken together, these data indicate that DYNLT1 binding to ActRIIB results in the inhibition of its signalling activity.

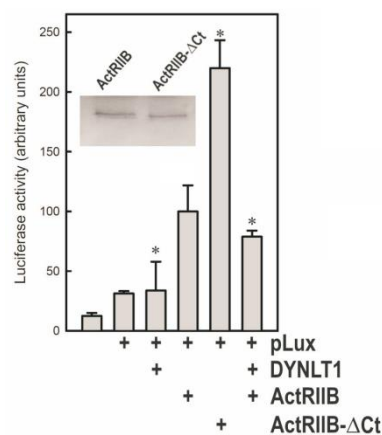


Figure 5.35. Effect of DYNLT1 on ActRIIB signalling was analysed in transfected HEK293 cells using a luciferase reporter construct. The insert shows a Western blot of transfected ActRIIB and ActRIIB-ΔCt. All experiments were performed in duplicate wells and repeated three times. The results show average-fold changes and are referred to transfected wild-type ActRIIB in the presence of p3TP-lux plasmid. S.D. values are indicated by error bars with *, $p < 0.05$.

6. THE BINDING SITE OF LFC WITHIN DYNLT1 EXTENDS WELL-BEYOND THE CANONICAL BINDING GROOVE

The association between DYNLT1 and guanine nucleotide exchange factor Lfc has been analysed extensively (Balan, 2013; Meiri et al., 2012; Meiri et al., 2014). Interestingly, it has been reported that Lfc associates within the DYNLT1 canonical binding groove and, at the same time, seems to be transported along microtubules, two observations that seem contradictory. Sequence comparison between DIC and Lfc reveals that the latter essentially maintains the five hydrophobic residues known to fit within the canonical binding groove but, surprisingly, various residues seem to be conserved upstream this canonical binding stretch (Figure 5.36).



Figure 5.36. Sequence alignment between DIC and Lfc. DYNLT1 and DYNLL1 binding regions are shown. The positions where hydrophobic amino acids are expected are marked at the bottom line using the symbol Φ.

Remarkably, upstream from the predicted stretch that would insert within the hydrophobic binding groove of DYNLT1, Lfc also displayed a polybasic motif. A yeast two-hybrid assay was used to try to identify the minimal DYNLT1-binding region of Lfc. If Lfc could behave in the same manner as DIC, a 20-amino acid stretch comprising Lfc residues 142 to 161 (SSL...FND) would be enough to render a positive interaction with DYNLT1. In contrast, not even an Lfc construct comprising residues 135–180 could result in a detectable binding to DYNLT1 (Figure 5.37).

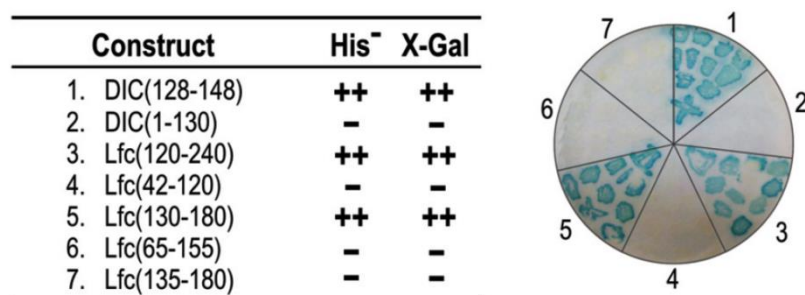


Figure 5.37. Yeast-two hybrid assay in which DYNLT1 was confronted to various DIC and Lfc constructs. Growth in the absence of histidine, in the presence of 3-amino triazole, as well as X-Gal activity is shown.

In agreement with previous published results (Meiri et al., 2014), it could be concluded that only large portions of Lfc could interact with DYNLT1. The most likely explanation to this piece of data would be that the Lfc-DYNLT1 interaction requires additional contacts involving residues lying outside the canonical binding groove. It is tempting to speculate that the N-terminal part of Lfc, lying outside the hydrophobic groove might interact with the α -helices of DYNLT1, as previously proposed. (Meiri et al., 2012) It must be mentioned, in this context, that DIC itself, displays both a polybasic stretch and also the PS and PQ motifs found in Lfc (Figure 5.36). This might indicate that even DIC might have secondary binding sites to the surface of DYNLT1 that were identified neither in the crystal structures of the DIC-DYNLT1 complexes, nor in the NMR solution structure presented here. To further prove that Lfc binding to DYNLT1 requires extensive contacts lying outside the consensus binding groove, ITC was employed, and two long peptides of Lfc were tested. As expected, this approach gave more quantitative information regarding the Lfc-DYNT1 interaction. Using ITC the binding of both Lfc(131–154) or Lfc(131–161) peptides to DYNLT1 could be detected. The longer peptide rendered a $K_d = 11.7 \pm 0.9 \mu\text{M}$ (thermodynamic parameters: $\Delta G = -6.72 \text{ kcal}\cdot\text{mol}^{-1}$, $\Delta H = -4.03 \text{ kcal}\cdot\text{mol}^{-1}$ and $-T\Delta S = 2.70 \text{ kcal}\cdot\text{mol}^{-1}$), whereas the shorter peptide rendered a $K_d = 78 \pm 2 \mu\text{M}$ (thermodynamic parameters: $\Delta G = -5.59 \text{ kcal}\cdot\text{mol}^{-1}$, $\Delta H = -9.05 \text{ kcal}\cdot\text{mol}^{-1}$ and $-T\Delta S = 3.46 \text{ kcal}\cdot\text{mol}^{-1}$).

Further confirmation of the extended DYNLT1 binding site within Lfc was obtained when a shorter peptide corresponding to Lfc(137–154) showed no binding heat in an ITC assay (Figure 5.38).

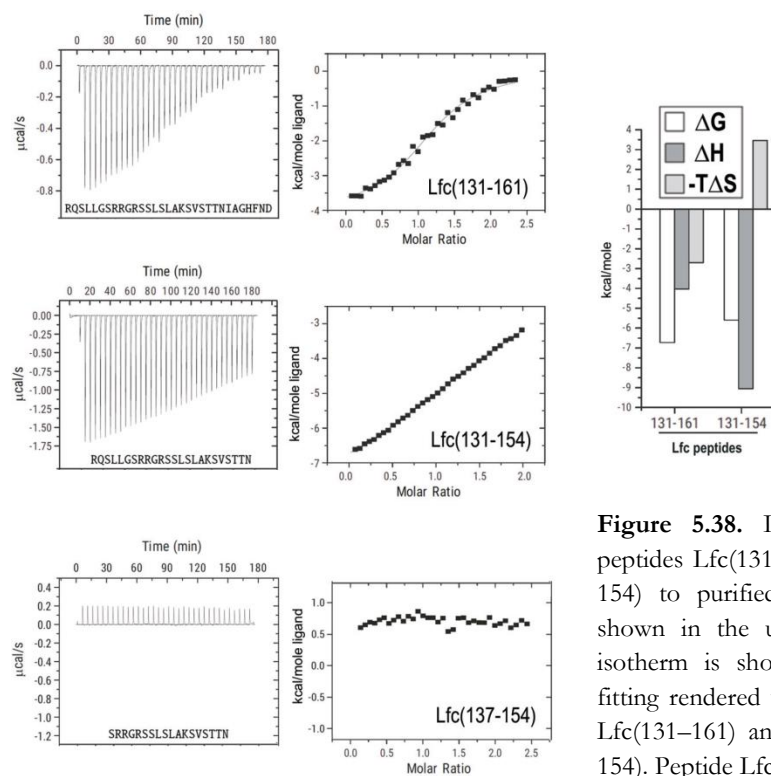


Figure 5.38. ITC analysis of the binding of peptides Lfc(131–161), Lfc(131–154) and Lfc(137–154) to purified DYNLT1. The thermogram is shown in the upper panel whereas the binding isotherm is shown in the bottom panel. Curve fitting rendered values of $K_d = 11.7 \pm 0.9 \mu\text{M}$ for Lfc(131–161) and $K_d = 78 \pm 2 \mu\text{M}$ for Lfc(131–154). Peptide Lfc(137–154) gave no binding heat.

To assess if the polybasic residues present in Lfc preceding the part predicted to insert within the canonical groove are involved in the interaction, the mutation Arg138/Arg139 was executed, and confirmed that the mutant Lfc polypeptide showed a significant lowered interaction in a yeast two-hybrid assay (Figure 5.39). Likewise, an Lfc Gly140Ala mutant bound poorly to DYNLT1. Somehow expectedly, mutant Lfc constructs of residues known to insert in the hydrophobic groove (Val150, Ser151 and Thr153) completely failed to interact with DYNLT1.



Construct	His ⁺	X-Gal
Lfc(130-180)	++	++
Lfc(130-180) R138/139A	+/-	+/-
Lfc(130-180) G140A	+/-	+/-
Lfc(130-180) V150A, S151A	-	-
Lfc(130-180) V150A, S151A, T153A	-	-

Figure 5.39. Yeast–two hybrid assay in which DYNLT1 was confronted to various Lfc constructs. Growth in the absence of histidine, in the presence of 3–amino triazole, as well as X–Gal activity is shown.

It has been previously shown that binding of microtubules to Lfc and to its homolog GEF-H1 might be regulated through Ser phosphorylation. For instance, polarity kinase Par1b phosphorylates GEF-H1 Ser143 (Yoshimura & Miki, 2011), a post-translational modification proposed to disrupt its interaction to DYNLT1 and to release it from the dynein motor. (Meiri et al., 2012) In fact, analysis of the Lfc stretch ¹³⁸RRGRSS¹⁴³ using kinase phosphorylation prediction programs easily reveals that both Ser142 and Ser143 could be substrates of several other protein kinases, such as ATM, GSK-3 or AKT1. Hence, the presence of these two phosphorylatable Ser residues within the DYNLT1-binding stretch of Lfc might be a way of regulating this interaction.

In order to determine whether DIC1 residues located next to those known to become inserted within the binding groove might also contribute to the overall binding energy the binding of a longer FITC-labelled DIC1 fragment (residues 118–178) was tested using fluorescence polarization measurements. The K_d value between probe and DYNLT1 was determined to be $1.6 \pm 0.8 \mu\text{M}$ (Fig. 6E). Hence, both the PS, QS motifs of DIC1 or other upstream residues contribute to the association to DYNLT1 since we previously obtained a K_d value of be $4.8 \pm 0.5 \mu\text{M}$ when using the FITC-labelled DIC1 (147–162) peptide (Figure 5.40).

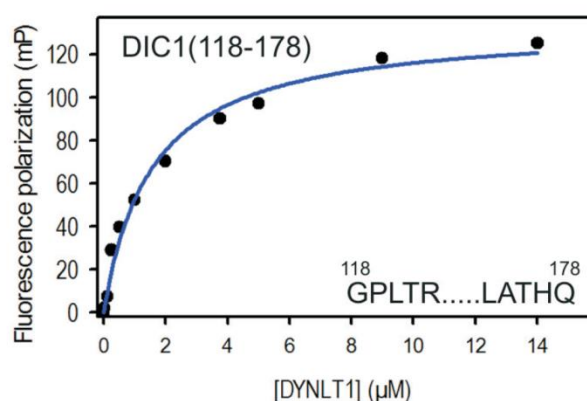


Figure 5.40. Representative binding curve for a FITC-labelled DIC1 peptide (118–178) to DYNLT1 measured by fluorescence polarization. A calculated $K_d = 1.6 \pm 0.8 \mu\text{M}$ could be obtained. Results are representative of three independent experiments.

7. OTHER DYNLT1-INTERACTING PARTNERS MIGHT ADOPT A SIMILAR WAY OF INTERACTION

Sequence comparison of DIC, ActRIIB, and Lfc with other proteins known to bind to DYNLT1, such as the parathyroid hormone-related protein receptor (Sugai et al., 2003), poliovirus receptor CD155 (Mueller, Cao, Welker, & Wimmer, 2002), or the neuronal protein Doc2 (Nagano et al., 1998), reveals a similar amino acid distribution (Figure 5.41). In general, these proteins do have hydrophobic amino acids at the five positions where residues are expected to fulfil this requirement. In most cases, the sequence known to insert within the canonical binding groove is preceded by a stretch of basic amino acids. Furthermore, not

only Lfc, but also PTHR display phosphorylatable Ser residues immediately following these basic residues, suggesting that phosphorylation might regulate DYNLT1 binding.

	polybasic motif Φ Φ Φ Φ Φ
DIC1 (UniProtKB O14576)	DSELG RRLH <u>KLGVSKVTQVD</u> FLP
ActRIIB (UniProtKB P27040-2)	VNGTTSDCLV <u>SLVTSVTNVD</u> LPP
Lfc (UniProtKB Q865S3)	LLGS RRGRSS <u>LSLAKSVST</u> TNIA
PTHR (CN-11) (UniProtKB H7C0Q3)	KRKARSGSSS <u>YSYGPMVSH</u> TSVT
Polio Rec CD155 (UniProtKB A0A0C4DG49)	YFYWS KCSRE <u>VLWHCHLCP</u> SSTE
Human Doc2 (UniProtKB H3BT11)	MRGRRGDRMT <u>INIQEHMAIN</u> VCP

Figure 5.41. Sequence comparison between six DYNLT1-binding partners showing the position of the stretch that inserts within the hydrophobic binding groove (in blue) and the basic residues (in red). Note that the Ser/Ser motif present in Lfc and PTHR probably becomes phosphorylated *in vivo*.

8. DISCUSSION

Microtubules are hollow cylindrical fibres of approximately 25 nm in diameter that are formed by the polymerization of the $\alpha\beta$ -tubulin heterodimer. Organelles, vesicles and other intracellular cargoes are transported by kinesin and dynein motors, which move in opposite directions along microtubules. In the case of the dynein motor, the identity of the precise dynein polypeptides involved in cargo binding remains the subject of much debate. Dynein light chains DYNLL (DLC8) and DYNLT (Tctex) have been proposed as cargo adaptors due to their ability to bind to DIC and also their ability to bind to multiple cellular and viral proteins. It has been recently demonstrated that the small GTPases RagA and Rab3D are bona fide DYNLT1-interacting proteins that associate to this dynein light chain when forming part of the dynein motor. (Merino-Gracia et al., 2015) However, among the over 20 DYNLT1-binding proteins it is not well established which ones use the canonical binding groove and which ones bind to the surface of DYNLT1.

To shed light on this question, information at atomic level should be necessary. The 3D structure of DYNLT1 auto-saturated with DIC in the canonical binding site has been resolved in this work. To do this, NMR spectroscopy was utilized to obtain the first mammalian structure available, allowing to the description of the exact sequence requirements that enable DIC association in the canonical DYNLT1 binding groove.

Although crystals of mouse apo-DYNLT1 have been obtained, diffraction was very weak. (Williams et al., 2005) Conversely, *D. melanogaster* apo-DYNLT1 crystals diffracting at sufficient quality have been obtained. Unfortunately, the absence of a binding peptide inserted in the canonical groove induced the formation of an aberrant hexameric structure in which each homodimer sequestered polypeptides from an adjacent homodimer to occupy the binding site. (Williams et al., 2005) NMR studies performed using mouse DYNLT1 titrated with a dynein intermediate chain peptide (that included the polybasic stretch) showed changes in the $\beta 2$ -strand and the $\alpha 2/\alpha 3$ -loop. Unfortunately, the exact binding mode of interaction for those peptides could not be obtained (Mok et al., 2001). Furthermore, NMR spectroscopy also indicated that titration of *Chlamydomonas* DYNLT1 with an intermediate

chain peptide resulted in changes in chemical shifts whose mapping onto the molecular surface revealed that they do not occur along the intermonomer grooves but rather are all located at one end of the molecule (Wu et al., 2005). This significantly contrasts with the two available crystal structures of DYNLT1 in complex with dynein intermediate chain peptides. (Hall et al., 2009; Williams et al., 2007)

We have shown human DYNLT1 in complex with DIC adopts a fold similar to that observed for both the *Chlamydomonas* (Wu et al., 2005) and fly complexes (Figure 5.42). (Hall et al., 2009; Williams et al., 2007) In general, the fly complex obtained by protein crystallography is more compact than either of the solution structures obtained by NMR spectroscopy. In all cases, the incoming peptide adopts an antiparallel β -strand conformation that extends the preformed β -sheet albeit with a clear kink that disrupts the expected hydrogen bond pattern (Figure 5.42, and references (Williams et al., 2007) and (Balan, 2013)). This feature probably makes it more difficult to predict putative DYNLT1-binding proteins just from the amino acid sequence.

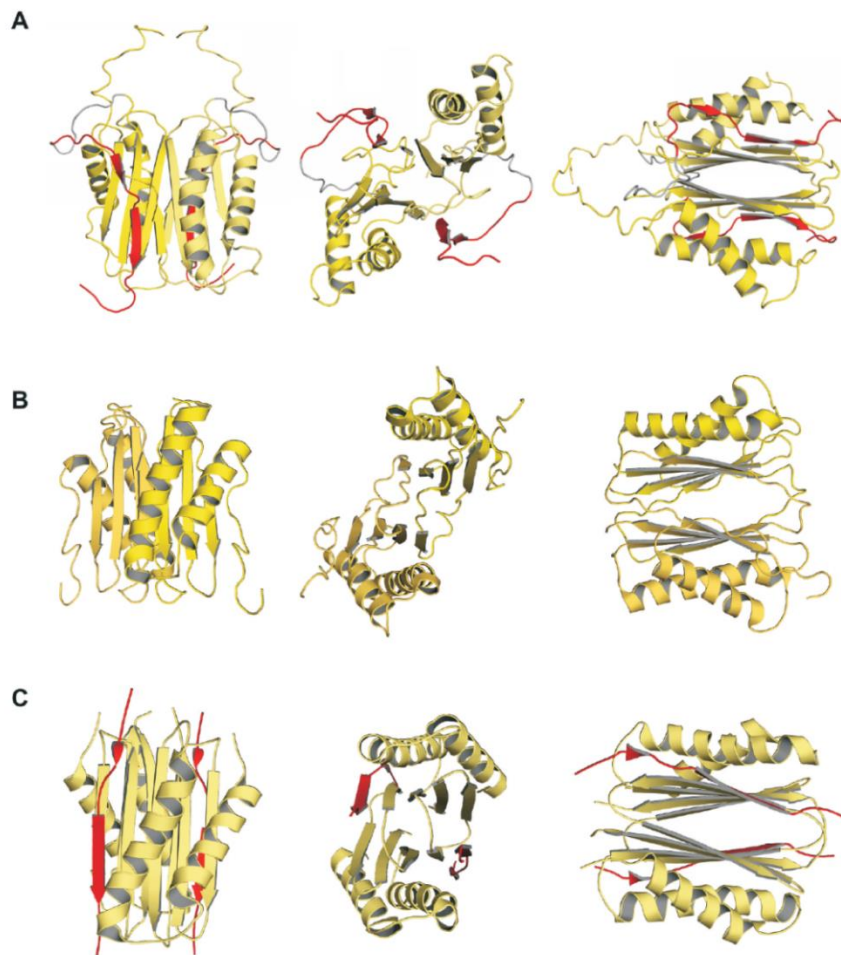


Figure 5.42. Structural comparison of the human, *C. reinhardtii*, and *D. melanogaster* DYNLT1 in complex with DIC peptides. **A.** Three views of the solution structure of human DYNLT1 in complex with human DIC are shown in comparison with the *C. reinhardtii* ortholog structure (**B**), also obtained by NMR spectroscopy (PDB 1XDX) (Wu et al., 2005), and the *D. melanogaster* crystal structure (**C**) (PDB 2PG1) (Williams et al., 2007).

The data presented here also suggest that not only Lfc, but also DIC itself very likely establish contacts with DYNLT1 in superficial areas distant from the hydrophobic groove. However, in the case of DIC a short stretch inserting in the groove renders a positive interaction whereas in the case of Lfc both contact areas seem to be necessary for binding. Furthermore, unlike DIC binding to DYNLT1, the lower affinity association of Lfc to DYNLT1 is regulated through Ser phosphorylation.

It has been also described herein the interaction between DYNLT1/Tctex-1 and Activin Receptor IIB. Interestingly, Tctex2 β , another dynein light chain family member binds to several members of the transforming growth factor- β receptors, such as endoglin, T β RII, betaglycan and ActRIIA (Meng et al., 2006). Tctex2 β has two clearly-defined domains and is larger than DYNLT1/Tctex-1 although not so dissimilar since its C-terminal part is highly homologous to DYNLT1/Tctex-1. In agreement with these results, overexpression of Tctex2 β also inhibits TGF β signalling. Likewise, there is a significant co-localization of endoglin and Tctex2 β in endomembranes when co-transfected in HEK293 cells. The suggestion that Tctex2 β might bridge together endoglin and T β RII (Meng et al., 2006) raises the possibility of DYNLT1/Tctex-1 also forming part of tripartite complexes with transmembrane TGF β receptors hence regulating their activity. Further studies will be necessary to address this issue.

With all the available data, different modes of protein binding to dynein light chain DYNLT1 can be proposed (Figure 5.43A). Small GTPases such as RagA can bind to the surface of both DYNLT1 and DYNLT1-DIC efficiently (Merino-Gracia et al., 2015), and they do not occupy the canonical binding groove. Then, they can form a tripartite complex with DYNLT1 and DIC. In this situation, DYNLT1 functions as a dimerization clamp that subsequently links these proteins to DIC and, finally, they become attached to microtubules via the dynein motor. On the other hand, Lfc and its homolog GEF-H1 bind to DYNLT1 using two sites: the canonical binding groove and part of the surface of DYNLT1. Eventually, the canonical binding groove becomes occupied by DIC itself that would displace Lfc from the binding groove without dissociating the trimeric complex. (Meiri et al., 2012) Then, the complex can be subsequently targeted to microtubules via the dynein motor.

Since it is very well established that both small GTPases of the RagA or Rab3D families as well as the guanine nucleotide exchange factors Lfc and GEF-H1 do associate to microtubules (Meiri et al., 2012; Merino-Gracia et al., 2015; Pavlos et al., 2011; Yoshimura & Miki, 2011), the formation of a tripartite complex involving DIC/DYNLT1/protein must take place. Conversely, a significant population of DYNLT1 is not attached to microtubules (Tai et al., 1998), being most likely involved in functioning as a molecular dimerization clamp of cytoplasmic proteins. In this work, it has been proposed that DYNLT1 functions as a dimerization clamp for the TGF β family receptor ActRIIB, in an association not involving the subsequent attachment to microtubules (Fig 5.43B).

Although other cargo adapters beyond DYNLT1 have been reported, such as intermediate chain itself upon binding to circovirus proteins (Cao et al., 2015) or the Hsp90–

immunophilin complex that is associated to dynactin necessary for p53 binding (Galigniana, Harrell, O'Hagen, Ljungman, & Pratt, 2004), data presented here put forward novel ways of cargo attachment through DYNLT1.

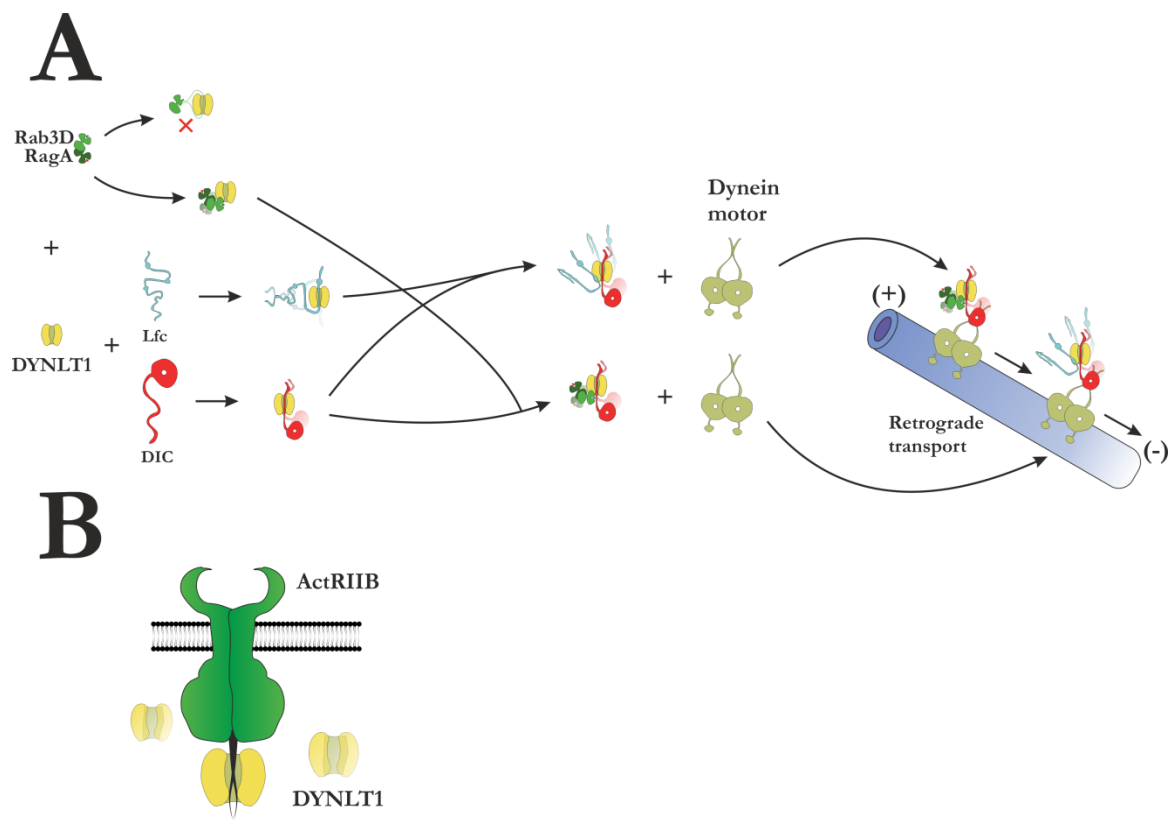


Figure 5.43. Proposed interaction model of DYNLT1 with various characterised targets. **A.** The small GTPases Raga and Rab3D are coloured in green, DIC in red, Lfc in blue. For description of the processes, see the text above. **B.** ActRIIB and other TGF- β receptors use DYNLT1 (or perhaps its homolog and Tctex2 β) as molecular clamps to anchor their C-termini.

CONCLUSIONS

1. The interaction of human DYNLT1 with DIC has been extensively characterised at residue level for the first time. A 15-residue sequence from DIC binds to DYNLT1 with high affinity, demonstrating that the preceding polybasic stretch is unessential for binding, contrasting with previous published reports.
2. Interaction of DYNLT1 with polypeptides through the canonical binding site is dictated by the presence of five hydrophobic binding pockets that, in the case of DIC, are occupied by residues Leu128, Val130, Val133, Val136, and Phe138.
3. 3D structure of the human DYNLT1–DIC chimera dimer, designed to describe the sequence requirements that enables DIC association in the DYNLT1 canonical binding groove, was elucidated by solution NMR. The structure is a well-defined swap dimer with a tertiary structure analogue to that shown by other members of the DYNLT family. DIC segments are located in the hydrophobic grooves on the structure surface extending the central β -sheet, and oriented in the same N to C terminus direction.
4. The β -structure of DIC in the designed construct is defined in two stretches (sequences Met–Ala–Lys, and Thr–Gln–Val–Asp–Phe). Some structural distortion (a bulge) is observed around position Ile132, which was also reported in previously published structures.
5. At least three of the defined hydrophobic binding pockets in the canonical binding groove of the DYNLT1–DIC chimera are clearly identified. The other two pockets are less defined in the calculated structure, as they correspond to regions which are more solvent-exposed than in previously reported structures.
6. The designed chimera allows the coexistence of two different conformational structures compatible with the observed restrictions used in the calculations. According to the position of the DIC segments, they can be extending the β -strand of its own monomer, or they can be located in the other side of the dimer, extending the β -strand of the other monomer. In the conditions used in this work, both conformers are indistinguishable.
7. Putative DYNLT1 binding motifs were identified in 15 proteins, including Activin Receptor IIB (ActRIIB). *In vitro* and *in vivo* interaction between DYNLT1 and ActRIIB was proved, and the interaction region of ActRIIB was identified as a stretch comprising residues 428–512. DYNLT1 binding to ActRIIB results in the inhibition of its signalling activity, suggesting the possible participation of DYNLT1 in the signalling pathway of TGF- β . It has been proposed that DYNLT1 can act as a dimerization clamp for the TGF- β family receptor ActRIIB, in an association not involving attachment to microtubules.

8. It was demonstrated that protein Lfc binds DYNLT1. However, only long fragments of Lfc are able to interact with DYNLT1, indicating that the interaction requires additional contacts involving residues lying outside the canonical binding groove. Lfc polybasic segment preceding the fragment predicted to insert into the canonical groove seems to be also involved in the DYNLT1 binding. This binding process could be regulated through Ser phosphorylation.

9. An interaction model of DYNLT1 with different characterised targets that agrees with experimental observations has been proposed. This model combines proteins that interact only by binding in the canonical groove with others that use exclusively or additionally some adjacent regions. This binding event is a preliminary step towards the interactions with the dynein motor and their attachment to microtubules for their transport.

BIBLIOGRAPHY

- Alberts B, Johnson A, Lewis J, & al., e. (2002). Molecular Motors. In G. Science (Ed.), *Molecular biology of the cell* (4th ed.).
- Angulo, I., Acebrón, I., de Las Rivas, B., Muñoz, R., Rodríguez-Crespo, I., Menéndez, M., . . . Pérez-Agote, B. (2011). High-resolution structural insights on the sugar-recognition and fusion tag properties of a versatile β -trefoil lectin domain from the mushroom *Laetiporus sulphureus*. *Glycobiology*, 21(10), 1349-1361.
- Ashokkumar, B., Nabokina, S. M., Ma, T. Y., & Said, H. M. (2009). Identification of dynein light chain road block-1 as a novel interaction partner with the human reduced folate carrier. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 297(3), G480.
- Attisano, L., Wrana, J. L., Montalvo, E., & Massague, J. (1996). Activation of signalling by the activin receptor complex. *Molecular and cellular biology*, 16(3), 1066-1073.
- Balan, M. (2013). *The regulation of RhoGEF Lfc by dynein light chain Tctex-1*. University of Toronto.
- Barbar, E. (2008). Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. *Biochemistry*, 47(2), 503-508.
- Benison, G., Karplus, P. A., & Barbar, E. (2008). The Interplay of Ligand Binding and Quaternary Structure in the Diverse Interactions of Dynein Light Chain LC8. *J Mol Biol*, 384(4), 954-966. doi:<http://dx.doi.org/10.1016/j.jmb.2008.09.083>
- Campbell, K. S., Cooper, S., Dessing, M., Yates, S., & Buder, A. (1998). Interaction of p59fyn kinase with the dynein light chain, Tctex-1, and colocalization during cytokinesis. *The Journal of Immunology*, 161(4), 1728-1737.
- Cao, J., Lin, C., Wang, H., Wang, L., Zhou, N., Jin, Y., . . . Zhou, J. (2015). Circovirus transport proceeds via direct interaction of the cytoplasmic dynein IC1 subunit with the viral capsid protein. *Journal of virology*, 89(5), 2777-2791.
- Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., . . . Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of computational chemistry*, 26(16), 1668-1688.
- Conde, C., Arias, C., Robin, M., Li, A., Saito, M., Chuang, J.-Z., . . . Cáceres, A. (2010). Evidence for the involvement of Lfc and Tctex-1 in axon formation. *The Journal of Neuroscience*, 30(19), 6793-6800.
- Cook, S. P., Galve-Roperh, I., del Pozo, Á. M. n., & Rodríguez-Crespo, I. (2002). Direct calcium binding results in activation of brain serine racemase. *Journal of Biological Chemistry*, 277(31), 27782-27792.
- Chuang, J.-Z., Yeh, T.-Y., Bollati, F., Conde, C., Canavosio, F., Cáceres, A., & Sung, C.-H. (2005). The dynein light chain Tctex-1 has a dynein-independent role in actin remodeling during neurite outgrowth. *Developmental cell*, 9(1), 75-86.
- Day, C. L., Puthalakath, H., Skea, G., Strasser, A., Barsukov, I., Lian, L.-Y., . . . Hinds, M. G. (2004). Localization of dynein light chains 1 and 2 and their pro-apoptotic ligands. *Biochemical Journal*, 377(Pt 3), 597-605. doi:10.1042/BJ20031251
- De Caestecker, M. (2004). The transforming growth factor- β superfamily of receptors. *Cytokine & growth factor reviews*, 15(1), 1-11.
- Diamant, A. G., & Carter, A. P. (2013). Dynein Family Classification *Encyclopedia of Biophysics* (pp. 552-558): Springer.
- DiBella, L. M., Benashski, S. E., Tedford, H. W., Harrison, A., Patel-King, R. S., & King, S. M. (2001). The Tctex1/Tctex2 class of dynein light chains Dimerization, differential expression, and interaction with the LC8 protein family. *Journal of Biological Chemistry*, 276(17), 14366-14373.

- Fan, J.-S., Zhang, Q., Tochio, H., Li, M., & Zhang, M. (2001). Structural basis of diverse sequence-dependent target recognition by the 8 kDa dynein light chain. *J Mol Biol*, 306(1), 97-108.
- Fan, J.-S., Zhang, Q., Tochio, H., & Zhang, M. (2002). Backbone dynamics of the 8 kDa dynein light chain dimer reveals molecular basis of the protein's functional diversity. *Journal of Biomolecular NMR*, 23(2), 103-114.
- Fletcher, D. A., & Mullins, R. D. (2010). Cell mechanics and the cytoskeleton. *Nature*, 463(7280), 485-492. doi:10.1038/nature08908
- Galigniana, M. D., Harrell, J. M., O'Hagen, H. M., Ljungman, M., & Pratt, W. B. (2004). Hsp90-binding immunophilins link p53 to dynein during p53 transport to the nucleus. *Journal of Biological Chemistry*, 279(21), 22483-22489.
- Goddard, T., & Kneller, D. (2004). SPARKY 3. *University of California, San Francisco*, 14, 15.
- Guilluy, C., Garcia-Mata, R., & Burridge, K. (2011). Rho protein crosstalk: another social network? *Trends in cell biology*, 21(12), 718-726.
- Güntert, P. (2004). Automated NMR structure calculation with CYANA. *Protein NMR Techniques*, 353-378.
- Hall, J., Karplus, P. A., & Barbar, E. (2009). Multivalency in the assembly of intrinsically disordered dynein intermediate chain. *Journal of Biological Chemistry*, 284(48), 33115-33121.
- Hall, J., Song, Y., Karplus, P. A., & Barbar, E. (2010). The crystal structure of dynein intermediate chain-light chain roadblock complex gives new insights into dynein assembly. *Journal of Biological Chemistry*, 285(29), 22566-22575.
- Huang, J.-r., Craggs, T. D., Christodoulou, J., & Jackson, S. E. (2007). Stable intermediate states and high energy barriers in the unfolding of GFP. *J Mol Biol*, 370(2), 356-371.
- Huang, X., Wang, H.-L., Qi, S.-T., Wang, Z.-B., Tong, J.-S., Zhang, Q.-H., . . . Qi, Z.-Q. (2011). DYNLT3 is required for chromosome alignment during mouse oocyte meiotic maturation. *Reproductive Sciences*, 18(10), 983-989.
- Jin, Q., Gao, G., & Mulder, K. M. (2009). Requirement of a dynein light chain in TGF β /Smad3 signaling. *Journal of cellular physiology*, 221(3), 707-715.
- Jin, Q., Gao, G., & Mulder, K. M. (2013). A dynein motor attachment complex regulates TGF β /Smad3 signaling. *International journal of biological sciences*, 9(6), 531.
- Kendrick-Jones, J., Hodge, T., Lister, I., Roberts, R., & Buss, F. (2001). Myosin superfamily. *eLS*.
- King, S. J., Bonilla, M., Rodgers, M. E., & Schroer, T. A. (2002). Subunit organization in cytoplasmic dynein subcomplexes. *Protein Science*, 11(5), 1239-1250.
- King, S. M. (2000). The dynein microtubule motor. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1496(1), 60-75.
- Koradi, R., Billeter, M., & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics*, 14(1), 51-55.
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, 8(4), 477-486.
- Liang, J., Jaffrey, S. R., Guo, W., Snyder, S. H., & Clardy, J. (1999). Structure of the PIN/LC8 dimer with a bound peptide. *Nature Structural & Molecular Biology*, 6(8), 735-740.
- Liu, C., Chuang, J.-Z., Sung, C.-H., & Mao, Y. (2015). A dynein independent role of tctex-1 at the kinetochore. *Cell Cycle*, 14(9), 1379-1388.
- Lo, K. W.-H., Kan, H.-M., & Pfister, K. K. (2006). Identification of a novel region of the cytoplasmic Dynein intermediate chain important for dimerization in the absence of the light chains. *Journal of Biological Chemistry*, 281(14), 9552-9559.

- Lo, K. W.-H., Kogoy, J. M., & Pfister, K. K. (2007). The DYNLT3 light chain directly links cytoplasmic dynein to a spindle checkpoint protein, Bub3. *Journal of Biological Chemistry*, 282(15), 11205-11212.
- Lo, K. W.-H., Kogoy, J. M., Rasoul, B. A., King, S. M., & Pfister, K. K. (2007). Interaction of the DYNLT (TCTEX1/RP3) light chains and the intermediate chains reveals novel intersubunit regulation during assembly of the dynein complex. *Journal of Biological Chemistry*, 282(51), 36871-36878.
- Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Scott, M. P., Bretscher, A., . . . Matsudaira, P. (2008). *Molecular Cell Biology* (W. H. Freeman Ed. 4th ed.).
- Lukashok, S. A., Tarassishin, L., Li, Y., & Horwitz, M. S. (2000). An adenovirus inhibitor of tumor necrosis factor alpha-induced apoptosis complexes with dynein and a small GTPase. *Journal of virology*, 74(10), 4705-4709.
- Machado, R. D., Rudarakanchana, N., Atkinson, C., Flanagan, J. A., Harrison, R., Morrell, N. W., & Trembath, R. C. (2003). Functional interaction between BMPR-II and Tctex-1, a light chain of Dynein, is isoform-specific and disrupted by mutations underlying primary pulmonary hypertension. *Human molecular genetics*, 12(24), 3277-3286.
- Makokha, M., Hare, M., Li, M., Hays, T., & Barbar, E. (2002). Interactions of cytoplasmic dynein light chains Tctex-1 and LC8 with the intermediate chain IC74. *Biochemistry*, 41(13), 4302-4311.
- Makokha, M., Huang, Y. J., Montelione, G., Edison, A. S., & Barbar, E. (2004). The solution structure of the pH-induced monomer of dynein light-chain LC8 from *Drosophila*. *Protein Science*, 13(3), 727-734.
- Martínez-Moreno, M., Navarro-Lérida, I., Roncal, F., Albar, J. P., Alonso, C., Gavilanes, F., & Rodríguez-Crespo, I. (2003). Recognition of novel viral sequences that associate with the dynein light chain LC8 identified through a pepscan technique. *FEBS Letters*, 544(1), 262-267.
- McNiven, M. A. (1998). Dynamin: A Molecular Motor with Pinchase Action. *Cell*, 94(2), 151-154. doi:[http://dx.doi.org/10.1016/S0092-8674\(00\)81414-2](http://dx.doi.org/10.1016/S0092-8674(00)81414-2)
- Meiri, D., Marshall, C. B., Greeve, M. A., Kim, B., Balan, M., Suarez, F., . . . Fine, N. (2012). Mechanistic insight into the microtubule and actin cytoskeleton coupling through dynein-dependent RhoGEF inhibition. *Molecular cell*, 45(5), 642-655.
- Meiri, D., Marshall, C. B., Mokady, D., LaRose, J., Mullin, M., Gingras, A.-C., . . . Rottapel, R. (2014). Mechanistic insight into GPCR-mediated activation of the microtubule-associated RhoA exchange factor GEF-H1. *Nature communications*, 5.
- Meng, Q., Lux, A., Holloschi, A., Li, J., Hughes, J. M., Foerg, T., . . . Hafner, M. (2006). Identification of Tctex2 β , a novel dynein light chain family member that interacts with different transforming growth factor- β receptors. *Journal of Biological Chemistry*, 281(48), 37069-37080.
- Merino-Gracia, J., Costas-Insua, C., Canales, M.-A., & Rodríguez-Crespo, I. (2016). Insights into the C-terminal Peptide Binding Specificity of the PDZ domain of neuronal Nitric Oxide Synthase. Characterisation of the Interaction with the tight-junction protein Claudin-3. *Journal of Biological Chemistry*, jbc. M116. 724427.
- Merino-Gracia, J. (2016). *Interacciones proteína-péptido: Estudio de DYNLT1 y del dominio PDZ de la óxido nítrico sintasa neuronal*. Universidad Complutense de Madrid.
- Merino-Gracia, J., García-Mayoral, M. F., Rapali, P., Valero, R. A., Bruix, M., & Rodríguez-Crespo, I. (2015). DYNLT (Tctex-1) forms a tripartite complex with dynein intermediate chain and RagA, hence linking this small GTPase to the dynein motor. *FEBS Journal*, 282(20), 3945-3958.
- Merino-Gracia, J., García-Mayoral, M. F., & Rodríguez-Crespo, I. (2011). The association of viral proteins with host cell dynein components during virus infection. *FEBS Journal*, 278(17), 2997-3011.

- Millar, A., Pavlos, N., Xu, J., & Zheng, M. (2002). Rab3D a regulator of exocytosis in non-neuronal cells.
- Mok, Y.-K., Lo, K. W.-H., & Zhang, M. (2001). Structure of Tctex-1 and its interaction with cytoplasmic dynein intermediate chain. *Journal of Biological Chemistry*, 276(17), 14067-14074.
- Mueller, S., Cao, X., Welker, R., & Wimmer, E. (2002). Interaction of the poliovirus receptor CD155 with the dynein light chain Tctex-1 and its implication for poliovirus pathogenesis. *Journal of Biological Chemistry*, 277(10), 7897-7904.
- Nagano, F., Orita, S., Sasaki, T., Naito, A., Sakaguchi, G., Maeda, M., . . . Takai, Y. (1998). Interaction of Doc2 with tctex-1, a Light Chain of Cytoplasmic Dynein IMPLICATION IN DYNEIN-DEPENDENT VESICLE TRANSPORT. *Journal of Biological Chemistry*, 273(46), 30065-30068.
- Navarro-Lérida, I., Álvarez-Barrientos, A., & Rodríguez-Crespo, I. (2006). N-terminal palmitoylation within the appropriate amino acid environment conveys on NOS2 the ability to progress along the intracellular sorting pathways. *Journal of cell science*, 119(8), 1558-1569.
- Navarro-Lérida, I., Martínez-Moreno, M., Ventoso, I., Álvarez-Barrientos, A., & Rodríguez-Crespo, I. (2007). Binding of CAP70 to inducible nitric oxide synthase and implications for the vectorial release of nitric oxide in polarized cells. *Molecular biology of the cell*, 18(7), 2768-2777.
- Nyarko, A., & Barbar, E. (2011). Light chain-dependent self-association of dynein intermediate chain. *Journal of Biological Chemistry*, 286(2), 1556-1566.
- Nyarko, A., Hare, M., Hays, T. S., & Barbar, E. (2004). The intermediate chain of cytoplasmic dynein is partially disordered and gains structure upon binding to light-chain LC8. *Biochemistry*, 43(49), 15595-15603.
- Patel-King, R. S., Benashski, S. E., Harrison, A., & King, S. M. (1997). A Chlamydomonas homologue of the putative murine t complex distorter Tctex-2 is an outer arm dynein light chain. *The Journal of cell biology*, 137(5), 1081-1090.
- Pavlos, N. J., Cheng, T. S., Qin, A., Ng, P. Y., Feng, H.-T., Ang, E. S., . . . Zheng, M.-H. (2011). Tctex-1, a novel interaction partner of Rab3D, is required for osteoclastic bone resorption. *Molecular and cellular biology*, 31(7), 1551-1564.
- Pfister, K., Salata, M., Dillman, J., Torre, E., & Lye, R. (1996). Identification and developmental regulation of a neuron-specific subunit of cytoplasmic dynein. *Molecular biology of the cell*, 7(2), 331-343.
- Pfister, K. K., Fisher, E. M., Gibbons, I. R., Hays, T. S., Holzbaur, E. L., McIntosh, J. R., . . . Witman, G. B. (2005). Cytoplasmic dynein nomenclature. *The Journal of cell biology*, 171(3), 411-413.
- Pfister, K. K., Salata, M. W., Dillman, J. F., Vaughan, K. T., Vallee, R. B., Torre, E., & Lye, R. J. (1996). Differential expression and phosphorylation of the 74-kDa intermediate chains of cytoplasmic dynein in cultured neurons and glia. *Journal of Biological Chemistry*, 271(3), 1687-1694.
- Pfister, K. K., Shah, P. R., Hummerich, H., Russ, A., Cotton, J., Annuar, A. A., . . . Fisher, E. M. (2006). Genetic analysis of the cytoplasmic dynein subunit families. *PLoS Genet*, 2(1), e1.
- Rodríguez-Crespo, I., Yélamos, B., Roncal, F., Albar, J. P., de Montellano, P. R. O., & Gavilanes, F. (2001). Identification of novel cellular proteins that bind to the LC8 dynein light chain using a pepscan technique. *FEBS Letters*, 503(2), 135-141.
- Roux, A., Uyhazi, K., Frost, A., & De Camilli, P. (2006). GTP-dependent twisting of dynamin implicates constriction and tension in membrane fission. *Nature*, 441(7092), 528-531.

- doi:http://www.nature.com/nature/journal/v441/n7092/supinfo/nature04718_S1.html
- Sachdev, P., Menon, S., Kastner, D. B., Chuang, J. Z., Yeh, T. Y., Conde, C., . . . Sakmar, T. P. (2007). G protein $\beta\gamma$ subunit interaction with the dynein light-chain component Tctex-1 regulates neurite outgrowth. *The EMBO journal*, 26(11), 2621-2632.
- Sánchez-Ruiloba, L., Aicart-Ramos, C., García-Guerra, L., Pose-Utrilla, J., Rodríguez-Crespo, I., & Iglesias, T. (2014). Protein kinase D interacts with neuronal nitric oxide synthase and phosphorylates the activatory residue serine 1412. *PLoS One*, 9(4), e95191.
- Sarma, N. J., & Yaseen, N. R. (2013). Dynein light chain 1 (DYNLT1) interacts with normal and oncogenic nucleoporins. *PLoS One*, 8(6), e67032.
- Schliwa, M., & Woehlke, G. (2003). Molecular motors. *Nature*, 422(6933), 759-765.
- Sebé-Pedrós, A., Grau-Bové, X., Richards, T. A., & Ruiz-Trillo, I. (2014). Evolution and classification of myosins, a paneukaryotic whole-genome approach. *Genome biology and evolution*, 6(2), 290-305.
- Shen, Y., Delaglio, F., Cornilescu, G., & Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR*, 44(4), 213-223.
- Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual review of biochemistry*, 76(1), 23-50. doi:doi:10.1146/annurev.biochem.76.052305.115300
- Sugai, M., Saito, M., Sukegawa, I., Katsushima, Y., Kinouchi, Y., Nakahata, N., . . . Sukegawa, J. (2003). PTH/PTH-related protein receptor interacts directly with Tctex-1 through its COOH terminus. *Biochemical and biophysical research communications*, 311(1), 24-31.
- Susalka, S. J., Nikulina, K., Salata, M. W., Vaughan, P. S., King, S. M., Vaughan, K. T., & Pfister, K. K. (2002). The roadblock light chain binds a novel region of the cytoplasmic dynein intermediate chain. *Journal of Biological Chemistry*, 277(36), 32939-32946.
- Tai, A. W., Chuang, J.-Z., Bode, C., Wolfrum, U., & Sung, C.-H. (1999). Rhodopsin's carboxy-terminal cytoplasmic tail acts as a membrane receptor for cytoplasmic dynein by binding to the dynein light chain Tctex-1. *Cell*, 97(7), 877-887.
- Tai, A. W., Chuang, J.-Z., & Sung, C.-H. (1998). Localization of Tctex-1, a cytoplasmic dynein light chain, to the Golgi apparatus and evidence for dynein complex heterogeneity. *Journal of Biological Chemistry*, 273(31), 19639-19649.
- Talbott, M., Hare, M., Nyarko, A., Hays, T. S., & Barbar, E. (2006). Folding is coupled to dimerization of Tctex-1 dynein light chain. *Biochemistry*, 45(22), 6793-6800.
- Tang, Q., Staub, C. M., Gao, G., Jin, Q., Wang, Z., Ding, W., . . . Mulder, K. M. (2002). A novel transforming growth factor- β receptor-interacting protein that is also a light chain of the motor protein dynein. *Molecular biology of the cell*, 13(12), 4484-4496.
- Tynan, S. H., Gee, M. A., & Vallee, R. B. (2000). Distinct but overlapping sites within the cytoplasmic dynein heavy chain for dimerization and for intermediate chain and light intermediate chain binding. *Journal of Biological Chemistry*, 275(42), 32769-32774.
- Tzolovsky, G., Millo, H., Pathirana, S., Wood, T., & Bownes, M. (2002). Identification and phylogenetic analysis of *Drosophila melanogaster* myosins. *Molecular biology and evolution*, 19(7), 1041-1052.
- Vaughan, P. S., Leszyk, J. D., & Vaughan, K. T. (2001). Cytoplasmic dynein intermediate chain phosphorylation regulates binding to dynactin. *Journal of Biological Chemistry*, 276(28), 26171-26179.
- Wanschers, B., van de Vorstenbosch, R., Wijers, M., Wieringa, B., King, S. M., & Fransen, J. (2008). Rab6 family proteins interact with the dynein light chain protein DYNLRB1. *Cell motility and the cytoskeleton*, 65(3), 183-196.

- Whitehead, B., Craven, C. J., & Waltho, J. P. (1997). Double and triple resonance NMR methods for protein assignment. *Protein NMR Techniques*, 29-52.
- Wickstead, B., & Gull, K. (2011). The evolution of the cytoskeleton. *The Journal of cell biology*, 194(4), 513-525.
- Wilkins, S. (2005). Rotary Molecular Motors *Advances in Protein Chemistry* (Vol. Volume 71, pp. 345-382): Academic Press.
- Williams, J. C., Roulhac, P. L., Roy, A. G., Vallee, R. B., Fitzgerald, M. C., & Hendrickson, W. A. (2007). Structural and thermodynamic characterisation of a cytoplasmic dynein light chain–intermediate chain complex. *Proceedings of the National Academy of Sciences*, 104(24), 10028-10033.
- Williams, J. C., Xie, H., & Hendrickson, W. A. (2005). Crystal structure of dynein light chain TcTex-1. *Journal of Biological Chemistry*, 280(23), 21981-21986.
- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., . . . Sykes, B. D. (1995). ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *Journal of Biomolecular NMR*, 6(2), 135-140.
- Wu, H., Maciejewski, M. W., Takebe, S., & King, S. M. (2005). Solution structure of the Tctex1 dimer reveals a mechanism for dynein-cargo interactions. *Structure*, 13(2), 213-223.
- Yeh, T.-Y., Chuang, J.-Z., & Sung, C.-H. (2005). Dynein light chain rp3 acts as a nuclear matrix-associated transcriptional modulator in a dynein-independent pathway. *Journal of cell science*, 118(15), 3431-3443.
- Yoshimura, Y., & Miki, H. (2011). Dynamic regulation of GEF-H1 localization at microtubules by Par1b/MARK2. *Biochemical and biophysical research communications*, 408(2), 322-328.

CHAPTER 6
SUMMARY AND
CONCLUDING REMARKS

SUMMARY AND GENERAL CONCLUSIONS

The work presented in this thesis was carried out in the group of Protein Structure, Dynamics and Interactions by NMR, belonging to the Department of Biological Physical Chemistry from Instituto de Química Física Rocasolano (CSIC), from December, 2012 to November, 2016.

Besides the specific conclusions detailed in each of the preceding chapters, the main general conclusions are highlighted here:

1. A series of peptides derived from the choline-binding domain of pneumococcal autolysin LytA were studied to find out whether they maintain their native structure and the ability to bind choline when isolated from the full-length protein. A thorough analysis using solution NMR, CD and fluorescence techniques led to the discovery of three peptides that show native-like, β -hairpin conformations in aqueous solution and undergo a reversible β -to- α transition in the presence of detergent micelles. In addition, these peptides are capable of binding choline with different affinities. A model has been proposed to explain the interaction between these peptides and detergent micelles. Moreover, the physicochemical bases governing the observed structural transition have been characterised by studying variants of one of the peptides, revealing the importance of the amphipathicity and the presence of aromatic residues.
2. To investigate the mechanism of action of BP100, an antimicrobial and cell-penetrating peptide, an approach based on ^2H -Ala scan and solid state NMR was used. Together with CD and microbiological experiments, it has been possible to propose a carpet model of action for BP100 which is responsible for inducing disorder in bacterial membranes, leading to cell death. Moreover, two BP100 variants were found to exhibit improved therapeutic index, with the same antimicrobial strength displayed by the parent peptide and a reduced cytotoxicity.
3. Crotalictidin, an antimicrobial peptide isolated from the venom of a rattlesnake, was dissected into two fragments and their structure and biological properties were extensively analysed in order to better understand the structure-function relationship. Results revealed that one of the fragments, Ctn₁₋₁₄, maintained the helical structure seen in the parent peptide in the presence of detergent micelles, and the other one, Ctn₁₅₋₃₄, displayed a poorly defined structure. Contrary to what is thought, it was demonstrated that the poorly structured fragment retained a high effectiveness against bacteria, whereas the helical fragment did not show any antibacterial activity. In addition, Ctn₁₅₋₃₄ showed good antitumour properties, with a low cytotoxicity against healthy cells, and high serum stability. In conclusion, the rational structural analysis of the crotalictidin fragments by solution NMR, combined with the CD and microbiological experiments, conducted to the finding of a promising therapeutic molecule with good bactericidal and antitumour properties.

4. Two homologous carbohydrate-binding modules (CBMs) from plant allergens, CtD-Ole e 9 (from *O. europaea*) and CtD-Fra e 9 (from *F. exvelsior*) were investigated to find the origin of their different affinity for a β -1,3-glucan substrate, laminarin. The carbohydrate-protein interaction was characterised by biochemical and physicochemical techniques, together with a comprehensive solution NMR analysis, which included structural and relaxation studies. This characterization paved the way to achieve a detailed description of the carbohydrate-protein interaction at atomic level, by combining solution NMR data with a driven molecular docking approach. The result was the proposal of a carbohydrate-binding mechanism for these CBMs, and the identification of the atomic features giving rise to the observed differences in the affinity for the substrate between CtD-Ole e 9 and CtD-Fra e 9. The relevance of slight changes in the primary structure was brought into light, as the pattern of interactions established between the amino acids and the substrate may result in significant variations in the affinity.
5. The interaction of human DYNLT1, a dynein light chain, and DIC, a dynein intermediate chain, components of the cytoskeletal dynein motor complex, has been examined in detail. The binding of DIC to DYNLT1 through the canonical binding site has been structurally characterised at residue level by solution NMR, using a chimera designed for this purpose. The chimeric protein complex, a dimer, shows many characteristics that are common in other known members of the family. Many other potential DYNLT1 binding partners were identified using biochemical approaches, and one of them, ActRIIB, was proved to bind DYNLT1, suggesting a possible role of this dynein light chain in the signalling pathway of TGF- β . The study of the interaction of protein Lfc with DYNLT1 points to the fact that some contacts necessary to the binding process can involve residues located outside the canonical binding groove. With all these data, an interaction model of DYNLT1 with different partners has been proposed. This model takes into account the fact that the interaction may imply contacts with regions from the canonical binding groove and from proximate zones.

CONCLUDING REMARKS

Throughout this dissertation, the versatility and the usefulness of NMR techniques have been evinced. It is very well known that NMR spectroscopy is a powerful tool to obtain detailed information of molecular recognition events and interactions at atomic level. This kind of information is crucial to understand the molecular bases behind innumerable biological and biochemical processes which may potentially have a huge impact in our lives.

In the preceding pages, I tried to highlight the outstanding performance of NMR spectroscopy by making use of an extensive repertoire of NMR techniques, and to get as much information as possible from the studied systems by NMR. However, it has been also demonstrated that the utilization of a single technique, even one as powerful as NMR, is not enough to achieve a complete study of the proposed systems and it is necessary the

concurrence of many different techniques to get a complete picture by a multidisciplinary approach.

In conclusion, the work included in this thesis is focused in the utilization of NMR techniques for the study of biomolecular recognition processes and interactions. Together with NMR data, the results obtained by other means, such as biochemical, physicochemical or computational techniques, and the work performed by our collaborators, have shed some light on open questions that, hopefully, will help other scientists to keep on moving in the path of knowledge.

PUBLICATIONS

Most of the work included in this thesis has been published in scientific journals:

- Zamora-Carreras, H., B. Maestro, E. Strandberg, A. S. Ulrich, J. M. Sanz and M. Á. Jiménez (2015). "Micelle-Triggered β -Hairpin to α -Helix Transition in a 14-Residue Peptide from a Choline-Binding Repeat of the Pneumococcal Autolysin LytA." Chemistry – A European Journal **21**(22): 8076-8089.
- Zamora-Carreras, H., E. Strandberg, P. Mühlhäuser, J. Bürck, P. Wadhwani, M. Ángeles Jiménez, M. Bruix and A. S. Ulrich (2016). "Alanine scan and ^2H -NMR analysis of the membrane-active peptide BP100 point to a distinct carpet mechanism of action." Biochimica et Biophysica Acta **1858**: 1328-1338.
- Borges Falcao, C., C. Pérez-Peinado, B. G. de la Torre, X. Mayol, H. Zamora-Carreras, M. Ángeles Jiménez, G. Rádis-Baptista and D. Andreu (2015). "Structural dissection of crotalicidin, a rattlesnake venom cathelicidin, retrieves a fragment with antimicrobial and antitumor activity." Journal of Medicinal Chemistry **58**(21): 8553-8563.
- Zamora-Carreras, H., M. Torres, N. Bustamante, A. L. Macedo, R. Rodríguez, M. Villalba and M. Bruix (2015). "The C-terminal domain of two homologous *Oleaceae* β -1,3-glucanases recognise carbohydrates differently: Laminarin binding by NMR." Archives of Biochemistry and Biophysics **580**: 93-101.
- Merino-Gracia, J., H. Zamora-Carreras, M. Bruix and I. Rodríguez-Crespo (2016). "Molecular basis for the protein recognition specificity of the dynein light chain DYNLT1/Tctex1. Characterization of the interaction with activin receptor IIB". Journal of Biological Chemistry **291**(40): 20962-20975.

Additional manuscripts derived from this thesis are in preparation to be published soon:

- Zamora-Carreras, H., B. Maestro, E. Strandberg, A. S. Ulrich, J. M. Sanz and M. Á. Jiménez. "Roles of amphipathicity, hydrophobic patches and Trp residues in the chameleonic structural behaviour of a 14-mer peptide." *In preparation*.
- Zamora-Carreras, H., M. Mompéan, R. Rodríguez, M. Villalba and M. Bruix. "Insights into protein-carbohydrate recognition: A novel binding mechanism for CBM family 43" *In preparation*.
- Zamora-Carreras, H., B. Maestro, E. Strandberg, A. S. Ulrich, J. M. Sanz and M. Á. Jiménez. "Choline binding and structural switch behaviour of peptides derived from two consecutive choline-binding repeats of the pneumococcal autolysin LytA." *In preparation*.

APPENDICES

TABLES

Table A1. ¹H and ¹³C chemical shifts (ppm from DSS) of LytA₂₃₉₋₂₅₂ under different solvent conditions at pH 3.0 and 25 °C. “nd” stands for not determined. ^aMeasured at 35°C. Simplified numbering in parentheses.

Condition		H ₂ O/D ₂ O 9:1 (v/v)	30 % vol. [D ₃]-TFE	0.45 mM [D ₃₈]-DPC	30 mM [D ₃₈]-DPC	30 mM [D ₂₅]- SDS	0.2 mM [D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	β-hairpin	α-helix	α-helix	β-hairpin
Residue	Resonance						
Thr 239 (Thr1)	C _α H	3.90	3.96	3.88	3.85	3.89	3.88
	¹³ C _α	61.2	61.8	61.6	63.6	63.7	61.4
	C _β H	4.07	4.11	4.04	4.01	4.07	4.05
	¹³ C _β	69.5	69.7	69.8	68.5	69.1	69.8
	C _γ H ₃	1.33	1.38	1.32	1.14	1.15	1.33
	¹³ C _γ	21.4	21.6	21.7	21.0	21.3	21.6
Gly 240 (Gly2)	H _N	8.54	8.57	8.56	9.02	8.65	8.54
	C _{αα} H	3.69, 4.48	3.73, 4.45	3.68, 4.47	3.92, 4.18	3.93, 4.07	3.66, 4.50
	¹³ C _α	44.5	44.7	44.5	45.5	46.4	44.6
Trp 241 (Trp3)	H _N	8.69	8.45	8.72	8.39	7.85	8.70
	C _α H	5.45	5.51	5.43	4.47	4.46	5.44
	¹³ C _α	57.5	57.6	57.7	59.2	59.1	57.7
	C _β H	3.19, 3.51	3.20, 3.56	3.19, 3.49	3.27, 3.29	3.32, 3.32	3.18, 3.50
	¹³ C _β	31.1	31.5	31.3	29.1 ^a	29.2	31.3
	C _{δ1} H	7.41	7.40	7.42	7.44	7.40	7.43
	¹³ C _{δ1}	127.3	127.1	127.6	127.7	127.8	127.4
	N _{ε1} H	10.21	9.96	10.25	10.70	10.00	10.22
	C _{ε3} H	7.68	7.72	7.66	7.41	7.44	7.68
	¹³ C _{ε3}	120.5	120.8	120.7	120.7	120.6	120.6
	C _{ζ3} H	7.03	7.05	7.02	6.91	6.94	7.04
	¹³ C _{ζ3}	122.1	122.1	122.3	121.4	121.7	122.3
	C _{η2} H	7.20	7.18	7.19	7.07	7.09	7.21
	¹³ C _{η2}	124.9	124.7	125.1	124.1	124.4	125.1
	C _{ζ2} H	7.34	7.33	7.35	7.48	7.48	7.35
	¹³ C _{ζ2}	115.0	114.8	115.2	115.0	114.9	115.2
Lys 242 (Lys4)	H _N	9.92	9.92	9.98	8.31	7.74	9.90
	C _α H	4.86	4.86	4.86	3.82	3.84	4.86
	¹³ C _α	55.9	55.9	56.0	59.5	59.2	56.1
	C _{ββ} H	1.90, 1.90	1.90, 1.94	1.85, 1.91	1.62, 1.73	1.60, 1.74	1.86, 1.92
	¹³ C _β	36.9	37.0	37.1	31.8 ^a	32.1	37.1
	C _{γγ} H	1.43, 1.43	1.44, 1.44	1.41, 1.41	1.08, 1.14	1.07, 1.07	1.42, 1.42
	¹³ C _γ	25.0	25.1	25.2	24.8 ^a	25.2	25.2
	C _{δδ} H	1.47, 1.54	1.53, 1.63	1.47, 1.53	1.62, 1.62	1.60, 1.60	1.46, 1.53
	¹³ C _δ	29.2	29.0	29.3	28.7	29.2	29.3
	C _{εε} H	2.74, 2.74	2.81, 2.81	2.74, 2.74	2.92, 2.92	2.94, 2.94	2.72, 2.72
	¹³ C _ε	41.7	42.0	41.9	41.6	42.3	41.9
	N _ζ H ₃	nd	nd	7.39	7.74	7.43	7.40
Lys 243 (Lys5)	H _N	8.51	8.32	8.52	7.68	7.72	8.53
	C _α H	4.47	4.53	4.45	4.11	4.11	4.46
	¹³ C _α	55.2	55.4	55.3	58.0	58.6	55.4
	C _{ββ} H	-0.19, 1.18	-0.10, 1.25	-0.19, 1.18	1.83, 1.83	1.87, 1.87	-0.25, 1.17
	¹³ C _β	32.1	32.3	32.3	31.8 ^a	32.2	32.2
	C _{γγ} H	-0.04, 0.60	0.08, 0.71	-0.05 0.60	1.42, 1.42	1.45, 1.50	-0.07, 0.58
	¹³ C _γ	24.5	24.7	24.8	24.5 ^a	25.2	24.7
	C _{δδ} H	1.15, 1.15	1.20, 1.20	1.16, 1.16	1.69, 1.69	1.71, 1.71	1.13, 1.13
	¹³ C _δ	29.1	29.3	29.4	28.5	29.2	29.3
	C _{εε} H	2.50, 2.62	2.59, 2.69	2.49, 2.62	2.96, 2.96	2.95, 2.95	2.48, 2.61
	¹³ C _ε	41.5	41.8	41.9	41.4	42.3	41.6
	N _ζ H ₃	7.48	7.54	7.48	nd	7.43	7.48

Table A1 (continuation)

Condition		H ₂ O/D ₂ O 9:1 (v/v)	30 % vol. [D ₃]-TFE	0.45 mM [D ₃₈]-DPC	30 mM [D ₃₈]-DPC	30 mM [D ₂₅]- SDS	0.2 mM [D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	β-hairpin	α-helix	α-helix	β-hairpin
Residue	Resonance						
Ile 244 (Ile6)	H _N	9.06	9.04	9.09	7.60	7.62	9.07
	C _α H	4.17	4.18	4.15	3.86	3.91	4.15
	¹³ C _α	60.9	61.1	61.3	60.9	61.4	61.1
	C _β H	1.95	2.00	1.94	2.03	2.09	1.94
	¹³ C _β	39.7	39.7	39.9	37.6	38.2	39.9
	C _{γ2} H ₃	0.88	0.90	0.88	0.93	1.01	0.88
	¹³ C _{γ2}	17	16.7	17.3	17.4	18.0	17.2
	C _{γγ1} H	1.10, 1.50	1.12, 1.54	1.10, 1.50	1.20, 1.63	1.29, 1.68	1.10, 1.50
	¹³ C _{γ1}	27.1	27.2	27.3	28.2	28.6	27.2
	C _{δ1} H ₃	0.76	0.78	0.76	0.80	0.92	0.75
	¹³ C _{δ1}	12.9	12.5	13.2	12.9	13.4	13.1
	H _N	9.20	8.97	9.23	8.26	8.07	9.23
	C _α H	3.93	3.89	3.94	4.09	4.12	3.92
	¹³ C _α	53.7	54.0	54.0	54.4	54.8	53.9
Ala 245 (Ala7)	C _β H ₃	1.49	1.54	1.49	1.54	1.53	1.49
	¹³ C _β	16.8	16.6	17.1	18.1	18.6	17.0
	H _N	8.52	8.47	8.53	8.44	8.21	8.52
	C _α H	4.33	4.43	4.32	4.63	4.63	4.31
Asp 246 (Asp8)	¹³ C _α	54.2	54.0	54.5	55.1	nd	54.5
	C _{ββ} H	2.92, 3.03	3.00, 3.00	2.91, 3.01	2.98, 2.98	2.97, 3.02	2.90, 3.00
	¹³ C _β	38.8	38.6	39.1	37.8	37.8	39.2
	H _N	7.88	7.94	7.87	7.80	7.79	7.89
Lys 247 (Lys9)	C _α H	4.62	4.63	4.60	4.19	4.17	4.61
	¹³ C _α	55.2	55.5	55.3	57.5	57.8	55.3
	C _{ββ} H	1.75, 1.75	1.75, 1.84	1.72, 1.77	1.82, 1.90	1.74, 1.84	1.71, 1.76
	¹³ C _β	36.0	36.1	36.2	32.0 ^a	32.2	36.3
	C _{γγ} H	1.27, 1.45	1.32, 1.50	1.26, 1.45	1.28, 1.42	1.27, 1.37	1.27, 1.46
	¹³ C _γ	25.2	25.3	25.4	24.4 ^a	24.8	25.4
	C _{δδ} H	1.73, 1.73	1.77, 1.77	1.72, 1.72	1.62, 1.62	1.59, 1.59	1.71, 1.71
	¹³ C _δ	29.0	29.1	29.3	28.5	28.9	29.1
	C _{εε} H	2.99, 2.99	3.03, 3.03	2.99, 2.99	2.85, 2.85	2.90, 2.90	2.99, 2.99
	¹³ C _ε	42.1	42.4	42.5	41.4	42.4	42.3
	N _ε H ₃	7.56	7.63	7.57	7.72	7.47	7.56
	H _N	8.39	8.19	8.40	7.99	7.96	8.41
	C _α H	5.07	5.13	5.04	4.50	4.57	5.06
	¹³ C _α	57.4	57.5	57.5	58.5	58.0	57.5
Trp 248 (Trp10)	C _{ββ} H	2.99, 3.31	3.02, 3.37	3.01, 3.30	3.24, 3.36	3.26, 3.37	2.99, 3.31
	¹³ C _β	30.9	31.1	31.1	29.4 ^a	30.0	31.1
	C _{δ1} H	7.22	7.21	7.21	7.14	7.03	7.21
	¹³ C _{δ1}	127.2	127.1	127.5	126.8	126.9	127.1
	N _{ε1} H	10.11	9.96	10.13	10.51	9.89	10.12
	C _{ε3} H	7.63	7.66	7.62	7.54	7.53	7.63
	¹³ C _{ε3}	120.1	120.5	120.3	120.8	120.8	120.2
	C _{ζ3} H	7.08	7.09	7.07	6.97	6.99	7.10
	¹³ C _{ζ3}	122.1	122.1	122.4	121.6	121.9	122.3
	C _{η2} H	7.29	7.26	7.27	7.08	7.11	7.31
	¹³ C _{η2}	124.8	124.6	125.0	124.0	124.5	125.0
	C _{ζ2} H	7.50	7.49	7.50	7.46	7.46	7.51
	¹³ C _{ζ2}	115.0	114.8	115.2	114.8	114.8	115.2

Table A1 (continuation)

Condition		H ₂ O/D ₂ O 9:1 (v/v)	30 % vol. [D ₃]-TFE	0.45 mM [D ₃₈]-DPC	30 mM [D ₃₈]-DPC	30 mM [D ₂₅]- SDS	0.2 mM [D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	β-hairpin	α-helix	α-helix	β-hairpin
Residue	Resonance						
Tyr 249 (Tyr11)	H _N	9.45	9.45	9.49	8.01	7.90	9.44
	C _α H	4.81	nd	nd	4.12	4.23	nd
	¹³ C _α	57.5	nd	nd	59.6	59.4	nd
	C _{ββ'} H	2.77, 2.77	2.78, 2.84	2.75, 2.75	2.78, 2.90	2.89, 2.89	2.76, 2.76
	¹³ C _β	42.6	42.7	42.7	38.6 ^a	38.5	42.9
	C _{δδ'} H	7.16, 7.16	7.15, 7.15	7.15, 7.15	6.60, 6.60	6.86, 6.86	7.18, 7.18
	¹³ C _δ	133.8	133.7	134.0	133.2	133.3	134.0
	C _{εε'} H	6.91, 6.91	6.90, 6.90	6.90, 6.90	6.70, 6.70	6.80, 6.80	6.92, 6.92
	¹³ C _ε	118.2	118.2	118.5	118.3	118.4	118.4
Tyr 250 (Tyr12)	H _N	8.46	8.17	8.48	7.62	7.47	8.48
	C _α H	4.09	4.23	4.08	4.24	4.23	4.06
	¹³ C _α	57.4	57.5	57.8	58.5	59.3	57.6
	C _{ββ'} H	0.94, 2.18	1.16, 2.25	0.92, 2.19	2.70, 2.88	2.74, 2.81	0.87, 2.16
	¹³ C _β	38.5	38.8	38.8	38.0 ^a	38.4	38.7
	C _{δδ'} H	5.78, 5.78	5.87, 5.87	5.78, 5.78	7.02, 7.02	6.84, 6.84	5.75, 5.75
	¹³ C _δ	132.6	132.8	132.8	133.4	133.3	132.7
	C _{εε'} H	6.54, 6.54	6.55, 6.55	6.55, 6.55	6.82, 6.82	6.77, 6.77	6.54, 6.54
	¹³ C _ε	117.5	117.6	117.8	118.4	118.4	117.6
Phe 251 (Phe13)	H _N	7.77	7.93	7.76	7.61	7.62	7.79
	C _α H	4.46	4.48	4.45	4.51	4.45	4.44
	¹³ C _α	56.8	57.0	56.9	57.3	58.3	57.0
	C _{ββ'} H	2.75, 3.01	2.78, 3.02	2.75, 3.01	3.03, 3.18	2.98, 3.16	2.73, 3.00
	¹³ C _β	40.3	40.3	40.4	39.0	39.9	40.4
	C _{δδ'} H	7.26, 7.26	7.27, 7.27	7.25, 7.25	7.28, 7.28	7.32, 7.32	7.29, 7.29
	¹³ C _δ	132.1	132.1	132.3	132.4	132.4	132.3
	C _{εε'} H	7.09, 7.09	7.06, 7.06	7.08, 7.08	7.20, 7.20	7.25, 7.25	7.10, 7.10
	¹³ C _ε	129.3	129.2	129.5	129.7	129.9	129.4
	C _ζ H	7.25	7.24	7.24, 7.24	7.28	7.33	7.25
Asn 252 (Asn14)	¹³ C _ζ	131.3	131.2	131.5	131.4	131.5	131.5
	H _N	8.20	8.08	8.20	7.98	7.96	8.25
	C _α H	4.38	4.45	4.37	4.56	4.58	4.34
	¹³ C _α	54.5	54.4	54.9	52.1	51.3	55.0
	C _{ββ'} H	2.68, 2.77	2.70, 2.80	2.67, 2.77	2.61, 2.71	2.43, 2.59	2.66, 2.77
	¹³ C _β	39.5	39.6	39.8	39.0	39.3	39.8
	N _{δδ'} H ₂	6.95, 7.58	6.86, 7.58	6.95, 7.60	6.75, 7.44	6.62, 7.15	6.97, 7.61

Table A2. ^1H and ^{13}C chemical shifts (ppm from DSS) of LytA₂₅₉₋₂₇₂ under different solvent conditions at pH 3.0 and 25 °C. “nd” stands for not determined. ^aMeasured at 35°C. ^bIn pure D₂O. Simplified numbering in parentheses.

Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE	[D ₃₈]-DPC ^a	[D ₂₅]-SDS ^b
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Thr 259 (Thr1)	C _α H	3.86 ^b	3.83	3.94	nd
	¹³ C _α	61.14 ^b	61.66	60.89 ^b	3.94
	C _β H	4.08 ^b	4.04	61.85	61.44
	¹³ C _β	69.26 ^b	69.63	4.16	4.15
	C _γ H ₃	1.26 ^b	1.24	69.39	69.10
	¹³ C _γ	21.42 ^b	21.59	1.26	1.25
Gly 260 (Gly2)	H _N	8.43	8.34	8.84	nd
	C _{αα} H	3.58, 4.33 ^b	3.55, 4.30	4.03, 4.03	4.02, 4.02
	¹³ C _α	44.49 ^b	44.86	45.99	45.62
Trp 261 (Trp3)	H _N	8.40	8.20	8.52	nd
	C _α H	5.18 ^b	5.32	4.75 ^b	4.15
	¹³ C _α	57.45 ^b	57.76	nd	nd
	C _β H	3.09, 3.35 ^b	3.08, 3.41	3.28, 3.40	3.28, 3.41
	¹³ C _β	30.70 ^b	31.28	30.48	29.87 ^b
	C _{δ1} H	7.28 ^b	7.28	7.38	7.38
	¹³ C _{δ1}	127.09 ^b	127.06	127.67	127.49
	N _{ε1} H	10.06	9.85	10.59	nd
	C _{ε3} H	7.57 ^b	7.63	7.56	7.56
	¹³ C _{ε3}	120.54 ^b	120.80	121.11	120.99
	C _{ζ3} H	7.03 ^b	7.00	6.95	6.92
	¹³ C _{ζ3}	122.04 ^b	122.10	121.82	121.64
	C _{η2} H	7.18 ^b	7.12	7.08	7.07
	¹³ C _{η2}	124.77 ^b	124.69	124.40	124.18
	C _{ζ2} H	7.30 ^b	7.25	7.47	7.47
	¹³ C _{ζ2}	114.84 ^b	114.76	115.04	114.91
Val 262 (Val4)	H _N	9.47	9.64	7.99	nd
	C _α H	4.50 ^b	4.57	3.96	3.89
	¹³ C _α	61.00 ^b	nd	60.93 ^b	64.46 ^b
	C _β H	1.86 ^b	1.99	2.06	2.09
	¹³ C _β	35.41 ^b	35.82	32.96	32.24 ^b
	C _γ H ₃	0.81 ^b	0.90	0.89	0.91
	¹³ C _γ	20.56 ^b	20.47	21.47	21.25
	C _{γ'} H ₃	0.84 ^b	0.90	0.92	0.95
	¹³ C _{γ'}	21.18 ^b	21.14	21.71	21.72
Lys 263 (Lys5)	H _N	8.27	8.06	nd	nd
	C _α H	4.36 ^b	4.42	4.20 ^b	4.12 ^b
	¹³ C _α	55.11 ^b	55.33	nd	57.554
	C _{ββ} H	nd, 1.20 ^b	0.11, 1.19	1.55, 1.55	1.54, 1.54
	¹³ C _β	32.68 ^b	32.83	32.82	32.67 ^b
	C _{γγ} H	nd, 0.64 ^b	0.13, 0.62	1.03, 1.07	1.06, 1.10
	¹³ C _γ	24.54 ^b	24.71	24.93	24.67 ^b
	C _{δδ} H	1.15, 1.15 ^b	1.08, 1.08	1.50, 1.50	1.51, 1.51
	¹³ C _δ	28.99 ^b	29.27	29.52	29.32 ^b
	C _{εε} H	2.51, 2.64 ^b	2.47, 2.58	2.82, 2.82	2.81, 2.81
	¹³ C _ε	41.55 ^b	42.05	42.40	42.17 ^b
	N _ζ H ₃	7.44	7.41	nd	nd

Table A2 (continuation)		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE	[D ₃₈]-DPC ^a	[D ₂₅]-SDS ^b
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Tyr 264 (Tyr6)	H _N	8.70	8.79	nd	nd
	C _α H	4.92 ^b	4.63	4.21 ^b	4.47 ^b
	¹³ C _α	57.28 ^b	57.50	56.87 ^b	59.40 ^b
	C _β H	2.91, 2.91 ^b	2.82, 2.87	3.26, 3.26	3.26, 3.26
	¹³ C _β	40.08 ^b	40.48	37.82 ^b	nd
	C _{δδ} H	7.06, 7.06 ^b	7.04, 7.04	6.99, 6.99	6.93, 6.93
	¹³ C _{δδ}	133.15, 133.15 ^b	133.26, 133.26	133.43, 133.43	133.43, 133.43
	C _{εε} H	6.75, 6.75 ^b	6.75, 6.75	6.89, 6.89	6.89, 6.89
	¹³ C _{εε}	118.18, 118.18 ^b	118.22, 118.22	118.57, 118.57	118.57, 118.57
	¹³ C _γ	21.18 ^b	21.14	21.71	21.72
	H _N	8.84	8.67	nd	nd
	C _α H	3.65 ^b	3.47	3.91	3.91
	¹³ C _α	58.04 ^b	58.48	58.31 ^b	58.45 ^b
Lys 265 (Lys7)	C _β H	1.73, 1.73 ^b	1.63, 1.68	1.74, 1.79	1.74, 1.79
	¹³ C _β	30.09 ^b	30.01	32.52	32.60 ^b
	C _{γγ} H	nd, 0.99 ^b	0.64, 0.86	1.27, 1.27	1.27, 1.27
	¹³ C _{γγ}	24.95 ^b	25.06	25.39	25.15 ^b
	C _{δδ} H	1.53, 1.53 ^b	1.47, 1.47	1.67, 1.67	1.67, 1.67
	¹³ C _{δδ}	29.16 ^b	29.50	29.89	29.41 ^b
	C _{εε} H	2.89, 2.89 ^b	2.82, 2.84	2.94, 2.94	2.94, 2.94
	¹³ C _{εε}	41.94 ^b	42.35	42.45	nd
	N _ε H	7.50	7.51	nd	nd
	H _N	8.47	8.33	8.32	8.33
	C _α H	4.52 ^b	4.41	4.58 ^b	4.52 ^b
	¹³ C _α	53.86 ^b	54.28	54.25 ^b	54.31 ^b
	C _{ββ} H	2.80, 3.01 ^b	2.79, 2.96	2.94, 2.94	2.91, 2.95 ^b
	¹³ C _β	39.02 ^b	39.05	38.66	37.71 ^b
Thr 267 (Thr9)	H _N	7.99	7.90	7.85	7.87
	C _α H	4.54 ^b	4.48	4.11 ^b	4.14 ^b
	¹³ C _α	61.67 ^b	62.01	nd	64.25 ^b
	C _β H	4.14 ^b	4.08	4.15	4.14
	¹³ C _β	71.20 ^b	71.64	69.92	69.62 ^b
	C _γ H ₃	1.09 ^b	1.03	1.18	1.18
	¹³ C _γ	21.07 ^b	20.93	22.36	nd
Trp 268 (Trp10)	H _N	8.46	8.30	nd	nd
	C _α H	5.19 ^b	5.22	4.53 ^b	4.55 ^b
	¹³ C _α	57.20 ^b	57.38	nd	58.78 ^b
	C _{ββ} H	3.04, 3.30 ^b	2.99, 3.32	3.26, 3.40	3.26, 3.40
	¹³ C _β	30.95 ^b	31.20	30.25	29.82 ^b
	C _{δ1} H	7.16 ^b	7.12	7.00	6.99
	¹³ C _{δ1}	127.09 ^b	127.05	127.47	126.96 ^b
	N _{ε1} H	10.04	9.83	10.24	nd
	C _{ε3} H	7.57 ^b	7.58	7.48	7.49
	¹³ C _{ε3}	nd	120.79	121.07	120.85 ^b
	C _{ζ3} H	7.04 ^b	6.98	6.99	6.98
	¹³ C _{ζ3}	nd	122.06	121.87	121.92 ^b
	C _{η2} H	7.23 ^b	7.16	7.10	7.10
	¹³ C _{η2}	124.69 ^b	124.57	124.42	124.53 ^b
	C _{ζ2} H	7.45 ^b	7.38	7.46	7.46
	¹³ C _{ζ2}	114.76 ^b	114.58	115.03	114.83 ^b

Table A2 (continuation)					
Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE	[D ₃₈]-DPC ^a	[D ₂₅]-SDS ^b
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Tyr 269 (Tyr11)	H _N	9.38	9.58	nd	nd
	C _α H	nd	4.86	4.5 ^b	4.24 ^b
	¹³ C _α	nd	nd	58.58 ^b	59.15 ^b
	C _{ββ} H	2.84, 2.95 ^b	2.78, 2.90	2.99, 3.14	2.99, 3.14
	¹³ C _β	41.37 ^b	41.89	38.30 ^b	38.40 ^b
	C _{δδ} H	7.09, 7.09 ^b	7.07, 7.07	7.10, 7.10	7.11, 7.11
	¹³ C _δ	133.61, 133.61 ^b	133.60, 133.60	133.19, 133.19	133.19, 133.19
	C _{εε} H	6.85, 6.85 ^b	6.84, 6.84	6.79, 6.79	6.77, 6.77
	¹³ C _ε	118.16, 118.16 ^b	118.19, 118.19	118.56, 118.56	118.56, 118.56
Tyr 270 (Tyr12)	H _N	8.45	8.22	nd	nd
	C _α H	4.09 ^b	4.14	4.35 ^b	4.34 ^b
	¹³ C _α	57.37 ^b	57.62	58.35 ^b	59.16 ^b
	C _{ββ} H	1.35, 2.42 ^b	1.24, 2.37	2.84, 2.84	2.84, 2.84
	¹³ C _β	38.25 ^b	38.47	38.31 ^b	nd
	C _{δδ} H	5.92, 5.92 ^b	5.84, 5.84	6.89, 6.89	6.89, 6.89
	¹³ C _δ	132.75, 132.75 ^b	132.85, 132.85	133.36, 133.36	133.36, 133.36
	C _{εε} H	6.54, 6.54 ^b	6.46, 6.46	6.82, 6.82	6.81, 6.81
	¹³ C _ε	117.52, 117.52 ^b	117.50, 117.50	118.59, 118.59	118.59, 118.59
Leu 271 (Leu13)	H _N	7.97	8.01	7.53	7.53
	C _α H	4.33 ^b	4.31	4.25	4.25
	¹³ C _α	53.86 ^b	54.05	55.44	55.50 ^b
	C _β H	1.42, 1.42 ^b	1.42, 1.42	1.52, 1.62	1.52, 1.62
	¹³ C _β	42.57 ^b	42.74	43.27	45.87 ^b
	C _γ H	1.48 ^b	1.51	1.73	1.73
	¹³ C _γ	26.85 ^b	27.07	27.32	27.28 ^b
	C _δ H ₃	0.74 ^b	0.74	0.85	0.85
	¹³ C _δ	23.66 ^b	23.40	24.17	23.76 ^b
	C _δ H ₃	0.79 ^b	0.78	0.88	0.88
	¹³ C _δ	25.01 ^b	24.97	26.00	25.85 ^b
Asp 272 (Asp14)	H _N	8.06	7.79	7.70	7.70
	C _α H	4.43 ^b	4.37	4.55	4.55
	¹³ C _α	53.48 ^b	53.51	52.56	52.01 ^b
	C _β H	2.75, 2.82 ^b	2.70, 2.77	2.50, 2.63	2.50, 2.63
	¹³ C _β	39.09 ^b	39.27	39.26	38.29 ^b

Table A3. ^1H and ^{13}C chemical shifts (ppm from DSS) of LytA₂₃₉₋₂₇₂ under different solvent conditions at pH 3.0 and 35 °C. “nd” stands for not determined. ^bIn pure D₂O. ^cIn 30 % vol. [D₃]-TFE only the assignment of the backbone was performed.

Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE ^c	[D ₃₈]-DPC	[D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Thr 239 (Thr1)	C _α H	3.89	3.98	3.92	3.89
	¹³ C _α	61.28 ^b	61.63	61.81	61.76
	C _β H	4.17	4.21	4.11	4.06
	¹³ C _β	68.95 ^b	69.25	69.14	69.41
	C _γ H ₃	1.30	1.36	1.18	1.14
Gly 240 (Gly2)	¹³ C _γ	21.62 ^b	21.55	21.88	21.68
	H _N	8.58	nd	9.04	8.59
	C _{αα} H	3.81, 4.16	4.02, 4.02	3.92, 4.16	3.89, 4.06
	¹³ C _α	44.89 ^b	45.43	46.70	46.64 ^b
Trp 241 (Trp3)	H _N	8.28	nd	8.38	7.80
	C _α H	5.11	4.94	4.43	4.48
	¹³ C _α	57.34 ^b	57.95	59.68 ^b	59.39 ^b
	C _β H	3.18, 3.36	3.29, 3.39	3.29, 3.29	3.31, 3.31
	¹³ C _β	30.68 ^b	30.45	29.82	29.68
	C _{δ1} H	7.29	nd	7.42	7.37
	¹³ C _{δ1}	127.04 ^b	nd	127.97	127.90
	N _{ε1} H	9.88	nd	10.66	9.96
	C _{ε3} H	7.57	nd	7.43	7.47
	¹³ C _{ε3}	120.85 ^b	nd	120.83	120.78
	C _{ζ3} H	6.99	nd	6.92	7.03
	¹³ C _{ζ3}	121.96 ^b	nd	121.62	121.99
	C _{η2} H	7.07	nd	7.07	7.12
	¹³ C _{η2}	124.62 ^b	nd	124.33	124.58
	C _{ζ2} H	7.28	nd	7.49	7.46
	¹³ C _{ζ2}	114.60 ^b	nd	115.05	114.97
Lys 242 (Lys4)	H _N	nd	nd	8.15	7.76
	C _α H	4.64 ^b	4.35	3.82	3.88
	¹³ C _α	55.96 ^b	57.10	60.16	59.59
	C _{ββ} H	1.71, 1.80	1.75, 1.75	1.67, 1.76	1.66, 1.77
	¹³ C _β	35.69 ^b	32.83	32.39	32.42
	C _{γγ} H	1.34, 1.34	nd	1.17, 1.50	1.14, 1.18
	¹³ C _γ	24.87 ^b	nd	25.54	25.53
	C _{δδ} H	1.56, 1.56	nd	1.67, 1.67	1.63, 1.63
	¹³ C _δ	29.12 ^b	nd	29.48	29.46
	C _{εε} H	2.85, 2.85	nd	2.96, 2.96	2.97, 2.97
	¹³ C _ε	41.90 ^b	nd	42.45	42.71
	N _ε H ₃	nd	nd	nd	7.42
Lys 243 (Lys4)	H _N	8.29	nd	7.68	7.72
	C _α H	4.38	4.34	4.14	4.12
	¹³ C _α	55.66 ^b	57.78	58.43	59.15
	C _{ββ} H	0.65, 1.47	nd, 1.48	1.87, 1.87	1.87, 1.92
	¹³ C _β	32.58 ^b	33.01	32.55	32.59
	C _{γγ} H	0.92, 0.92	nd	1.46, 1.53	1.45, 1.53
	¹³ C _γ	24.77 ^b	nd	25.54	25.55
	C _{δδ} H	1.38, 1.38	nd	1.72, 1.72	1.72, 1.72
	¹³ C _δ	29.15 ^b	nd	29.49	29.55
	C _{εε} H	2.72, 2.77	nd	2.99, 2.99	3.00, 3.00
	¹³ C _ε	41.69 ^b	nd	42.53	42.66
	N _ε H ₃	nd	nd	nd	7.42

Table A3 (continuation)					
Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE ^c	[D ₃₈]-DPC	[D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Ile 244 (Ile6)	H _N	8.56	nd	7.55	7.61
	C _α H	4.14	4.16	3.86	3.89
	¹³ C _α	60.02 ^b	62.07	64.15	64.21
	C _β H	1.85	1.95	2.04	2.08
	¹³ C _β	39.42 ^b	39.28	38.47	38.44
	C _{γ2} H ₃	0.85	0.94	0.94	0.99
	¹³ C _{γ2}	17.24 ^b	17.41	18.24	18.31
	C _{γγ1} H	1.06, 1.42	1.19, 1.53	1.22, 1.64	1.26, 1.68
	¹³ C _{γ1}	27.15 ^b	27.74	29.00	27.35
	C _{δ1} H ₃	0.70	0.84	0.83	0.90
	¹³ C _{δ1}	12.89 ^b	12.81	13.60	13.75
	H _N	8.72	nd	8.16	8.07
	C _α H	4.10	4.21	4.05	4.09
	¹³ C _α	53.35 ^b	53.82	55.39	55.38
Ala 245 (Ala7)	C _β H ₃	1.44	1.48	1.50	1.52
	¹³ C _β	17.79 ^b	18.43	18.85	18.83
	H _N	8.33	nd	8.34	8.24
	C _α H	4.49	4.64	4.59	4.58
Asp 246 (Asp8)	¹³ C _α	54.09 ^b	54.15	53.32	55.46 ^b
	C _{ββ} H	2.89, 2.89	2.94, 2.94	2.93, 2.96	2.99, 3.05
	¹³ C _β	41.97 ^b	nd	38.66	38.16
	H _N	7.96	nd	7.82	7.84
Lys 247 (Lys9)	C _α H	4.39	4.29	4.21	4.23
	¹³ C _α	55.95 ^b	57.88	58.20	57.10
	C _{ββ} H	1.60, 1.60	1.81, 1.81	1.86, 1.98	1.88, 1.99
	¹³ C _β	34.84 ^b	33.56	32.59	32.59
	C _{γγ} H	1.22, 1.30	nd	1.48, 1.48	1.47, 1.47
	¹³ C _γ	24.90 ^b	nd	25.50	25.42
	C _{δδ} H	1.61, 1.61	nd	1.65, 1.65	1.66, 1.66
	¹³ C _δ	29.09 ^b	nd	29.47	29.11
	C _{εε} H	2.91, 2.91	nd	2.98, 2.98	2.94, 2.99
	¹³ C _ε	42.05 ^b	nd	42.44	42.74
	N _ε H ₃	nd	nd	nd	nd
	H _N	7.98	nd	8.03	8.11
	C _α H	4.87 ^b	5.07	4.44	4.47
	¹³ C _α	57.45 ^b	59.63	59.26	nd
Trp 248 (Trp10)	C _{ββ} H	3.02, 3.19	3.33, 3.46	3.29, 3.40	3.30, 3.43
	¹³ C _β	30.46 ^b	30.48	30.10	30.12
	C _{δ1} H	7.14	nd	7.05	7.02
	¹³ C _{δ1}	127.02 ^b	nd	127.28	127.12
	N _{ε1} H	9.98	nd	10.39	9.86
	C _{ε3} H	7.51	nd	7.46	7.48
	¹³ C _{ε3}	120.55 ^b	nd	121.00	120.93
	C _{ζ3} H	7.01	nd	6.92	6.95
	¹³ C _{ζ3}	122.05 ^b	nd	121.74	121.82
	C _{γ2} H	7.20	nd	7.06	7.11
	¹³ C _{γ2}	124.64 ^b	nd	124.26	124.55
	C _{ζ2} H	7.44	nd	7.43	7.44
	¹³ C _{ζ2}	114.75 ^b	nd	114.94	114.91

Table A3		(continuation)			
	Condition	H ₂ O/D ₂ O 9:1 (v/v)	30 % vol. [D ₃]-TFE ^c	30 mM [D ₃₈]-DPC	30 mM [D ₂₅]-SDS
	Peptide structure	β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Tyr 249 (Tyr11)	H _N	nd	nd	7.98	8.00
	C _α H	4.67 ^b	4.38	4.08	4.10
	¹³ C _α	57.67 ^b	56.97	58.30 ^b	60.70
	C _{ββ} H	2.67, 2.75	2.89, 2.98	2.87, 2.98	2.94, 3.01
	¹³ C _β	41.68 ^b	38.86	38.49	38.04
	C _{δδ} H	6.97, 6.97	6.98, 6.98	6.81, 6.81	6.94, 6.94
	¹³ C _δ	133.57 ^b	nd	133.41	133.33
	C _{εε} H	6.80, 6.80	nd	6.78, 6.78	6.81, 6.81
	¹³ C _ε	118.16 ^b	nd	118.60	118.62
	Tyr 250 (Tyr12)	H _N	nd	nd	7.67
C _α H		4.34 ^b	4.32	4.22	4.18
¹³ C _α		57.38 ^b	57.80	58.54	57.20
C _{ββ} H		1.90, 2.52	2.89, 2.89	2.85, 2.99	2.94, 2.98
¹³ C _β		39.05 ^b	38.86	38.58 ^b	37.93
C _{δδ} H		6.34, 6.34	nd	7.00, 7.00	6.94, 6.94
¹³ C _δ		132.83 ^b	nd	133.45	133.41
C _{εε} H		6.63, 6.63	nd	6.82, 6.82	6.81, 6.81
¹³ C _ε		117.86 ^b	nd	118.63	118.56
Phe 251 (Phe13)		H _N	7.81	nd	7.80
	C _α H	4.48	4.47	4.37	4.37
	¹³ C _α	57.19 ^b	57.75	59.27 ^b	59.68 ^b
	C _{ββ} H	2.84, 2.91	3.05, 3.12	3.10, 3.10	3.07, 3.13
	¹³ C _β	40.31 ^b	39.71	39.52 ^b	39.54 ^b
	C _{δδ} H	7.09, 7.09	7.22, 7.22	7.21, 7.21	7.25, 7.25
	¹³ C _δ	131.89 ^b	131.94	132.25	132.17
	C _{εε} H	7.19, 7.19	7.30, 7.30	7.27, 7.27	7.32, 7.32
	¹³ C _ε	131.31 ^b	131.46	131.48	131.56
	C _ζ H	7.06	7.23	7.20	7.27
Asn 252 (Asn14)	¹³ C _ζ	129.59 ^b	129.87	129.83	130.01
	H _N	8.22	nd	7.98	8.07
	C _α H	4.62	4.56	4.55	4.44
	¹³ C _α	52.56 ^b	54.54	53.88	54.68 ^b
	C _{ββ} H	2.80, 2.91	2.66, 2.66	2.48, 2.65	2.32, 2.53
	¹³ C _β	39.27 ^b	39.26	39.36	39.20 ^b
	N _{δδ} H ₂	nd	nd	6.41, 7.21	nd
	H _N	8.31	nd	8.02	7.87
	C _α H	4.28	4.21	4.19	4.17
	¹³ C _α	57.18 ^b	57.83	57.14	60.67
Glu253 (Glu15)	C _{ββ} H	2.00, 2.12	2.04, 2.14	1.90, 2.10	1.90, 2.08
	¹³ C _β	28.66 ^b	28.56	32.39	28.46
	C _{γγ} H	2.45, 2.45	2.42, 2.42	2.35, 2.35	2.34, 2.34
	¹³ C _γ	33.45 ^b	28.56	33.17	33.17
	H _N	8.12	nd	8.11	7.99
	C _α H	4.34	4.24	4.15	4.23
Glu254 (Glu16)	¹³ C _α	56.35 ^b	57.70	58.04	58.19
	C _{ββ} H	2.03, 2.18	2.03, 2.14	2.01, 2.09	1.99, 2.10
	¹³ C _β	28.58 ^b	28.56	33.03	28.45
	C _{γγ} H	2.46, 2.46	2.45, 2.45	2.49, 2.64	2.39, 2.39
	¹³ C _γ	33.53 ^b	33.29	33.12	33.14
	H _N	8.20	nd	8.20	8.08
Gly255 (GLy17)	C _{αα} H	3.83, 4.01	3.89, 3.89	3.87, 3.92	3.94, 4.01
	¹³ C _α	45.38 ^b	46.30	46.44 ^b	47.19 ^b

Table A3 (continuation)					
Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE ^c	[D ₃₈]-DPC	[D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Ala256 (Ala18)	H _N	7.98	nd	7.98	7.96
	C _α H	4.32	4.26	4.22	4.23
	¹³ C _α	52.20 ^b	53.77	54.05	54.32 ^b
	C _β H ₃	1.36	1.42	1.42	1.43
	¹³ C _β	19.44 ^b	18.96	19.41	19.76
Met257 (Met19)	H _N	nd	nd	8.15	8.20
	C _α H	4.37	4.43	4.34	4.35
	¹³ C _α	55.61 ^b	56.82	57.14	57.85 ^b
	C _{ββ} H	1.94, 2.00	2.14, 2.14	2.02, 2.09	2.11, 2.19
	¹³ C _β	33.27 ^b	32.98	32.53	32.84 ^b
	C _{γγ} H	2.35, 2.40	2.53, 2.64	2.51, 2.65	2.53, 2.63
	¹³ C _γ	32.11 ^b	32.50	33.49	33.26
	C _{εε} H	1.96, 1.96	nd	2.02, 2.02	2.02, 2.02
	¹³ C _ε	17.02 ^b	nd	17.75	17.91
Lys258 (Lys20)	H _N	8.25	nd	7.98	7.93
	C _α H	4.43	4.39	4.22	4.21
	¹³ C _α	55.99 ^b	57.33	58.16 ^b	58.18
	C _{ββ} H	1.71, 1.84	1.90, 1.96	1.85, 1.93	1.87, 1.95
	¹³ C _β	33.50 ^b	33.20	32.99	32.63
	C _{γγ} H	1.40, 1.40	1.47, 1.55	1.43, 1.53	1.43, 1.55
	¹³ C _γ	24.82 ^b	25.21	25.85	25.61
	C _{δδ} H	1.61, 1.61	1.71, 1.71	1.71, 1.71	1.73, 1.73
	¹³ C _δ	28.97 ^b	29.27	29.59	29.64
	C _{εε} H	2.93, 2.93	2.99, 2.99	2.97, 2.97	3.00, 3.00
	¹³ C _ε	42.14 ^b	42.226	42.57	42.62
	N _ε H ₃	7.49	nd	nd	nd
Thr259 (Thr21)	H _N	7.93	nd	7.94	7.91
	C _α H	4.40	4.40	4.22	4.14
	¹³ C _α	61.54 ^b	62.94	63.94 ^b	64.70
	C _β H	4.20	4.30	4.23	4.26
	¹³ C _β	70.39 ^b	70.29	69.64 ^b	69.78
	C _γ H ₃	1.19	1.29	1.23	1.22
	¹³ C _γ	21.59 ^b	21.61	22.27	22.29
Gly260 (Gly22)	H _N	8.39	nd	8.24	8.36
	C _{αα} H	3.56, 4.18	3.82, 4.15	3.93, 3.93	3.94, 3.94
	¹³ C _α	44.92 ^b	46.07	46.35 ^b	46.51 ^b
Trp261 (Trp23)	H _N	8.20	nd	8.24	8.14
	C _α H	5.22	4.68	4.22	4.21
	¹³ C _α	57.43 ^b	59.01	58.38	59.29
	C _{ββ} H	3.19, 3.37	3.28, 3.40	3.41, 3.41	3.43, 3.43
	¹³ C _β	30.64 ^b	30.37	29.85	30.08
	C _{δ1} H	7.29	nd	7.29	7.25
	¹³ C _{δ1}	127.13 ^b	nd	127.32	127.14
	N _{ε1} H	10.00	nd	10.44	9.88
	C _{ε3} H	7.58	nd	7.51	7.49
	¹³ C _{ε3}	120.75 ^b	nd	121.15	120.95
	C _{ζ3} H	7.03	nd	6.93	6.97
	¹³ C _{ζ3}	122.08 ^b	nd	121.68	121.84
	C _{η2} H	7.15	nd	7.08	7.11
	¹³ C _{η2}	124.71 ^b	nd	124.35	124.65
	C _{ζ2} H	7.31	nd	7.46	7.45
	¹³ C _{ζ2}	114.81 ^b	nd	114.98	114.86

Table A3 (continuation)					
Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE ^c	[D ₃₈]-DPC	[D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Val262 (Val24)	H _N	nd	nd	7.70	7.64
	C _α H	4.54 ^b	4.25	4.22	3.49
	¹³ C _α	61.21 ^b	62.49	61.19	59.70 ^b
	C _β H	1.90	2.07	2.09	2.06
	¹³ C _β	33.24 ^b	34.28	32.64	32.10
	C _γ H ₃	0.82	0.94	0.89	0.86
	¹³ C _γ	21.35 ^b	21.18	21.48	21.51
	C _{γ'} H ₃	0.82	0.96	0.98	0.97
	¹³ C _{γ'}	20.78 ^b	21.05	22.10	22.38
	N _ε H ₃	7.41	nd	nd	nd
Lys263 (Lys25)	H _N	8.21	nd	7.78	7.64
	C _α H	4.32	4.38	4.11	4.02
	¹³ C _α	55.54 ^b	57.31	57.90 ^b	58.85 ^b
	C _{ββ'} H	0.37, 1.28	nd	1.53, 1.55	1.61, 1.69
	¹³ C _β	32.87 ^b	nd	32.63	32.43 ^b
	C _{γγ'} H	0.73, 0.73	nd	1.07, 1.15	0.99, 1.11
	¹³ C _γ	24.68 ^b	nd	24.81 ^b	24.87
	C _{δδ'} H	1.20, 1.20	nd	1.53, 1.53	1.53, 1.53
	¹³ C _δ	29.06 ^b	nd	29.64	29.70
	C _{εε'} H	2.56, 2.67	nd	2.84, 2.84	2.85, 2.85
Tyr264 (Tyr26)	¹³ C _ε	41.69 ^b	nd	42.63	42.59
	N _ε H ₃	7.41	nd	nd	nd
	H _N	8.61	nd	nd	7.91
	C _α H	4.63 ^b	4.64	nd	4.45
	¹³ C _α	57.13 ^b	58.28	nd	60.04
	C _{ββ'} H	2.77, 2.85	2.94, 3.08	2.90, 3.17	2.89, 3.14
	¹³ C _β	39.64 ^b	39.69	38.58 ^b	38.78
	C _{δδ'} H	6.89, 6.89	6.85, 6.85	7.11, 7.11	7.11, 7.11
	¹³ C _δ	133.06 ^b	nd	133.31	133.32
	C _{εε'} H	6.64, 6.64	nd	6.79, 6.79	6.79, 6.79
Lys265 (Lys27)	¹³ C _ε	117.86 ^b	nd	118.55	118.55
	H _N	8.59	nd	8.23	7.99
	C _α H	3.69	3.79	3.78	3.74
	¹³ C _α	57.96 ^b	58.18	59.98 ^b	59.02
	C _{ββ'} H	1.69, 1.69	1.76, 1.76	1.55, 1.55	1.45, 1.45
	¹³ C _β	34.89 ^b	31.26	32.75	32.05
	C _{γγ'} H	0.81, 0.96	nd	1.20, 1.23	1.14, 1.14
	¹³ C _γ	24.85 ^b	nd	25.32	24.80
	C _{δδ'} H	1.51, 1.51	nd	1.55, 1.55	1.52, 1.52
	¹³ C _δ	29.11 ^b	nd	29.64	29.68
Asp266 (Asp28)	C _{εε'} H	2.87, 2.87	nd	2.90, 2.90	2.88, 2.88
	¹³ C _ε	41.95 ^b	nd	42.55	42.61
	N _ε H ₃	nd	nd	nd	7.38
	H _N	8.37	nd	8.23	7.76
	C _α H	4.50	4.64	4.59	4.48 ^b
	¹³ C _α	54.03 ^b	54.83	55.00 ^b	54.54
	C _{ββ'} H	2.82, 2.98	2.91, 2.98	2.91, 2.91	2.94, 2.94
	¹³ C _β	41.95 ^b	nd	38.99	38.36 ^b
	H _N	7.91	nd	7.85	7.76
	C _α H	4.48	4.42	4.09	4.14
Thr267 (Thr29)	¹³ C _α	61.85 ^b	62.89	60.64 ^b	64.62
	C _β H	4.13	4.18	4.16	4.12
	¹³ C _β	71.11 ^b	70.85	69.61 ^b	70.07
	C _γ H ₃	1.07	1.15	1.19	1.12
	¹³ C _γ	21.14 ^b	21.32	21.98	22.07

Table A3 (continuation)					
Condition		H ₂ O/D ₂ O	30 % vol.	30 mM	30 mM
		9:1 (v/v)	[D ₃]-TFE ^c	[D ₃₈]-DPC	[D ₂₅]-SDS
Peptide structure		β-hairpin	β-hairpin	α-helix	α-helix
Residue	Resonance				
Trp268 (Trp30)	H _N	8.35	nd	7.81	7.77
	C _α H	5.08	4.94	3.78	4.51
	¹³ C _α	57.26 ^b	57.97	58.99 ^b	58.79 ^b
	C _{ββ} H	3.06, 3.26	3.17, 3.32	3.23, 3.23	3.23, 3.23
	¹³ C _β	30.84 ^b	30.38	29.85	29.81
	C _{δ1} H	7.14	nd	6.96	7.05
	¹³ C _{δ1}	127.04 ^b	nd	127.43	127.08
	N _{ε1} H	9.95	nd	10.17	9.78
	C _{ε3} H	7.55	nd	7.46	7.49
	¹³ C _{ε3}	120.73 ^b	nd	120.92	120.84
	C _{ζ3} H	7.03	nd	6.98	6.97
	¹³ C _{ζ3}	122.07 ^b	nd	121.86	121.92
	C _{η2} H	7.19	nd	7.09	7.12
	¹³ C _{η2}	124.64 ^b	nd	124.4	124.67
	C _{ζ2} H	7.44	nd	7.43	7.45
	¹³ C _{ζ2}	114.72 ^b	nd	114.91	115.00
Tyr269 (Tyr31)	H _N	9.18	nd	nd	7.44
	C _α H	4.93 ^b	58.24	4.13 ^b	4.21
	¹³ C _α	57.37 ^b	2.94, 3.08	57.92 ^b	59.44
	C _{ββ} H	2.79, 2.87	39.52	2.79, 2.79	2.77, 2.77
	¹³ C _β	39.17 ^b	7.12, 7.12	38.77 ^b	38.52
	C _{δδ} H	7.02, 7.02	nd	6.86, 6.86	6.86, 6.86
	¹³ C _δ	133.52 ^b	nd	133.31	133.20
	C _{εε} H	6.81, 6.81	nd	6.82, 6.82	6.80, 6.80
	¹³ C _ε	118.24 ^b	nd	118.67	118.51
Tyr270 (Tyr32)	H _N	nd	nd	nd	7.44
	C _α H	4.34 ^b	4.63	4.34	4.36
	¹³ C _α	57.12 ^b	58.19	59.32 ^b	59.99 ^b
	C _{ββ} H	1.59, 2.52	nd, 2.83	2.96, 3.07	2.87, 3.00
	¹³ C _β	38.51 ^b	nd	nd	38.69 ^b
	C _{δδ} H	6.19, 6.19	6.97, 6.97	7.06, 7.06	7.06, 7.06
	¹³ C _δ	132.86 ^b	nd	133.47	133.45
	C _{εε} H	6.64, 6.64	nd	6.82, 6.82	6.83, 6.83
	¹³ C _ε	117.85 ^b	nd	118.59	118.58
Leu271 (Leu33)	H _N	8.07	nd	7.52	7.59
	C _α H	4.30	4.38	4.20	4.25
	¹³ C _α	54.43 ^b	55.09	55.06	55.77
	C _β H	1.46	1.59, 1.59	1.52, 1.66	1.56, 1.68
	¹³ C _β	42.41 ^b	42.88	43.00	42.94
	C _γ H	1.47	1.60	1.73	1.67
	¹³ C _γ	27.01 ^b	27.19	27.27	27.35
	C _δ H ₃	0.71	0.85	0.87	0.89
	¹³ C _δ	23.66 ^b	23.43	24.08	23.88
	C _{δ'} H ₃	0.73	0.87	0.90	0.93
	¹³ C _{δ'}	25.31 ^b	25.08	26.01	25.88
Asp272 (Asp34)	H _N	7.73	nd	7.67	7.87
	C _α H	4.50	4.60	4.50	4.64 ^b
	¹³ C _α	53.01 ^b	52.78	54.43 ^b	52.13 ^b
	C _{ββ} H	2.75, 2.84	2.77, 2.91	2.49, 2.64	2.66, 2.76
	¹³ C _β	41.73 ^b	39.34	39.43	38.63

Table A4. ^1H and ^{13}C chemical shifts (ppm from DSS) of SESYW11 in 30 mM $[\text{D}_{38}]$ -DPC in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v) at pH 5.5 and 25 °C. “nd” stands for not determined.

Condition			30 mM $[\text{D}_{38}]$ -DPC		
Residue	Resonance		Residue	Resonance	
Ser1	C_αH	4.18	Thr10	H_N	7.89
	$^{13}\text{C}_\alpha$	57.4		C_αH	4.36
	C_βH	3.93, 3.98		$^{13}\text{C}_\alpha$	61.8
	$^{13}\text{C}_\beta$	63.1		C_βH	4.19
Glu2	H_N	nd	Trp11	$^{13}\text{C}_\beta$	70.6
	C_αH	4.48		$\text{C}_\gamma\text{H}_3$	1.12
	$^{13}\text{C}_\alpha$	56.1		$^{13}\text{C}_\gamma$	21.6
	C_βH	1.90, 2.04		H_N	8.46
	$^{13}\text{C}_\beta$	30.8		C_αH	4.87
	C_γH	2.27, 2.27		$^{13}\text{C}_\alpha$	nd
	$^{13}\text{C}_\gamma$	35.8		$\text{C}_{\beta\beta'}\text{H}$	3.09, 3.18
Ser3	H_N	8.41		$^{13}\text{C}_\beta$	30.1
	C_αH	4.43		$\text{C}_{\delta 1}\text{H}$	7.21
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_{\delta 1}$	127.3
	C_βH	3.75, 3.75		$\text{N}_{\epsilon 1}\text{H}$	10.1
Tyr4	$^{13}\text{C}_\beta$	64.5		$\text{C}_{\epsilon 3}\text{H}$	7.47
	H_N	8.46		$^{13}\text{C}_{\epsilon 3}$	120.9
	C_αH	4.72		$\text{C}_{\zeta 3}\text{H}$	7.02
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_{\zeta 3}$	122.0
	C_βH	2.89, 3.02		$\text{C}_{\eta 2}\text{H}$	7.17
	$^{13}\text{C}_\beta$	39.6		$^{13}\text{C}_{\eta 2}$	124.6
	C_δH	6.99, 6.99		$\text{C}_{\zeta 2}\text{H}$	7.44
Ile5	$^{13}\text{C}_\delta$	133.7	Thr12	$^{13}\text{C}_{\zeta 2}$	114.9
	$\text{C}_\epsilon\text{H}$	6.75, 6.75		H_N	8.55
	$^{13}\text{C}_\epsilon$	118.3		C_αH	4.34
	H_N	8.02		$^{13}\text{C}_\alpha$	61.9
	C_αH	4.01	Val13	C_βH	3.75
	$^{13}\text{C}_\alpha$	60.5		$^{13}\text{C}_\beta$	70.3
	C_βH	1.32		$\text{C}_\gamma\text{H}_3$	1.07
	$^{13}\text{C}_\beta$	38.8		$^{13}\text{C}_\gamma$	21.5
	$\text{C}_{\gamma 2}\text{H}_3$	0.72		H_N	8.27
	$^{13}\text{C}_{\gamma 2}$	16.9		C_αH	4.38
	$\text{C}_{\gamma 1'}\text{H}$	0.79, 0.96		$^{13}\text{C}_\alpha$	61.8
	$^{13}\text{C}_{\gamma 1}$	27.4		C_βH	2.02
	$\text{C}_{\delta 1}\text{H}_3$	0.62		$^{13}\text{C}_\beta$	33.2
	$^{13}\text{C}_{\delta 1}$	12.9		$\text{C}_\gamma\text{H}_3$	0.91
Asn6	H_N	8.55		$^{13}\text{C}_\gamma$	20.6
	C_αH	nd		$\text{C}_{\gamma'}\text{H}_3$	0.91
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_{\gamma'}$	21.2
	$\text{C}_{\beta\beta'}\text{H}$	2.84, 2.84	Thr14	H_N	8.46
	$^{13}\text{C}_\beta$	39.5		C_αH	4.43
Ser7	$\text{N}_{\delta\delta'}\text{H}_2$	6.90, 7.42		$^{13}\text{C}_\alpha$	61.5
	H_N	8.66		C_βH	4.23
	C_αH	4.26		$^{13}\text{C}_\beta$	70.2
	$^{13}\text{C}_\alpha$	60.1		$\text{C}_\gamma\text{H}_3$	1.20
	C_βH	3.85, 3.92	Glu15	$^{13}\text{C}_\gamma$	21.5
Asp8	$^{13}\text{C}_\beta$	63.1		H_N	8.11
	H_N	8.01		C_αH	4.14
	C_αH	4.61		$^{13}\text{C}_\alpha$	58.0
	$^{13}\text{C}_\alpha$	nd		$\text{C}_{\beta\beta'}\text{H}$	1.89, 2.05
	$\text{C}_{\beta\beta'}\text{H}$	2.67, 2.85		$^{13}\text{C}_\beta$	30.8
Gly9	$^{13}\text{C}_\beta$	40.8		$\text{C}_{\gamma'}\text{H}$	2.24, 2.24
	H_N	8.15		$^{13}\text{C}_{\gamma'}$	35.9
	$\text{C}_{\alpha\alpha'}\text{H}$	3.67, 4.09			
	$^{13}\text{C}_\alpha$	45.5			

Table A5. ^1H and ^{13}C chemical shifts (ppm from DSS) of SESYV11 in 30 mM $[\text{D}_{38}]\text{-DPC}$ in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 (v/v) at pH 5.5 and 25 °C. “nd” stands for not determined.

Condition			30 mM $[\text{D}_{38}]\text{-DPC}$		
Residue	Resonance		Residue	Resonance	
Ser1	C_αH	4.19	Thr10	H_N	8.03
	$^{13}\text{C}_\alpha$	57.6		C_αH	4.33
	C_βH	3.92, 3.98		$^{13}\text{C}_\alpha$	62.4
	$^{13}\text{C}_\beta$	63.4		C_βH	4.20
Glu2	H_N	8.78	Val11	$^{13}\text{C}_\beta$	70.7
	C_αH	4.46		$\text{C}_\gamma\text{H}_3$	1.16
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\gamma$	21.9
	C_βH	1.88, 2.00		H_N	8.32
	$^{13}\text{C}_\beta$	31.1		C_αH	4.41
	C_γH	2.25, 2.25		$^{13}\text{C}_\alpha$	61.7
	$^{13}\text{C}_\gamma$	36.0		C_βH	1.98
Ser3	H_N	8.40		$^{13}\text{C}_\beta$	33.5
	C_αH	4.46		$\text{C}_\gamma\text{H}_3$	0.86
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\gamma$	21.4
	C_βH	3.71, 3.71		$\text{C}_\gamma\text{H}_3$	0.91
Tyr4	$^{13}\text{C}_\beta$	64.6	Thr12	$^{13}\text{C}_{\gamma'}$	20.3
	H_N	8.43		H_N	8.66
	C_αH	4.66		C_αH	4.39
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\alpha$	62.2
	C_βH	2.91, 3.02		C_βH	3.89
	$^{13}\text{C}_\beta$	39.7		$^{13}\text{C}_\beta$	70.5
	C_δH	7.03, 7.03		$\text{C}_\gamma\text{H}_3$	1.14
Ile5	$^{13}\text{C}_\delta$	nd	Val13	$^{13}\text{C}_\gamma$	22.0
	$\text{C}_\epsilon\text{H}$	6.78, 6.78		H_N	8.42
	$^{13}\text{C}_\epsilon$	nd		C_αH	4.39
	H_N	8.09		$^{13}\text{C}_\alpha$	62.2
	C_αH	4.23		C_βH	2.03
	$^{13}\text{C}_\alpha$	60.9		$^{13}\text{C}_\beta$	33.5
	C_βH	1.72		$\text{C}_\gamma\text{H}_3$	0.90
	$^{13}\text{C}_\beta$	39.0		$^{13}\text{C}_\gamma$	20.9
	$\text{C}_{\gamma 2}\text{H}_3$	0.86		$\text{C}_\gamma\text{H}_3$	0.90
	$^{13}\text{C}_{\gamma 2}$	17.5		$^{13}\text{C}_{\gamma'}$	20.9
	$\text{C}_{\gamma 1}\text{H}$	1.07, 1.42	Thr14	H_N	8.48
	$^{13}\text{C}_{\gamma 1}$	27.6		C_αH	4.42
Asn6	$\text{C}_{\delta 1}\text{H}_3$	0.79		$^{13}\text{C}_\alpha$	61.6
	$^{13}\text{C}_{\delta 1}$	13.1		C_βH	4.21
	H_N	8.64		$^{13}\text{C}_\beta$	70.3
	C_αH	4.61		$\text{C}_\gamma\text{H}_3$	1.18
	$^{13}\text{C}_\alpha$	nd	Glu15	$^{13}\text{C}_\gamma$	21.7
	$\text{C}_{\beta\beta'}\text{H}$	2.88, 2.88		H_N	8.11
	$^{13}\text{C}_\beta$	39.7		C_αH	4.13
Ser7	$\text{N}_{\delta\delta'}\text{H}_2$	7.03, 7.48		$^{13}\text{C}_\alpha$	58.2
	H_N	8.68		$\text{C}_{\beta\beta'}\text{H}$	1.89, 2.05
	C_αH	4.31		$^{13}\text{C}_\beta$	31.1
	$^{13}\text{C}_\alpha$	60.1		$\text{C}_{\gamma\gamma'}\text{H}$	2.23, 2.23
	C_βH	3.87, 3.95		$^{13}\text{C}_\gamma$	36.2
Asp8	$^{13}\text{C}_\beta$	63.4			
	H_N	8.08			
	C_αH	4.61			
	$^{13}\text{C}_\alpha$	nd			
Gly9	$\text{C}_{\beta\beta'}\text{H}$	2.70, 2.86			
	$^{13}\text{C}_\beta$	40.9			
	H_N	8.24			
	$\text{C}_{\alpha\alpha'}\text{H}$	3.76, 4.13			
	$^{13}\text{C}_\alpha$	45.7			

Table A6. ^1H and ^{13}C chemical shifts (ppm from DSS) of K3W5–LytA₂₃₉₋₂₅₂ in aqueous solution and in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 25 °C. “nd” stands for not determined. ^aIn pure D_2O . ^bBroad signal.

Condition		$\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1(v/v)		30 mM $[\text{D}_{38}]$ –DPC	
Peptide structure		β -hairpin		α -helix	
Residue	Resonance	^1H δ , ppm	^{13}C δ , ppm	^1H δ , ppm	^{13}C δ , ppm
Thr239 (Thr1)	C_αH	4.04	61.8	3.99	61.5
	C_βH	4.16	69.5	4.30	69.3
	$\text{C}_\gamma\text{H}_3$	1.35	21.8	1.34	21.8
Gly240 (Gly2)	H_N	8.57		9.12	
	$\text{C}_{\alpha\alpha'}\text{H}$	3.83, 4.13	45.1	3.99, 4.07	46.0
Lys241 (Lys3)	H_N	8.31		8.56	
	C_αH	5.19	55.5 ^a	4.22	57.8
	$\text{C}_{\beta\beta'}\text{H}$	1.42, 1.58	35.4	1.73, 1.73	33.1
	$\text{C}_{\gamma\gamma'}\text{H}$	1.04, 1.04	24.0	1.38, 1.45	25.2
	$\text{C}_{\delta\delta'}\text{H}$	0.84, 1.21	29.9	10.00	29.5
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.07, 2.17	41.8	2.90, 2.90	42.4
	N_ζH_3	6.96		7.71	
	H_N	9.17		8.27	
	C_αH	4.89 ^a	55.7 ^a	4.23	57.8
	$\text{C}_{\beta\beta'}\text{H}$	1.84, 1.91	36.8	1.78, 1.83	33.1
Lys242 (Lys4)	$\text{C}_{\gamma\gamma'}\text{H}$	1.64, 1.64	25.2	1.42, 1.45	25.2
	$\text{C}_{\delta\delta'}\text{H}$	1.53, 1.63	29.7	1.68, 1.68	29.3
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.84, 2.84	42.3	2.98, 2.98	42.4
	N_ζH_3	7.51		7.64	
	H_N	8.84		8.33	
	C_αH	4.65	56.5 ^a	4.49	59.6 ^a
	$\text{C}_{\beta\beta'}\text{H}$	2.18, 2.94	29.2	3.33, 3.33	29.6
	$\text{C}_{\delta 1}\text{H}$	6.83	127.3	7.36	127.6
	$\text{N}_{\epsilon 1}\text{H}$	9.75		10.61	
	$\text{C}_{\epsilon 3}\text{H}$	5.71 ^b	nd	7.47	121.0
Trp243 (Trp5)	$\text{C}_{\epsilon 3}\text{H}$	6.46	121.4	6.97	121.7
	$\text{C}_{\eta 2}\text{H}$	6.93	124.6	7.07	124.2
	$\text{C}_{\zeta 2}\text{H}$	7.24	114.1	7.48	114.8
	H_N	8.38		7.88	
	C_αH	4.01	60.8	3.89	62.7
	C_βH	1.65	40.4	1.83	38.9
	$\text{C}_{\gamma 2}\text{H}_3$	0.70	17.1	0.86	18.0
	$\text{C}_{\gamma\gamma' 1}\text{H}$	0.94, 1.34	27.1	1.06, 1.44	28.3
	$\text{C}_{\delta 1}\text{H}_3$	0.62	13.2	0.83	13.7
	H_N	8.59		8.00	
Ala245 (Ala7)	C_αH	3.75	54.2	4.07	54.2
	C_βH_3	1.28	17.6	1.42	19.0
	H_N	8.00		8.27	
Asp246 (Asp8)	C_αH	3.91	54.2	4.56 ^a	54.9 ^a
	$\text{C}_{\beta\beta'}\text{H}$	2.77, 2.86	38.4	2.97, 3.02	42.3
	H_N	6.85		7.86	
Lys247 (Lys9)	C_αH	3.96	55.7	4.13	57.2
	$\text{C}_{\beta\beta'}\text{H}$	1.52, 1.57	35.2	1.60, 1.65	32.5
	$\text{C}_{\gamma\gamma'}\text{H}$	1.08, 1.16	24.0	0.99, 1.15	24.5
	$\text{C}_{\delta\delta'}\text{H}$	1.53, 1.53	29.4	1.40, 1.45	28.8
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.85, 2.85	42.3	2.58, 2.76	42.3
	N_ζH_3	7.46		7.47	

Table A6 (continuation)					
Condition		H ₂ O/D ₂ O 9:1(v/v)		30 mM [D ₃₈]-DPC	
Peptide structure		β -hairpin		β -hairpin	
Residue	Resonance	¹ H δ , ppm	¹³ C δ , ppm	¹ H δ , ppm	¹³ C δ , ppm
Trp248 (Trp10)	H _N	7.99		7.98	
	C _{α} H	4.96	56.9	4.53	58.8 ^a
	C _{$\beta\beta$} H	2.67, 3.04	30.9	3.23, 3.23	29.9
	C _{δ1} H	7.42	128.1	7.02	127.5
	N _{ϵ1} H	10.20		10.52	
	C _{ϵ3} H	7.33	120.4	7.46	121.0
	C _{ζ3} H	7.04	122.7	6.98	121.7
	C _{η2} H	7.18	124.7	7.10	124.4
	C _{ζ2} H	7.35	115.3	7.47	114.8
	¹³ C _{ζ2}				
Tyr249 (Tyr11)	H _N	9.11		7.98	
	C _{α} H	4.54	57.5 ^a	4.16	59.5
	C _{$\beta\beta$} H	2.79, 2.79	40.8	2.70, 2.88	39.0
	C _{$\delta\delta$} H	7.01, 7.01	133.6	6.77, 6.77	133.4
	C _{$\epsilon\epsilon$} H	6.76, 6.76	118.5	6.77, 6.77	118.4
Tyr250 (Tyr12)	H _N	8.31		7.61	
	C _{α} H	4.83 ^a	57.8 ^a	4.21	59.5
	C _{$\beta\beta$} H	2.71, 2.92	39.8	2.71, 2.83	39.0
	C _{$\delta\delta$} H	7.09, 7.09	133.5	6.83, 6.83	133.1
	C _{$\epsilon\epsilon$} H	6.69, 6.69	118.1	6.74, 6.74	118.3
Phe251 (Phe13)	H _N	8.15		7.66	
	C _{α} H	4.76 ^a	56.9 ^a	4.47	58.1 ^a
	C _{$\beta\beta$} H	2.95, 3.12	40.6	2.90, 3.18	39.9
	C _{$\delta\delta$} H	7.20, 7.20	132.1	7.27, 7.27	132.4
	C _{$\epsilon\epsilon$} H	7.24, 7.24	129.2	7.26, 7.26	129.4
	C _{ζ} H	7.20	131.2	7.17	131.2
Asn252 (Asn14)	H _N	8.40		7.99	
	C _{α} H	4.57	54.0 ^a	4.54	53.2 ^a
	C _{$\beta\beta$} H	2.70, 2.83	40.3	2.56, 2.68	39.3
	N _{$\delta\delta$} H ₂	6.89, 7.59		6.63, 7.36	

Table A7. ^1H and ^{13}C chemical shifts (ppm from DSS) of W5K10–LytA₂₃₉₋₂₅₂ in aqueous solution and in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 25 °C. “nd” stands for not determined. ^aIn pure D_2O .

Condition		$\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1(v/v)		30 mM $[\text{D}_{38}]$ –DPC	
Peptide structure		β -hairpin		β -hairpin	
Residue	Resonance	^1H δ , ppm	^{13}C δ , ppm	^1H δ , ppm	^{13}C δ , ppm
Thr239 (Thr1)	C_αH	3.92	61.8	3.94	61.5
	C_βH	4.09	69.5	4.21	69.2
	$\text{C}_\gamma\text{H}_3$	1.30	21.7	1.28	21.6
Gly240 (Gly2)	H_N	8.54		9.06	
	$\text{C}_{\alpha\alpha'}\text{H}$	3.90, 4.01	45.1	3.98, 4.05	45.8
Trp241 (Trp3)	H_N	8.04		8.53	
	C_αH	5.07	56.4 ^a	4.58	57.8 ^a
	C_βH	2.75, 2.75	31.1	3.13, 3.13	29.7
	$\text{C}_{\delta 1}\text{H}$	7.08	127.7	7.33	127.5
	$\text{N}_{\epsilon 1}\text{H}$	10.09		10.62	
	$\text{C}_{\epsilon 3}\text{H}$	7.15	120.7	7.46	121.1
	$\text{C}_{\zeta 3}\text{H}$	6.65	122.2	7.00	121.5
	$\text{C}_{\eta 2}\text{H}$	7.16	124.7	7.09	124.1
	$\text{C}_{\zeta 2}\text{H}$	7.43	114.9	7.49	114.9
	H_N	8.89		8.37	
Lys242 (Lys4)	C_αH	4.53	55.5 ^a	4.06	58.5
	$\text{C}_{\beta\beta'}\text{H}$	1.74, 1.79	35.8	1.79, 1.79	32.4
	$\text{C}_{\gamma\gamma'}\text{H}$	1.30, 1.36	24.8	1.30, 1.30	25.0
	$\text{C}_{\delta\delta'}\text{H}$	1.55, 1.55	29.6	1.63, 1.63	29.3
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.76, 2.76	42.3	2.89, 2.93	42.2
	$\text{N}_{\zeta}\text{H}_3$	7.46		7.63	
	H_N	8.65		7.78	
	C_αH	5.04	57.3 ^a	4.40	59.3 ^a
	$\text{C}_{\beta\beta'}\text{H}$	3.14, 3.23	30.3	3.26, 3.31	29.3
	$\text{C}_{\delta 1}\text{H}$	7.29	127.5	7.37	127.8
Trp243 (Trp5)	$\text{N}_{\epsilon 1}\text{H}$	10.21		10.43	
	$\text{C}_{\epsilon 3}\text{H}$	7.36	120.7	7.45	121.2
	$\text{C}_{\zeta 3}\text{H}$	6.97	122.2	6.95	121.7
	$\text{C}_{\eta 2}\text{H}$	7.14	124.8	7.04	124.3
	$\text{C}_{\zeta 2}\text{H}$	7.42	114.8	7.40	114.8
	H_N	8.88		7.46	
	C_αH	4.20	61.5	3.77	63.4
	C_βH	1.86	39.8	1.77	38.2
	$\text{C}_{\gamma 2}\text{H}_3$	0.86	17.3	0.83	17.8
	$\text{C}_{\gamma\gamma'}\text{H}$	1.08, 1.43	27.4	1.05, 1.34	28.6
Ile244 (Ile6)	$\text{C}_{\delta 1}\text{H}_3$	0.69	13.1	0.83	13.7
	H_N	9.00		7.86	
	C_αH	4.00	53.8	4.12	54.3
	C_βH_3	1.45	17.8	1.44	18.7
	H_N	8.58		8.23	
Asp246 (Asp8)	C_αH	4.51	54.2 ^a	4.66	54.6 ^a
	$\text{C}_{\beta\beta'}\text{H}$	2.95, 3.03	38.6	2.94, 3.05	38.5
Lys247 (Lys9)	H_N	8.10		8.05	
	C_αH	4.53	55.7 ^a	4.15	57.8
	$\text{C}_{\beta\beta'}\text{H}$	1.76, 1.76	35.2	1.86, 1.86	32.3
	$\text{C}_{\gamma\gamma'}\text{H}$	1.32, 1.42	25.2	1.30, 1.43	24.9
	$\text{C}_{\delta\delta'}\text{H}$	1.68, 1.68	29.3	1.58, 1.58	29.2
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.97, 2.97	42.5	2.67, 2.85	42.2
	$\text{N}_{\zeta}\text{H}_3$	7.54		7.75	

Table A7 (continuation)					
Condition		H ₂ O/D ₂ O 9:1(v/v)		30 mM [D ₃₈]-DPC	
Peptide structure		β -hairpin		β -hairpin	
Residue	Resonance	¹ H δ , ppm	¹³ C δ , ppm	¹ H δ , ppm	¹³ C δ , ppm
Lys248	H _N	8.09		7.86	
(Lys10)	C _{α} H	4.18	56.4	4.11	58.2
	C _{$\beta\beta$} H	0.95, 0.95	33.4	1.78, 1.78	32.4
	C _{$\gamma\gamma$} H	0.40, 0.58	25.0	1.30, 1.30	25.1
	C _{$\delta\delta$} H	0.77, 0.96	29.4	1.63, 1.63	29.5
	C _{$\epsilon\epsilon$} H	2.02, 2.17	41.8	2.89, 2.96	42.3
	N ϵ H ₃	7.01		7.64	
Tyr249	H _N	8.75		7.70	
(Tyr11)	C _{α} H	4.59	56.9 ^a	4.31	59.4
	C _{$\beta\beta$} H	2.71, 2.84	41.0	2.89, 2.89	38.6
	C _{$\delta\delta$} H	7.04, 7.04	133.7	6.78, 6.78	133.4
	C _{$\epsilon\epsilon$} H	6.81, 6.81	118.4	6.75, 6.75	118.3
Tyr250	H _N	8.31		7.71	
(Tyr12)	C _{α} H	4.44	58.0	4.16	60.2
	C _{$\beta\beta$} H	2.27, 2.69	39.4	2.63, 2.70	39.5
	C _{$\delta\delta$} H	6.64, 6.64	133.3	6.90, 6.90	133.2
	C _{$\epsilon\epsilon$} H	6.58, 6.58	118.0	6.76, 6.76	118.5
Phe251	H _N	7.97		7.84	
(Phe13)	C _{α} H	4.60	57.4 ^a	4.62	57.6 ^a
	C _{$\beta\beta$} H	2.87, 3.05	40.4	2.96, 3.27	40.1
	C _{$\delta\delta$} H	7.18, 7.18	132.3	7.35, 7.35	132.5
	C _{$\epsilon\epsilon$} H	7.21, 7.21	131.4	7.27, 7.27	131.2
	C ζ H	7.15	129.8	7.14	129.4
Asn252	H _N	8.24		8.06	
(Asn14)	C _{α} H	4.49	54.3 ^a	4.64	53.4
	C _{$\beta\beta$} H	2.70, 2.77	40.0	2.77, 2.87	39.7
	N $\delta\delta$ H ₂	6.89, 7.56		6.89, 7.64	

Table A8. ^1H and ^{13}C chemical shifts (ppm from DSS) of S3S10–LytA_{239–252} in aqueous solution and in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 25 °C. “nd” stands for not determined. ^aIn pure D_2O .

Condition		$\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1(v/v)		30 mM $[\text{D}_{38}]$ –DPC	
Peptide structure		Random coil		α –helix	
Residue	Resonance	^1H δ , ppm	^{13}C δ , ppm	^1H δ , ppm	^{13}C δ , ppm
Thr239 (Thr1)	C_αH	3.96	61.8	3.94	61.7
	C_βH	4.21	69.2	4.18	69.1
	$\text{C}_\gamma\text{H}_3$	1.34	21.6	1.32	21.6
Gly240 (Gly2)	H_N	8.84		8.83	
	$\text{C}_{\alpha\alpha}\text{H}$	4.08, 4.08	45.4	4.06, 4.06	45.3
Ser241 (Ser3)	H_N	8.40		8.39	
	C_αH	4.48	58.4	4.45	58.4
	$\text{C}_{\beta\beta}\text{H}$	3.85, 3.85	64.3	3.83, 3.83	64.2
Lys242 (Lys4)	H_N	8.45		8.46	
	C_αH	4.33	56.5	4.30	56.6
	$\text{C}_{\beta\beta}\text{H}$	1.74, 1.82	33.4	1.72, 1.80	33.3
	$\text{C}_{\gamma\gamma}\text{H}$	1.39, 1.43	25.0	1.39, 1.44	25.0
	$\text{C}_{\delta\delta}\text{H}$	1.66, 1.66	29.4	1.66, 1.66	29.3
	$\text{C}_{\epsilon\epsilon}\text{H}$	2.98, 2.98	42.4	2.96, 2.96	42.2
	N_ζH_3	7.53		7.54	
	H_N	8.29		8.36	
	C_αH	4.30	56.8	4.29	56.8
	$\text{C}_{\beta\beta}\text{H}$	1.72, 1.78	33.3	1.71, 1.71	33.3
Lys243 (Lys5)	$\text{C}_{\gamma\gamma}\text{H}$	1.37, 1.37	25.0	1.37, 1.42	25.1
	$\text{C}_{\delta\delta}\text{H}$	1.65, 1.65	29.3	1.66, 1.66	29.4
	$\text{C}_{\epsilon\epsilon}\text{H}$	2.97, 2.97	42.4	2.96, 2.96	42.2
	N_ζH_3	7.52		7.55	
	H_N	8.24		8.20	
	C_αH	4.13	61.1	4.09	61.3
	C_βH	1.83	39.0	1.82	38.9
	$\text{C}_{\gamma_2}\text{H}_3$	0.91	17.7	0.89	17.8
	$\text{C}_{\gamma\gamma_1}\text{H}$	1.18, 1.46	27.5	1.16, 1.46	27.6
	$\text{C}_{\delta_1}\text{H}_3$	0.85	13.0	0.83	13.1
Ala245 (Ala7)	H_N	8.44		8.40	
	C_αH	4.30	52.7	4.28	52.8
	C_βH_3	1.37	19.5	1.36	19.4
Asp246 (Asp8)	H_N	8.37		8.37	
	C_αH	4.64	53.5	4.62	53.5 ^a
	$\text{C}_{\beta\beta}\text{H}$	2.80, 2.88	39.2	2.82, 2.82	39.2
Lys247 (Lys9)	H_N	8.39		8.31	
	C_αH	4.32	56.6	4.26	57.0
	$\text{C}_{\beta\beta}\text{H}$	1.71, 1.78	33.3	1.72, 1.72	33.2
	$\text{C}_{\gamma\gamma}\text{H}$	1.40, 1.40	25.0	1.37, 1.37	25.0
	$\text{C}_{\delta\delta}\text{H}$	1.67, 1.67	29.3	1.65, 1.65	29.2
	$\text{C}_{\epsilon\epsilon}\text{H}$	2.97, 2.97	42.5	2.93, 2.93	42.2
	N_ζH_3	7.52		7.60	
Ser248 (Ser10)	H_N	8.23		8.18	
	C_αH	4.37	58.7	4.36	58.9
	$\text{C}_{\beta\beta}\text{H}$	3.76, 3.76	64.0	3.76, 3.76	63.9
Tyr249 (Tyr11)	H_N	8.06		8.13	
	C_αH	4.46	58.2	4.38	58.7
	$\text{C}_{\beta\beta}\text{H}$	2.84, 2.84	39.2	2.83, 2.83	39.1
	$\text{C}_{\delta\delta}\text{H}$	6.98, 6.98	133.5	6.89, 6.89	133.4
	$\text{C}_{\epsilon\epsilon}\text{H}$	6.77, 6.77	118.4	6.75, 6.75	118.4

Table A8 (continuation)					
Condition		H ₂ O/D ₂ O 9:1(v/v)		30 mM [D ₃₈]-DPC	
Peptide structure		Random coil		α -helix	
Residue	Resonance	¹ H δ , ppm	¹³ C δ , ppm	¹ H δ , ppm	¹³ C δ , ppm
Tyr250 (Tyr12)	H _N	7.91		7.94	
	C _{α} H	4.47	57.9	4.30	59.1
	C _{$\beta\beta$} H	2.79, 2.92	39.3	2.71, 2.82	39.5
	C _{$\delta\delta$} H	7.03, 7.03	133.4	6.94, 6.94	133.3
	C _{$\epsilon\epsilon$} H	6.81, 6.81	118.4	6.76, 6.76	118.4
Phe251 (Phe13)	H _N	7.97		7.88	
	C _{α} H	4.58	57.7	4.57	57.5 ^a
	C _{$\beta\beta$} H	2.97, 3.11	40.1	2.95, 3.16	40.1
	C _{$\delta\delta$} H	7.24, 7.24	132.2	7.28, 7.28	132.4
	C _{$\epsilon\epsilon$} H	7.35, 7.35	131.7	7.30, 7.30	131.4
Asn252 (Asn14)	C _{γ} H	7.28	130.1	7.20	129.7
	H _N	8.11		8.14	
	C _{α} H	4.58	53.1	4.59	53.0 ^a
	C _{$\beta\beta$} H	2.71, 2.79	39.6	2.71, 2.78	39.7
	N _{$\delta\delta$} H ₂	6.87, 7.55		6.86, 7.58	

Table A9. ^1H and ^{13}C chemical shifts (ppm from DSS) of I3V10–LytA₂₃₉₋₂₅₂ in aqueous solution and in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 25 °C. “nd” stands for not determined. ^aIn pure D_2O .

Condition		$\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1(v/v)		30 mM $[\text{D}_{38}]$ –DPC	
Peptide structure		β -hairpin		α -helix	
Residue	Resonance	^1H δ , ppm	^{13}C δ , ppm	^1H δ , ppm	^{13}C δ , ppm
Thr239 (Thr1)	C_αH	3.95	61.8	3.97	61.6
	C_βH	4.19	69.2	4.23	69.2
	$\text{C}_\gamma\text{H}_3$	1.33	21.6	1.34	21.7
Gly240 (Gly2)	H_N	8.79		9.00	
	$\text{C}_{\alpha\alpha'}\text{H}$	3.97, 4.09	45.2	4.01, 4.20	45.9
Ile241 (Ile3)	H_N	8.23		8.59	
	C_αH	4.22	61.3	3.99	63.2
	C_βH	1.80	39.4	1.89	38.6
	$\text{C}_{\gamma 2}\text{H}_3$	0.88	17.7	0.92	18.2
	$\text{C}_{\gamma 1'}\text{H}$	1.16, 1.46	27.5	1.25, 1.56	28.9
	$\text{C}_{\delta 1}\text{H}_3$	0.86	13.2	0.90	13.7
	H_N	8.48		8.49	
Lys242 (Lys4)	C_αH	4.34	56.3	3.99	59.5
	$\text{C}_{\beta\beta'}\text{H}$	1.69, 1.77	33.6	1.84, 1.84	32.6
	$\text{C}_{\gamma'}\text{H}$	1.33, 1.42	25.0	1.37, 1.49	25.8
	$\text{C}_{\delta\delta'}\text{H}$	1.65, 1.65	29.4	1.71, 1.71	29.3
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.96, 2.96	42.3	2.97, 2.97	42.2
	N_ζH_3	7.53		7.70	
	H_N	8.43		7.90	
Lys243 (Lys5)	C_αH	4.40	56.3	4.18	58.5
	$\text{C}_{\beta\beta'}\text{H}$	1.70, 1.77	33.5	1.85, 1.85	32.6
	$\text{C}_{\gamma'}\text{H}$	1.33, 1.40	25.0	1.42, 1.51	25.4
	$\text{C}_{\delta\delta'}\text{H}$	1.65, 1.65	29.4	1.71, 1.71	29.3
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.96, 2.96	42.4	2.96, 2.96	42.2
	N_ζH_3	7.53		7.62	
	H_N	8.31		7.81	
Ile244 (Ile6)	C_αH	4.15	61.1	3.87	63.7
	C_βH	1.83	39.2	1.95	38.4
	$\text{C}_{\gamma 2}\text{H}_3$	0.91	17.7	0.91	18.0
	$\text{C}_{\gamma 1'}\text{H}$	1.16, 1.46	27.5	1.18, 1.64	28.7
	$\text{C}_{\delta 1}\text{H}_3$	0.83	13.0	0.85	13.4
	H_N	8.52		8.44	
Ala245 (Ala7)	C_αH	4.28	52.8	4.03	54.9
	C_βH_3	1.39	19.3	1.45	18.8
	H_N	8.40		8.39	
Asp246 (Asp8)	C_αH	4.61	53.7	4.58	55.4 ^a
	$\text{C}_{\beta\beta'}\text{H}$	2.78, 2.87	39.2	2.96, 2.96	38.6
	H_N	8.19		7.88	
Lys247 (Lys9)	C_αH	4.27	56.5	4.24	58.0
	$\text{C}_{\beta\beta'}\text{H}$	1.69, 1.69	33.5	1.91, 1.98	32.9
	$\text{C}_{\gamma'}\text{H}$	1.25, 1.31	25.0	1.51, 1.55	25.4
	$\text{C}_{\delta\delta'}\text{H}$	1.62, 1.62	29.3	1.71, 1.71	29.4
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.91, 2.91	42.4	2.94, 2.94	42.2
	N_ζH_3	7.52		7.78	
Val248 (Val10)	H_N	8.00		7.85	
	C_αH	4.01	62.4	3.88	64.3
	C_βH	1.85	33.2	2.14	32.5
	$\text{C}_{\gamma 1}\text{H}_3$	0.67	21.3	0.87	22.0
	$\text{C}_{\gamma 2}\text{H}_3$	0.82	20.9	0.97	22.0

Table A9 (continuation)					
Condition		H ₂ O/D ₂ O 9:1(v/v)		30 mM [D ₃₈]-DPC	
Peptide structure		β -hairpin		α -helix	
Residue	Resonance	¹ H δ , ppm	¹³ C δ , ppm	¹ H δ , ppm	¹³ C δ , ppm
Tyr249	H _N	8.26		8.09	
(Tyr11)	C _{α} H	4.48	57.8	4.23	60.0
	C _{$\beta\beta$} H	2.74, 2.82	39.7	2.82, 2.82	39.6
	C _{$\delta\delta$} H	7.04, 7.04	133.5	6.62, 6.62	133.2
	C _{$\epsilon\epsilon$} H	6.77, 6.77	118.4	6.66, 6.66	118.3
Tyr250	H _N	8.02		7.74	
(Tyr12)	C _{α} H	4.54	57.7	4.32	59.0
	C _{$\beta\beta$} H	2.77, 2.92	39.6	2.79, 2.99	38.6
	C _{$\delta\delta$} H	7.04, 7.04	133.5	7.09, 7.09	133.5
	C _{$\epsilon\epsilon$} H	6.77, 6.77	118.4	6.82, 6.82	118.4
Phe251	H _N	8.12		7.75	
(Phe13)	C _{α} H	4.60	57.5	4.49	58.3 ^a
	C _{$\beta\beta$} H	2.96, 3.12	40.2	3.13, 3.19	39.7
	C _{$\delta\delta$} H	7.23, 7.23	132.3	7.28, 7.28	132.4
	C _{$\epsilon\epsilon$} H	7.32, 7.32	131.7	7.29, 7.29	131.4
	C _{ζ} H	7.25	130.1	7.21	129.7
Asn252	H _N	8.17		7.99	
(Asn14)	C _{α} H	4.57	53.2	4.57	52.9 ^a
	C _{$\beta\beta$} H	2.73, 2.78	39.7	2.67, 2.78	39.8
	N _{$\delta\delta$} H ₂	6.87, 7.55		6.77, 7.56	

Table A10. ^1H and ^{13}C chemical shifts (ppm from DSS) of I5Y6T11T13–LytA₂₃₉₋₂₅₂ in aqueous solution and in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 25 °C. “nd” stands for not determined. ^aIn pure D_2O .

Condition		$\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1(v/v)		30 mM $[\text{D}_{38}]$ –DPC	
Peptide structure		β -hairpin		α -helix	
Residue	Resonance	^1H δ , ppm	^{13}C δ , ppm	^1H δ , ppm	^{13}C δ , ppm
Thr239 (Thr1)	C_αH	3.84	61.7	3.89	63.0
	C_βH	4.08	69.2	4.14	69.2
	$\text{C}_\gamma\text{H}_3$	1.25	21.6	1.23	21.7
Gly240 (Gly2)	H_N	8.58		8.94	
	$\text{C}_{\alpha\alpha'}\text{H}$	3.90, 4.06	45.2	4.01, 4.14	45.9
Trp241 (Trp3)	H_N	8.27		8.58	
	C_αH	4.84 ^a	57.4 ^a	4.58	59.0 ^a
	C_βH	3.16, 3.24	30.4	3.32, 3.32	
	$\text{C}_{\delta 1}\text{H}$	7.25	127.6	7.43	127.6
	$\text{N}_{\epsilon 1}\text{H}$	10.14		10.70	
	$\text{C}_{\epsilon 3}\text{H}$	7.55	121.0	7.46	120.8
	$\text{C}_{\zeta 3}\text{H}$	7.13	122.3	6.90	121.5
	$\text{C}_{\eta 2}\text{H}$	7.21	124.9	7.07	124.1
	$\text{C}_{\zeta 2}\text{H}$	7.46	115.0	7.49	115.0
	H_N	8.45		8.20	
	C_αH	4.32	56.4	3.83	59.3
Lys242 (Lys4)	$\text{C}_{\beta\beta'}\text{H}$	1.56, 1.62	34.3	1.56, 1.65	32.4
	$\text{C}_{\gamma\gamma'}\text{H}$	1.18, 1.18	24.7	1.01, 1.05	25.3
	$\text{C}_{\delta\delta'}\text{H}$	1.57, 1.57	29.3	1.58, 1.58	29.3
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.86, 2.86	42.4	2.89, 2.89	42.2
	N_ζH_3	7.48		7.69	
	H_N	8.05		7.59	
Ile243 (Ile5)	C_αH	4.11	60.9	3.89	61.7
	C_βH	1.46	38.7	1.76	37.9
	$\text{C}_{\gamma 2}\text{H}_3$	0.70	17.3	0.70	17.7
	$\text{C}_{\gamma 1'}\text{H}$	0.79, 1.11	27.5	1.15, 1.31	28.2
	$\text{C}_{\delta 1}\text{H}_3$	0.65	12.9	0.78	13.2
	H_N	8.35		7.48	
Tyr244 (Tyr6)	C_αH	4.51	58.0	4.28	60.3 ^a
	$\text{C}_{\beta\beta'}\text{H}$	2.86, 2.99	39.6	3.00, 3.06	39.2
	$\text{C}_{\delta\delta'}\text{H}$	7.09, 7.09	133.5	7.11, 7.11	133.0
	$\text{C}_{\epsilon\epsilon'}\text{H}$	6.80, 6.80	118.4	6.81, 6.81	118.6
Ala245 (Ala7)	H_N	8.35		8.05	
	C_αH	4.09	53.0	4.12	54.3
	C_βH_3	1.27	18.7	1.48	19.0
Asp246 (Asp8)	H_N	8.28		8.13	
	C_αH	4.42	53.8	4.56	54.6 ^a
	$\text{C}_{\beta\beta'}\text{H}$	2.80, 2.80	38.7	2.90, 2.90	38.4
Lys247 (Lys9)	H_N	8.03		7.82	
	C_αH	4.33	56.2	4.19	57.3
	$\text{C}_{\beta\beta'}\text{H}$	1.65, 1.71	33.7	1.77, 1.77	32.8
	$\text{C}_{\gamma\gamma'}\text{H}$	1.26, 1.26	24.6	1.32, 1.39	25.1
	$\text{C}_{\delta\delta'}\text{H}$	1.60, 1.60	29.3	1.58, 1.58	29.4
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.91, 2.91	42.3	2.84, 2.89	42.2
	N_ζH_3	7.51		7.65	

Table A10 (continuation)

Condition		H ₂ O/D ₂ O 9:1(v/v)		30 mM DPC-d38	
Peptide structure		β -hairpin		α -helix	
Residue	Resonance	¹ H δ , ppm	¹³ C δ , ppm	¹ H δ , ppm	¹³ C δ , ppm
Trp248	H _N	8.31		7.94	
(Trp10)	C _{α} H	4.78	nd	4.65	58.0 ^a
	C _{$\beta\beta'$} H	3.18, 3.26	29.9	3.23, 3.40	29.9
	C _{δ_1} H	7.15	127.5	7.22	127.0
	N _{ϵ_1} H	10.01		10.33	
	C _{ϵ_3} H	7.56	121.0	7.59	121.1
	C _{ζ_3} H	7.08	122.2	7.01	121.7
	C _{η_2} H	7.24	125.0	7.07	124.1
	C _{ζ_2} H	7.46	115.0	7.43	114.6
Thr249	H _N	8.27		7.69	
(Thr11)	C _{α} H	4.40	61.9	4.28	62.0
	C _{β} H	4.06	70.8	4.13	70.0
	C _{γ} H ₃	1.09	21.5	1.09	21.6
Tyr250	H _N	8.20		7.67	
(Tyr12)	C _{α} H	4.44	58.1	4.47	58.5 ^a
	C _{$\beta\beta'$} H	2.42, 2.83	38.89	2.79, 2.98	39.2
	C _{$\delta\delta'$} H	6.76, 6.76	133.4	6.99, 6.99	133.4
	C _{$\epsilon\epsilon'$} H	6.73, 6.73	118.4	6.78, 6.78	118.5
Thr251	H _N	7.92		7.88	
(Thr13)	C _{α} H	4.31	61.5	4.30	61.8
	C _{β} H	4.11	70.5	4.19	70.1
	C _{γ} H ₃	1.13	21.6	1.17	21.8
Asn252	H _N	8.14		8.15	
(Asn14)	C _{α} H	4.54	53.7 ^a	4.61	53.0 ^a
	C _{$\beta\beta'$} H	2.71, 2.80	39.8	2.73, 2.80	39.6
	N _{$\delta\delta'$} H ₂	6.88, 7.55		6.87, 7.58	

Table A11. ^1H and ^{13}C chemical shifts (ppm, from DSS) of crotonalidin in 30 mM $[\text{D}_{38}]\text{-DPC}$ at pH 3.0 and 35°C. ^aMeasured at 25 °C.

Condition			30 mM $[\text{D}_{38}]\text{-DPC}$		
Residue	Resonance		Residue	Resonance	
Lys1	H_N	nd	Lys5	N_ζH_3	nd
	C_αH	4.03	Phe6	H_N	7.98
	$^{13}\text{C}_\alpha$	55.4		C_αH	4.35
	$\text{C}_{\beta\beta'}\text{H}$	1.87, 1.87		$^{13}\text{C}_\alpha$	60.1
	$^{13}\text{C}_\beta$	31.6		$\text{C}_{\beta\beta'}\text{H}$	3.08, 3.26
	$\text{C}_{\gamma\gamma'}\text{H}$	1.45, 1.45		$^{13}\text{C}_\beta$	38.8
	$^{13}\text{C}_\gamma$	nd		$\text{C}_{\delta\delta'}\text{H}$	7.11, 7.11
	$\text{C}_{\delta\delta'}\text{H}$	1.74, 1.74		$^{13}\text{C}_\delta$	132.1
	$^{13}\text{C}_\delta$	nd		$\text{C}_{\epsilon\epsilon'}\text{H}$	7.23, 7.23
	$\text{C}_{\epsilon\epsilon'}\text{H}$	3.02, 3.02		$^{13}\text{C}_\epsilon$	131.3
Arg2	$^{13}\text{C}_\epsilon$	41.4		C_ζH	7.18
	N_ζH_3	nd		$^{13}\text{C}_\zeta$	129.7
	H_N	nd	Phe7	H_N	8.49
	C_αH	4.34		C_αH	4.14
	$^{13}\text{C}_\alpha$	56.5		$^{13}\text{C}_\alpha$	60.2
	$\text{C}_{\beta\beta'}\text{H}$	1.72, 1.72		$\text{C}_{\beta\beta'}\text{H}$	3.17, 3.17
	$^{13}\text{C}_\beta$	30.2		$^{13}\text{C}_\beta$	38.1
	$\text{C}_{\gamma\gamma'}\text{H}$	1.48, 1.55		$\text{C}_{\delta\delta'}\text{H}$	7.25, 7.25
	$^{13}\text{C}_\gamma$	26.5		$^{13}\text{C}_\delta$	132.2
	$\text{C}_{\delta\delta'}\text{H}$	3.12, 3.12		$\text{C}_{\epsilon\epsilon'}\text{H}$	7.32, 7.32
Phe3	$^{13}\text{C}_\delta$	42.9		$^{13}\text{C}_\epsilon$	131.4
	$\text{N}_\epsilon\text{H}$	7.39		C_ζH	7.22
	H_N	8.59		$^{13}\text{C}_\zeta$	129.7
	C_αH	4.50	Lys8	H_N	8.12
	$^{13}\text{C}_\alpha$	59.4		C_αH	3.85
	$\text{C}_{\beta\beta'}\text{H}$	3.19, 3.19		$^{13}\text{C}_\alpha$	59.3
	$^{13}\text{C}_\beta$	38.2		$\text{C}_{\beta\beta'}\text{H}$	1.94, 1.94
	$\text{C}_{\delta\delta'}\text{H}$	7.23, 7.23		$^{13}\text{C}_\beta$	nd
	$^{13}\text{C}_\delta$	132.1		$\text{C}_{\gamma\gamma'}\text{H}$	1.49, 1.49
	$\text{C}_{\epsilon\epsilon'}\text{H}$	7.32, 7.32		$^{13}\text{C}_\gamma$	24.0
Lys4	$^{13}\text{C}_\epsilon$	131.3		$\text{C}_{\delta\delta'}\text{H}$	1.67, 1.73
	C_ζH	nd		$^{13}\text{C}_\delta$	nd
	$^{13}\text{C}_\zeta$	nd		$\text{C}_{\epsilon\epsilon'}\text{H}$	3.01, 3.01
	H_N	8.23		$^{13}\text{C}_\epsilon$	nd
	C_αH	3.99	Lys9	N_ζH_3	nd
	$^{13}\text{C}_\alpha$	59.5		H_N	7.62
	$\text{C}_{\beta\beta'}\text{H}$	1.78, 1.91		C_αH	4.10
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\alpha$	58.3
	$\text{C}_{\gamma\gamma'}\text{H}$	1.44, 1.58		$\text{C}_{\beta\beta'}\text{H}$	1.95, 1.95
	$^{13}\text{C}_\gamma$	nd		$^{13}\text{C}_\beta$	nd
Lys5	$\text{C}_{\delta\delta'}\text{H}$	nd		$\text{C}_{\gamma\gamma'}\text{H}$	1.43, 1.57
	$^{13}\text{C}_\delta$	nd		$^{13}\text{C}_\gamma$	nd
	$\text{C}_{\epsilon\epsilon'}\text{H}$	3.02, 3.02		$\text{C}_{\delta\delta'}\text{H}$	1.72, 1.72
	$^{13}\text{C}_\epsilon$	41.2		$^{13}\text{C}_\delta$	nd
	N_ζH_3	nd		$\text{C}_{\epsilon\epsilon'}\text{H}$	2.96, 2.96
	H_N	8.16		$^{13}\text{C}_\epsilon$	nd
	C_αH	4.06	Val10	N_ζH_3	nd
	$^{13}\text{C}_\alpha$	58.6		H_N	7.92
	$\text{C}_{\beta\beta'}\text{H}$	1.86, 1.86		C_αH	3.76
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\alpha$	64.5
	$\text{C}_{\gamma\gamma'}\text{H}$	1.35, 1.43		C_βH	2.06
	$^{13}\text{C}_\gamma$	nd		$^{13}\text{C}_\beta$	31.2
	$\text{C}_{\delta\delta'}\text{H}$	1.69, 1.69		$\text{C}_\gamma\text{H}_3$	0.72
	$^{13}\text{C}_\delta$	nd		$^{13}\text{C}_\gamma$	21.4
	$\text{C}_{\epsilon\epsilon'}\text{H}$	2.96, 2.96		$\text{C}_{\gamma'}\text{H}_3$	0.81
	$^{13}\text{C}_\epsilon$	nd		$^{13}\text{C}_{\gamma'}$	20.9

Table A11 (continuation)

Condition		30 mM [D ₃₈]-DPC				
Residue	Resonance					
Lys11	H _N	8.19	Lys16	C _{εε} H	2.99, 2.99	
	C _α H	3.70		¹³ C _ε	nd	
	¹³ C _α	59.8		N _ζ H ₃	nd	
	C _{ββ} H	1.68, 1.75	Arg17	H _N	7.71	
	¹³ C _β	31.7		C _α H	4.21	
	C _{γγ} H	1.30, 1.34		¹³ C _α	57.2	
	¹³ C _γ	nd		C _{ββ} H	1.95, 1.95	
	C _{δδ} H	1.67, 1.67		¹³ C _β	29.6	
	¹³ C _δ	nd		C _{γγ} H	1.78, 1.78	
	C _{εε} H	2.83, 2.83		¹³ C _γ	27.1	
	¹³ C _ε	41.1		C _{δδ} H	3.19, 3.23	
	N _ζ H ₃	nd		¹³ C _δ	42.2	
Lys12	H _N	7.88	Leu18	N _ε H	7.67	
	C _α H	3.98		H _N	8.31	
	¹³ C _α	59.1		C _α H	4.16	
	C _{ββ} H	1.94, 1.94		¹³ C _α	56.7	
	¹³ C _β	nd		C _β H	1.58, 1.94	
	C _{γγ} H	1.46, 1.60		¹³ C _β	41.3	
	¹³ C _γ	nd		C _γ H	1.94	
	C _{δδ} H	nd		¹³ C _γ	26.4	
	¹³ C _δ	nd		C _δ H ₃	0.91	
	C _{εε} H	2.99, 2.99		¹³ C _δ	22.8	
	¹³ C _ε	nd		C _δ H ₃	0.91	
	N _ζ H ₃	nd		¹³ C _δ	25.4	
Ser13	H _N	7.95	Lys19	H _N	7.90	
	C _α H	4.27		C _α H	4.08	
	¹³ C _α	61.1		¹³ C _α	58.4	
	C _β H	3.93, 3.99		C _{ββ} H	1.87, 1.93	
Val14	¹³ C _β	62.5	¹³ C _β	nd		
	H _N	8.38	C _{γγ} H	1.52, 1.63		
	C _α H	3.66	¹³ C _γ	nd		
	¹³ C _α	65.9	C _{δδ} H	1.75, 1.75		
	C _β H	2.18	¹³ C _δ	nd		
	¹³ C _β	31.1	C _{εε} H	3.00, 3.00		
	C _γ H ₃	0.94	¹³ C _ε	nd		
	¹³ C _γ	21.0	N _ζ H ₃	nd		
	C _γ H ₃	1.04	Lys20	H _N	7.47	
	¹³ C _γ	22.7		C _α H	4.17	
	Lys15	H _N		8.27	¹³ C _α	57.3
		C _α H		3.84	C _{ββ} H	1.94, 1.94
¹³ C _α		60.2	¹³ C _β	nd		
C _{ββ} H		1.92, 1.92	C _{γγ} H	1.47, 1.58		
¹³ C _β		nd	¹³ C _γ	nd		
C _{γγ} H		1.38, 1.61	C _{δδ} H	1.72, 1.72		
¹³ C _γ		25.3	¹³ C _δ	nd		
C _{δδ} H		1.74, 1.74	C _{εε} H	3.00, 3.00		
¹³ C _δ		nd	¹³ C _ε	nd		
C _{εε} H		2.89, 2.92	Ile21	N _ζ H ₃	nd	
¹³ C _ε		41.1		H _N	7.57	
N _ζ H ₃		nd		C _α H	3.92	
Lys16	H _N	7.77		¹³ C _α	62.4	
	C _α H	4.03	C _β H	1.76		
	¹³ C _α	58.8	¹³ C _β	38.2		
	C _{ββ} H	1.96, 1.96	C _{γ2} H ₃	0.57		
	¹³ C _β	nd	¹³ C _{γ2}	16.7		
	C _{γγ} H	1.47, 1.60	C _{γγ1} H	1.17, 1.53		
	¹³ C _γ	nd	¹³ C _{γ1}	27.3		
	C _{δδ} H	1.73, 1.73	C _{δ1} H ₃	0.81		

Table A11 (continuation)

Condition		30 mM [D ₅ 8]-DPC	
Residue	Resonance		
Ile21	¹³ C _δ	13.1	
Phe22	H _N	7.89	
	C _α H	4.63	
	¹³ C _α	57.4 ^a	
	C _{ββ} H	2.95, 3.29	
	¹³ C _β	38.2	
	C _{δδ} H	7.38, 7.38	
	¹³ C _δ	132.4	
	C _{εε} H	7.24, 7.24	
	¹³ C _ε	131.3	
	C _ζ H	7.13	
	¹³ C _ζ	129.3	
Lys23	H _N	7.81	
	C _α H	4.29	
	¹³ C _α	56.3	
	C _{ββ} H	1.82, 1.92	
	¹³ C _β	32.4	
	C _{γγ} H	1.47, 1.47	
	¹³ C _γ	24.1	
	C _{δδ} H	1.71, 1.71	
	¹³ C _δ	nd	
	C _{εε} H	3.00, 3.03	
	¹³ C _ε	nd	
	N _ζ H ₃	nd	
Lys24	H _N	7.98	
	C _α H	4.61	
	¹³ C _α	53.6	
	C _{ββ} H	1.78, 1.89	
	¹³ C _β	nd	
	C _{γγ} H	1.48, 1.48	
	¹³ C _γ	nd	
	C _{δδ} H	1.75, 1.75	
	¹³ C _δ	nd	
	C _{εε} H	3.01, 3.01	
	¹³ C _ε	nd	
	N _ζ H ₃	nd	
Pro25	H _N	-	
	C _α H	4.42	
	¹³ C _α	63.0	
	C _{ββ} H	1.91, 2.27	
	¹³ C _β	31.2	
	C _{γγ} H	2.00, 2.09	
	¹³ C _γ	27.0	
	C _{δδ} H	3.66, 3.84	
	¹³ C _δ	49.9	
Met26	H _N	8.34	
	C _α H	4.50	
	¹³ C _α	55.1	
	C _{ββ} H	2.01, 2.01	
	¹³ C _β	33.1	
	C _{γγ} H	2.51, 2.51	
	¹³ C _γ	31.5	
	C _{εε} H	2.01	
	¹³ C _ε	16.6	
Val27	H _N	8.15	
	C _α H	4.16	
	¹³ C _α	60.7	
	C _β H	1.94	
Val27	¹³ C _β	32.3	
Val27	C _γ H ₃	0.90	
	¹³ C _γ	20.7	
	C _γ H ₃	0.93	
	¹³ C _{γ'}	20.2	
Ile28	H _N	8.21	
	C _α H	4.16	
	¹³ C _α	60.7	
	C _β H	1.94	
	¹³ C _β	38.1	
	C _{γ2} H ₃	0.95	
	¹³ C _{γ2}	17.3	
	C _{γγ1} H	1.20, 1.56	
	¹³ C _{γ1}	27.0	
	C _{δ1} H ₃	0.88	
	¹³ C _δ	12.6	
Gly29	H _N	8.50	
	C _{αα'} H	3.95, 3.95	
	¹³ C _α	45.1	
Val30	H _N	7.63	
	C _α H	4.27	
	¹³ C _α	61.0	
	C _β H	2.11	
	¹³ C _β	32.8	
	C _γ H ₃	0.92	
	¹³ C _γ	19.9	
	C _γ H ₃	0.94	
	¹³ C _{γ'}	20.8	
Thr31	H _N	8.19	
	C _α H	4.35	
	¹³ C _α	61.3	
	C _β H	4.09	
	¹³ C _β	69.1	
	C _γ H ₃	1.14	
	¹³ C _γ	21.0	
Ile32	H _N	8.22	
	C _α H	4.33	
	¹³ C _α	57.8	
	C _β H	1.87	
	¹³ C _β	37.8	
	C _{γ2} H ₃	0.81	
	¹³ C _{γ2}	17.0	
	C _{γγ1} H	1.19, 1.56	
	¹³ C _{γ1}	26.8	
	C _{δ1} H ₃	0.85	
	¹³ C _δ	12.1	
Pro33	H _N	-	
	C _α H	4.47	
	¹³ C _α	62.8	
	C _{ββ} H	1.87, 2.07	
	¹³ C _β	30.4	
	C _{γγ} H	1.74, 1.94	
	¹³ C _γ	26.5	
	C _{δδ} H	3.41, 3.88	
	¹³ C _δ	50.1	
Phe34	H _N	7.42	
	C _α H	4.61	
	¹³ C _α	56.2	

Table A11 (continuation)		
Condition		30 mM [D ₃₈]-DPC
Residue	Resonance	
Phe34	C _{ββ} H	3.02, 3.21
	¹³ C _β	39.5
	C _{δδ} H	7.24, 7.24
	¹³ C _δ	132.4
	C _{εε} H	7.26, 7.26
	¹³ C _ε	131.4
	C _ζ H	nd
	¹³ C _ζ	nd
CONH ₂	H _N	7.04, 7.62

Table A12. ^1H and ^{13}C chemical shifts (ppm, from DSS) of crotalicidin–Nt in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 35°C.

Condition			30 mM $[\text{D}_{38}]$ –DPC		
Residue	Resonance		Residue	Resonance	
Lys1	H_N	nd	Lys5	N_ζH_3	nd
	C_αH	4.01	Phe6	H_N	8.05
	$^{13}\text{C}_\alpha$	55.3		C_αH	4.41
	$\text{C}_{\beta\beta'}\text{H}$	1.89, 1.89		$^{13}\text{C}_\alpha$	58.9
	$^{13}\text{C}_\beta$	32.6		$\text{C}_{\beta\beta'}\text{H}$	3.03, 3.10
	$\text{C}_{\gamma\gamma'}\text{H}$	1.45, 1.45		$^{13}\text{C}_\beta$	39.0
	$^{13}\text{C}_\gamma$	23.8		$\text{C}_{\delta\delta'}\text{H}$	7.10, 7.10
	$\text{C}_{\delta\delta'}\text{H}$	1.65, 1.65		$^{13}\text{C}_\delta$	131.4
	$^{13}\text{C}_\delta$	28.6		$\text{C}_{\text{ee}'}\text{H}$	7.22, 7.22
	$\text{C}_{\text{ee}'}\text{H}$	2.97, 2.97		$^{13}\text{C}_\text{e}$	130.8
	$^{13}\text{C}_\text{e}$	41.3		C_ζH	7.19
Arg2	N_ζH_3	nd		$^{13}\text{C}_\zeta$	129.5
	H_N	nd	Phe7	H_N	8.24
	C_αH	4.31		C_αH	4.29
	$^{13}\text{C}_\alpha$	56.3		$^{13}\text{C}_\alpha$	58.9
	$\text{C}_{\beta\beta'}\text{H}$	1.67, 1.67		$\text{C}_{\beta\beta'}\text{H}$	3.04, 3.14
	$^{13}\text{C}_\beta$	31.0		$^{13}\text{C}_\beta$	38.7
	$\text{C}_{\gamma\gamma'}\text{H}$	1.47, 1.55		$\text{C}_{\delta\delta'}\text{H}$	7.24, 7.24
	$^{13}\text{C}_\gamma$	26.6		$^{13}\text{C}_\delta$	131.5
	$\text{C}_{\delta\delta'}\text{H}$	3.10, 3.10		$\text{C}_{\text{ee}'}\text{H}$	7.31, 7.31
	$^{13}\text{C}_\delta$	42.9		$^{13}\text{C}_\text{e}$	130.9
	$\text{N}_\epsilon\text{H}$			C_ζH	7.25
Phe3	H_N	8.51		$^{13}\text{C}_\zeta$	129.6
	C_αH	4.51	Lys8	H_N	8.07
	$^{13}\text{C}_\alpha$	58.5		C_αH	4.02
	$\text{C}_{\beta\beta'}\text{H}$	3.15, 3.15		$^{13}\text{C}_\alpha$	57.9
	$^{13}\text{C}_\beta$	38.6		$\text{C}_{\beta\beta'}\text{H}$	1.82, 1.82
	$\text{C}_{\delta\delta'}\text{H}$	7.24, 7.24		$^{13}\text{C}_\beta$	32.2
	$^{13}\text{C}_\delta$	131.6		$\text{C}_{\gamma\gamma'}\text{H}$	1.41, 1.55
	$\text{C}_{\text{ee}'}\text{H}$	7.27, 7.27		$^{13}\text{C}_\gamma$	24.7
	$^{13}\text{C}_\text{e}$	131.2		$\text{C}_{\delta\delta'}\text{H}$	nd
	C_ζH	nd		$^{13}\text{C}_\delta$	nd
	$^{13}\text{C}_\zeta$	nd		$\text{C}_{\text{ee}'}\text{H}$	2.98, 2.98
Lys4	H_N	8.19		$^{13}\text{C}_\text{e}$	nd
	C_αH	4.07	Lys9	N_ζH_3	nd
	$^{13}\text{C}_\alpha$	57.8		H_N	7.97
	$\text{C}_{\beta\beta'}\text{H}$	1.80, 1.80		C_αH	4.17
	$^{13}\text{C}_\beta$	32.5		$^{13}\text{C}_\alpha$	57.3
	$\text{C}_{\gamma\gamma'}\text{H}$	1.37, 1.46		$\text{C}_{\beta\beta'}\text{H}$	1.86, 1.86
	$^{13}\text{C}_\gamma$	24.7		$^{13}\text{C}_\beta$	32.1
	$\text{C}_{\delta\delta'}\text{H}$	1.70, 1.70		$\text{C}_{\gamma\gamma'}\text{H}$	1.41, 1.48
	$^{13}\text{C}_\delta$	28.6		$^{13}\text{C}_\gamma$	24.6
	$\text{C}_{\text{ee}'}\text{H}$	nd		$\text{C}_{\delta\delta'}\text{H}$	1.69, 1.69
	$^{13}\text{C}_\text{e}$	nd		$^{13}\text{C}_\delta$	nd
Lys5	N_ζH_3	nd		$\text{C}_{\text{ee}'}\text{H}$	2.96, 2.96
	H_N	8.15	Val10	$^{13}\text{C}_\text{e}$	nd
	C_αH	4.09		N_ζH_3	nd
	$^{13}\text{C}_\alpha$	57.7		H_N	8.07
	$\text{C}_{\beta\beta'}\text{H}$	1.75, 1.75		C_αH	3.78
	$^{13}\text{C}_\beta$	32.7		$^{13}\text{C}_\alpha$	63.8
	$\text{C}_{\gamma\gamma'}\text{H}$	1.28, 1.36		C_βH	2.02
	$^{13}\text{C}_\gamma$	24.5		$^{13}\text{C}_\beta$	31.7
	$\text{C}_{\delta\delta'}\text{H}$	1.67, 1.67		$\text{C}_\gamma\text{H}_3$	0.78
	$^{13}\text{C}_\delta$	nd		$^{13}\text{C}_\gamma$	21.2
	$\text{C}_{\text{ee}'}\text{H}$	2.92, 2.92		$\text{C}_{\gamma'}\text{H}_3$	0.85
	$^{13}\text{C}_\text{e}$	nd		$^{13}\text{C}_{\gamma'}$	21.0

Table A12 (continuation)		
Condition		30 mM [D ₃₈]-DPC
Residue	Resonance	
Lys11	H _N	8.25
	C _α H	4.02
	¹³ C _α	57.9
	C _{ββ} H	1.82, 1.82
	¹³ C _β	32.2
	C _{γγ} H	1.37, 1.37
	¹³ C _γ	23.6
	C _{δδ} H	nd
	¹³ C _δ	nd
	C _{εε} H	2.93, 2.93
	¹³ C _ε	41.3
	N _ε H ₃	nd
Lys12	H _N	8.05
	C _α H	4.20
	¹³ C _α	56.9
	C _{ββ} H	1.85, 1.85
	¹³ C _β	32.5
	C _{γγ} H	1.47, 1.47
	¹³ C _γ	24.3
	C _{δδ} H	1.68, 1.68
	¹³ C _δ	nd
	C _{εε} H	nd
	¹³ C _ε	nd
	N _ε H ₃	nd
Ser13	H _N	8.12
	C _α H	4.40
	¹³ C _α	59.9
	C _β H	3.87, 3.87
	¹³ C _β	63.8
Val14	H _N	7.93
	C _α H	4.11
	¹³ C _α	61.8
	C _β H	2.13
	¹³ C _β	32.0
	C _γ H ₃	0.95
	¹³ C _γ	20.0
	C _{γ'} H ₃	0.95
	¹³ C _{γ'}	20.9
CONH ₂	H _N	nd, nd

Table A13. ^1H and ^{13}C chemical shifts (ppm, from DSS) of crotaledidin–Ct in 30 mM $[\text{D}_{38}]$ –DPC at pH 3.0 and 35°C.

Condition		30 mM [D ₃₈]-DPC			
Residue	Resonance				
Lys15	H _N	nd	Lys19	N _ζ H ₃	nd
	C _α H	4.02	Lys20	H _N	8.13
	¹³ C _α	55.5		C _α H	4.24
	C _{ββ'} H	1.89, 1.89		¹³ C _α	56.4
	¹³ C _β	33.2		C _{ββ'} H	1.69, 1.69
	C _{γγ'} H	1.44, 1.44		¹³ C _β	33.1
	¹³ C _γ	24.0		C _{γγ'} H	1.27, 1.35
	C _{δδ'} H	1.70, 1.70		¹³ C _γ	24.9
	¹³ C _δ	29.1		C _{δδ'} H	1.65, 1.65
	C _{εε'} H	2.98, 2.98		¹³ C _δ	29.0
	¹³ C _ε	41.8		C _{εε'} H	2.92, 2.92
	N _ζ H ₃	7.62		¹³ C _ε	41.9
Lys16	H _N	8.70	Ile21	N _ζ H ₃	7.57
	C _α H	4.34		H _N	8.01
	¹³ C _α	56.4		C _α H	4.07
	C _{ββ'} H	1.75, 1.75		¹³ C _α	61.2
	¹³ C _β	33.3		C _β H	1.73
	C _{γγ'} H	1.40, 1.44		¹³ C _β	38.9
	¹³ C _γ	24.8		C _{γ2} H ₃	1.10, 1.36
	C _{δδ'} H	nd		¹³ C _{γ2}	27.3
	¹³ C _δ	nd		C _{γγ1'} H	0.79
	C _{εε'} H	2.97, 2.97		¹³ C _δ	13.1
	¹³ C _ε	nd	Phe22	H _N	8.20
	N _ζ H ₃	nd		C _α H	4.62
Arg17	H _N	7.28		¹³ C _α	57.4
	C _α H	4.34		C _{ββ'} H	2.95, 3.12
	¹³ C _α	56.0		¹³ C _β	39.8
	C _{ββ'} H	1.76, 1.80		C _{δδ'} H	7.27, 7.27
	¹³ C _β	31.1		¹³ C _δ	132.0
	C _{γγ'} H	1.61, 1.65		C _{εε'} H	7.24, 7.24
	¹³ C _γ	27.3		¹³ C _ε	131.0
	C _{δδ'} H	3.18, 3.18		C _ζ H	7.18
	¹³ C _δ	43.4		¹³ C _ζ	129.3
	N _ε H	7.28	Lys23	H _N	8.15
	H _N	8.37		C _α H	4.26
C _α H	4.32	¹³ C _α		56.1	
¹³ C _α	55.3	C _{ββ'} H		1.77, 1.77	
C _β H	1.57, 1.64	¹³ C _β		33.3	
¹³ C _β	42.5	C _{γγ'} H		1.36, 1.36	
C _γ H	1.64	¹³ C _γ		24.7	
¹³ C _γ	27.2	C _{δδ'} H		1.67, 1.67	
C _δ H ₃	0.87	¹³ C _δ		29.1	
¹³ C _δ	23.9	C _{εε'} H		2.96, 2.96	
C _{δ'} H ₃	0.92	¹³ C _ε		nd	
¹³ C _{δ'}	25.1	Lys24	N _ζ H ₃	7.54	
Lys19	H _N		8.23	H _N	8.21
	C _α H		4.25	C _α H	4.47
	¹³ C _α		56.5	¹³ C _α	54.5
	C _{ββ'} H		1.72, 1.77	C _{ββ'} H	1.72, 1.82
	¹³ C _β		33.5	¹³ C _β	32.7
	C _{γγ'} H		1.36, 1.43	C _{γγ'} H	1.46, 1.46
	¹³ C _γ		25.0	¹³ C _γ	24.7
	C _{δδ'} H		nd	C _{δδ'} H	1.70, 1.70
	¹³ C _δ		nd	¹³ C _δ	29.3
	C _{εε'} H		2.96, 2.96	C _{εε'} H	2.97, 2.97
	¹³ C _ε		nd	¹³ C _ε	41.9

Table A13 (continuation)

Condition		30 mM [D ₃₈]-DPC				
Residue	Resonance					
Lys24	N ϵ H ₃	7.62	Thr31	¹³ C β	69.6	
Pro25	H _N	-		C γ H ₃	1.12	
	C α H	4.42		¹³ C γ	21.6	
	¹³ C α	63.4	Ile32	H _N	8.20	
	C $\beta\beta'$ H	1.87, 2.27		C α H	4.35	
	¹³ C β	32.1		¹³ C α	58.3	
	C $\gamma\gamma'$ H	1.99, 2.02		C β H	1.84	
	¹³ C γ	27.5		¹³ C β	38.5	
	C $\delta\delta'$ H	3.63, 3.85		C γ_2 H ₃	0.83	
	¹³ C δ	50.6		¹³ C γ_2	17.6	
	Met26	H _N		8.37	C $\gamma\gamma_1'$ H	1.16, 1.52
C α H		4.47		¹³ C γ_1	27.3	
¹³ C α		55.6		C δ_1 H ₃	0.82	
C $\beta\beta'$ H		1.97, 1.97	¹³ C δ	12.7		
¹³ C β		33.6	Pro33	H _N	-	
C $\gamma\gamma'$ H		2.49, 2.53		C α H	4.42	
¹³ C γ		32.1		¹³ C α	63.4	
C $\epsilon\epsilon'$ H		2.03		C $\beta\beta'$ H	1.84, 2.09	
¹³ C ϵ		17.1		¹³ C β	31.3	
Val27		H _N		8.13	C $\gamma\gamma'$ H	1.79, 1.92
	C α H	4.15		¹³ C γ	27.2	
	¹³ C α	62.1		C $\delta\delta'$ H	3.46, 3.85	
	C β H	2.05		¹³ C δ	50.8	
	¹³ C β	33.1		Phe34	H _N	7.58
	C γ H ₃	0.88	C α H		4.58	
	¹³ C γ	20.5	¹³ C α		56.9	
	C γ' H ₃	0.89	C $\beta\beta'$ H		3.01, 3.16	
	¹³ C γ'	21.3	¹³ C β		40.1	
	Ile28	H _N	8.24		C $\delta\delta'$ H	7.23, 7.23
C α H		4.13	¹³ C δ		132.0	
¹³ C α		61.2	C $\epsilon\epsilon'$ H		7.28, 7.28	
C β H		1.90	¹³ C ϵ		131.1	
¹³ C β		38.7	C ζ H		7.20	
C γ_2 H ₃		0.92	¹³ C ζ	129.5		
¹³ C γ_2		17.9	CONH ₂	H _N	7.00, 7.56	
C $\gamma\gamma_1'$ H		1.17, 1.52				
¹³ C γ_1		27.5				
C δ_1 H ₃		0.86				
¹³ C δ	13.2					
Gly29	H _N	8.45				
	C $\alpha\alpha'$ H	3.92, 3.92				
	¹³ C α	45.5				
	Val30	H _N		7.74		
		C α H		4.24		
		¹³ C α	61.6			
		C β H	2.08			
		¹³ C β	33.4			
		C γ H ₃	0.89			
		¹³ C γ	20.7			
C γ' H ₃		0.90				
¹³ C γ'		21.4				
Thr31		H _N	8.22			
	C α H	4.33				
	¹³ C α	61.8				
	C β H	4.06				

Table A14. Backbone ^1H and ^{13}C chemical shifts (ppm) of rCtD–Fra e 9 in $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v at pH 6.0 and 25°C.

Residue	Resonance		Residue	Resonance	
Asp354	H_N	nd	Pro368	$^{15}\text{N}_\text{H}$	nd
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\alpha$	63.0
	$^{13}\text{C}_\alpha$	54.3		$^{13}\text{C}_\beta$	32.0
	$^{13}\text{C}_\beta$	41.4	Thr369	H_N	8.25
Glu355	H_N	nd		$^{15}\text{N}_\text{H}$	114.3
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\alpha$	62.0
	$^{13}\text{C}_\alpha$	57.5	Gln370	$^{13}\text{C}_\beta$	70.0
	$^{13}\text{C}_\beta$	33.9		H_N	8.48
Pro356	H_N	-		$^{15}\text{N}_\text{H}$	122.9
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\alpha$	56.0
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\beta$	29.6
	$^{13}\text{C}_\beta$	nd		$\text{H}_{\text{N}\epsilon}$	nd
Val357	H_N	nd		$^{15}\text{N}_\epsilon$	nd
	$^{15}\text{N}_\text{H}$	nd		H_N	8.46
	$^{13}\text{C}_\alpha$	nd		$^{15}\text{N}_\text{H}$	110.0
	$^{13}\text{C}_\beta$	nd	Asn372	$^{13}\text{C}_\alpha$	45.3
Pro358	H_N	-		H_N	8.33
	$^{15}\text{N}_\text{H}$	nd		$^{15}\text{N}_\text{H}$	118.7
	$^{13}\text{C}_\alpha$	63.2		$^{13}\text{C}_\alpha$	53.2
	$^{13}\text{C}_\beta$	32.1		$^{13}\text{C}_\beta$	39.0
Thr359	H_N	7.97		$\text{H}_{\text{N}\delta}$	6.94, 7.60
	$^{15}\text{N}_\text{H}$	113.0	Lys373	$^{15}\text{N}_\delta$	112.8
	$^{13}\text{C}_\alpha$	57.5		H_N	8.16
	$^{13}\text{C}_\beta$	79.4		$^{15}\text{N}_\text{H}$	121.0
Pro360	H_N	-		$^{13}\text{C}_\alpha$	56.8
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\beta$	33.6
	$^{13}\text{C}_\alpha$	nd	Lys374	H_N	6.79
	$^{13}\text{C}_\beta$	nd		$^{15}\text{N}_\text{H}$	117.9
Ser361	H_N	nd		$^{13}\text{C}_\alpha$	55.2
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\beta$	37.0
	$^{13}\text{C}_\alpha$	nd		H_N	9.13
	$^{13}\text{C}_\beta$	nd	Trp375	$^{15}\text{N}_\text{H}$	119.5
Ser362	H_N	nd		$^{13}\text{C}_\alpha$	58.1
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\beta$	33.6
	$^{13}\text{C}_\alpha$	nd		$\text{H}_{\text{N}\epsilon 1}$	nd
	$^{13}\text{C}_\beta$	nd		$^{15}\text{N}_{\epsilon 1}$	nd
Pro363	H_N	-	Cys376	H_N	8.35
	$^{15}\text{N}_\text{H}$	nd		$^{15}\text{N}_\text{H}$	121.2
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\alpha$	54.4
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\beta$	41.2
Val364	H_N	nd		H_N	8.50
	$^{15}\text{N}_\text{H}$	nd		$^{15}\text{N}_\text{H}$	122.1
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\alpha$	57.0
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\beta$	30.0
Pro365	H_N	-	Pro378	H_N	-
	$^{15}\text{N}_\text{H}$	nd		$^{15}\text{N}_\text{H}$	nd
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\alpha$	61.7
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\beta$	32.5
Lys366	H_N	nd		H_N	8.71
	$^{15}\text{N}_\text{H}$	nd		$^{15}\text{N}_\text{H}$	124.0
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\alpha$	57.1
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\beta$	33.7
Pro367	H_N	-	Ala380	H_N	9.01
	$^{15}\text{N}_\text{H}$	nd		$^{15}\text{N}_\text{H}$	127.7
	$^{13}\text{C}_\alpha$	nd		$^{13}\text{C}_\alpha$	55.5
	$^{13}\text{C}_\beta$	nd		$^{13}\text{C}_\beta$	18.3
Pro368	H_N	-		H_N	9.19

Table A14 (continuation)

Residue	Resonance		Residue	Resonance	
Glu381	¹⁵ N _H	112.6	Val394	¹³ C _α	67.3
	¹³ C _α	57.4		¹³ C _β	32.1
	¹³ C _β	28.8	Cys395	H _N	8.65
Ala382	H _N	7.20		¹⁵ N _H	111.6
	¹⁵ N _H	123.2		¹³ C _α	56.3
	¹³ C _α	52.5		¹³ C _β	38.8
	¹³ C _β	19.4	Ser396	H _N	7.50
Thr383	H _N	8.83		¹⁵ N _H	113.5
	¹⁵ N _H	113.6		¹³ C _α	58.7
	¹³ C _α	60.1		¹³ C _β	64.1
	¹³ C _β	71.4	Gln397	H _N	7.70
Asp384	H _N	9.02		¹⁵ N _H	121.5
	¹⁵ N _H	121.6		¹³ C _α	55.1
	¹³ C _α	58.0		¹³ C _β	28.9
	¹³ C _β	39.4		H _{Nε}	nd
Ala385	H _N	8.33		¹⁵ N _ε	nd
	¹⁵ N _H	121.1	Gly398	H _N	8.67
	¹³ C _α	58.8		¹⁵ N _H	110.0
	¹³ C _β	29.1		¹³ C _α	46.6
Gln386	H _N	7.74	Gly399	H _N	8.66
	¹⁵ N _H	118.8		¹⁵ N _H	111.4
	¹³ C _α	58.8		¹³ C _α	45.6
	¹³ C _β	29.1	Met400	H _N	7.53
	H _{Nε}	nd		¹⁵ N _H	118.6
	¹⁵ N _ε	nd		¹³ C _α	53.5
Leu387	H _N	8.49		¹³ C _β	32.7
	¹⁵ N _H	118.9	Asp401	H _N	8.76
	¹³ C _α	57.7		¹⁵ N _H	122.5
	¹³ C _β	40.4		¹³ C _α	53.1
Gln388	H _N	8.87		¹³ C _β	41.1
	¹⁵ N _H	120.4	Cys402	H _N	9.20
	¹³ C _α	58.3		¹⁵ N _H	123.2
	¹³ C _β	28.2		¹³ C _α	54.9
	H _{Nε}	nd		¹³ C _β	42.7
	¹⁵ N _ε	nd	Gly403	H _N	9.03
Ser389	H _N	7.71		¹⁵ N _H	111.6
	¹⁵ N _H	114.9		¹³ C _α	48.5
	¹³ C _α	62.1	Pro404	H _N	-
	¹³ C _β	62.7		¹⁵ N _H	nd
Asn390	H _N	7.73		¹³ C _α	65.6
	¹⁵ N _H	121.5		¹³ C _β	33.0
	¹³ C _α	56.0	Val405	H _N	6.99
	¹³ C _β	36.7		¹⁵ N _H	102.5
	H _{Nδ}	6.47, 7.55		¹³ C _α	59.5
	¹⁵ N _δ	107.8		¹³ C _β	30.2
Ile391	H _N	nd	Gln406	H _N	7.32
	¹⁵ N _H	nd		¹⁵ N _H	119.7
	¹³ C _α	66.1		¹³ C _α	54.0
	¹³ C _β	38.5		¹³ C _β	17.8
Asp392	H _N	8.56		H _{Nε}	nd
	¹⁵ N _H	119.0		¹⁵ N _ε	nd
	¹³ C _α	57.1	Ala407	H _N	8.14
	¹³ C _β	40.0		¹⁵ N _H	121.0
Tyr393	H _N	8.20		¹³ C _α	54.1
	¹⁵ N _H	119.2		¹³ C _β	17.8
	¹³ C _α	61.4	Asn408	H _N	7.84
	¹³ C _β	38.4		¹⁵ N _H	114.0
Val394	H _N	8.29		¹³ C _α	54.5
	¹⁵ N _H	119.3		¹³ C _β	37.7

Table A14 (continuation)

Residue	Resonance		Residue	Resonance	
Asn408	$^{15}\text{N}_\delta$	113.6	Ser422	H_N	8.26
	$\text{H}_{\text{N}\delta}$	6.87, 7.49		$^{15}\text{N}_\text{H}$	112.1
Gly409	H_N	8.69		$^{13}\text{C}_\alpha$	62.3
	$^{15}\text{N}_\text{H}$	109.0		$^{13}\text{C}_\beta$	63.7
	$^{13}\text{C}_\alpha$	44.7	Tyr423	H_N	7.27
Ala410	H_N	9.12		$^{15}\text{N}_\text{H}$	116.7
	$^{15}\text{N}_\text{H}$	124.0		$^{13}\text{C}_\alpha$	63.1
	$^{13}\text{C}_\alpha$	54.9		$^{13}\text{C}_\beta$	37.9
	$^{13}\text{C}_\beta$	19.2	Val424	H_N	7.92
Cys411	H_N	8.53		$^{15}\text{N}_\text{H}$	114.2
	$^{15}\text{N}_\text{H}$	112.9		$^{13}\text{C}_\alpha$	65.7
	$^{13}\text{C}_\alpha$	55.2		$^{13}\text{C}_\beta$	31.2
	$^{13}\text{C}_\beta$	47.1	Met425	H_N	9.18
Phe412	H_N	7.53		$^{15}\text{N}_\text{H}$	121.3
	$^{15}\text{N}_\text{H}$	122.0		$^{13}\text{C}_\alpha$	59.5
	$^{13}\text{C}_\alpha$	61.5		$^{13}\text{C}_\beta$	33.7
	$^{13}\text{C}_\beta$	39.2	Asn426	H_N	8.27
Asn413	H_N	8.42		$^{15}\text{N}_\text{H}$	116.6
	$^{15}\text{N}_\text{H}$	115.0		$^{13}\text{C}_\alpha$	59.9
	$^{13}\text{C}_\alpha$	50.3		$^{13}\text{C}_\beta$	39.9
	$^{13}\text{C}_\beta$	39.7		$\text{H}_{\text{N}\delta}$	nd
	$\text{H}_{\text{N}\delta}$	7.22, 7.71		$^{15}\text{N}_\delta$	nd
	$^{15}\text{N}_\delta$	114.7	Ser427	H_N	8.73
Pro414	H_N	-		$^{15}\text{N}_\text{H}$	116.7
	$^{15}\text{N}_\text{H}$	nd		$^{13}\text{C}_\alpha$	63.2
	$^{13}\text{C}_\alpha$	62.3		$^{13}\text{C}_\beta$	62.9
	$^{13}\text{C}_\beta$	34.1	Trp428	H_N	8.53
Asn415	H_N	8.89		$^{15}\text{N}_\text{H}$	124.1
	$^{15}\text{N}_\text{H}$	124.0		$^{13}\text{C}_\alpha$	60.8
	$^{13}\text{C}_\alpha$	51.9		$^{13}\text{C}_\beta$	30.5
	$^{13}\text{C}_\beta$	35.9		$\text{H}_{\text{N}\epsilon 1}$	nd
	$\text{H}_{\text{N}\delta}$	6.72, 7.56		$^{15}\text{N}_{\epsilon 1}$	nd
	$^{15}\text{N}_\delta$	110.4	Tyr429	H_N	8.59
Thr416	H_N	7.22		$^{15}\text{N}_\text{H}$	120.8
	$^{15}\text{N}_\text{H}$	113.1		$^{13}\text{C}_\alpha$	61.1
	$^{13}\text{C}_\alpha$	58.0		$^{13}\text{C}_\beta$	39.4
	$^{13}\text{C}_\beta$	71.8	Gln430	H_N	8.99
Val417	H_N	8.53		$^{15}\text{N}_\text{H}$	116.0
	$^{15}\text{N}_\text{H}$	122.9		$^{13}\text{C}_\alpha$	57.5
	$^{13}\text{C}_\alpha$	66.1		$^{13}\text{C}_\beta$	29.7
	$^{13}\text{C}_\beta$	34.8		$\text{H}_{\text{N}\epsilon}$	nd
Arg418	H_N	8.54		$^{15}\text{N}_\epsilon$	nd
	$^{15}\text{N}_\text{H}$	118.9	Ser431	H_N	8.02
	$^{13}\text{C}_\alpha$	60.5		$^{15}\text{N}_\text{H}$	115.8
	$^{13}\text{C}_\beta$	30.6		$^{13}\text{C}_\alpha$	61.1
	$\text{H}_{\text{N}\epsilon}$	nd		$^{13}\text{C}_\beta$	63.0
	$^{15}\text{N}_\epsilon$	nd	Lys432	H_N	7.30
Ala419	H_N	8.14		$^{15}\text{N}_\text{H}$	119.6
	$^{15}\text{N}_\text{H}$	122.3		$^{13}\text{C}_\alpha$	53.5
	$^{13}\text{C}_\alpha$	55.5		$^{13}\text{C}_\beta$	30.1
	$^{13}\text{C}_\beta$	18.7	Gly433	H_N	7.65
His420	H_N	8.23		$^{15}\text{N}_\text{H}$	119.6
	$^{15}\text{N}_\text{H}$	117.6		$^{13}\text{C}_\alpha$	46.2
	$^{13}\text{C}_\alpha$	56.9	Arg434	H_N	7.63
	$^{13}\text{C}_\beta$	28.7		$^{15}\text{N}_\text{H}$	112.5
Ala421	H_N	9.50		$^{13}\text{C}_\alpha$	56.9
	$^{15}\text{N}_\text{H}$	120.9		$^{13}\text{C}_\beta$	27.2
	$^{13}\text{C}_\alpha$	54.9		$\text{H}_{\text{N}\epsilon}$	nd
	$^{13}\text{C}_\beta$	19.3	Asn435	$^{15}\text{N}_\epsilon$	nd

Table A14 (continuation)					
Residue	Resonance		Residue	Resonance	
Asn435	H _N	7.26	Ser450	H _N	8.38
	¹⁵ N _H	117.3		¹⁵ N _H	120.4
	¹³ C _α	53.0		¹³ C _α	57.1
	¹³ C _β	40.3		¹³ C _β	63.7
	H _{Nδ}	7.06, 7.52	Asp451	H _N	7.97
	¹⁵ N _δ	114.7		¹⁵ N _H	128.4
Asp436	H _N	9.00		¹³ C _α	50.8
	¹⁵ N _H	124.7	Pro452	¹³ C _β	41.8
	¹³ C _α	58.1		H _N	-
	¹³ C _β	39.6		¹⁵ N _H	nd
Phe437	H _N	8.23		¹³ C _α	61.6
	¹⁵ N _H	117.5		¹³ C _β	30.7
	¹³ C _α	58.8	Ser453	H _N	8.04
	¹³ C _β	44.3		¹⁵ N _H	115.8
Gly438	H _N	8.08		¹³ C _α	61.4
	¹⁵ N _H	108.5	Asn454	¹³ C _β	63.4
	¹³ C _α	45.6		H _N	8.37
Cys439	H _N	7.76		¹⁵ N _H	120.2
	¹⁵ N _H	117.4		¹³ C _α	52.8
	¹³ C _α	52.4		¹³ C _β	41.2
	¹³ C _β	39.2		H _{Nδ}	7.08, 7.49
Asp440	H _N	7.26	Gly455	¹⁵ N _δ	112.8
	¹⁵ N _H	119.1		H _N	8.91
	¹³ C _α	56.0		¹⁵ N _H	112.4
	¹³ C _β	42.2	Ser456	¹³ C _α	47.1
Phe441	H _N	8.29		H _N	8.91
	¹⁵ N _H	113.9		¹⁵ N _H	120.8
	¹³ C _α	59.1		¹³ C _α	58.8
	¹³ C _β	36.0		¹³ C _β	63.9
Ser442	H _N	8.74	Cys457	H _N	8.63
	¹⁵ N _H	114.5		¹⁵ N _H	123.7
	¹³ C _α	58.4		¹³ C _α	55.4
	¹³ C _β	61.4	Ser458	¹³ C _β	41.5
Gly443	H _N	8.01		H _N	8.59
	¹⁵ N _H	106.6		¹⁵ N _H	124.5
	¹³ C _α	46.0		¹³ C _α	57.2
Thr444	H _N	7.90		¹³ C _β	63.4
	¹⁵ N _H	107.6	Phe459	H _N	9.41
	¹³ C _α	62.5		¹⁵ N _H	129.6
	¹³ C _β	71.5		¹³ C _α	58.0
Gly445	H _N	8.07	Leu460	¹³ C _β	38.4
	¹⁵ N _H	110.6		H _N	8.01
	¹³ C _α	44.7		¹⁵ N _H	125.1
Ala446	H _N	9.55		¹³ C _α	56.4
	¹⁵ N _H	127.5		¹³ C _β	42.3
	¹³ C _α	50.8	Ser461	H _N	7.72
	¹³ C _β	22.2		¹⁵ N _H	119.2
Ile447	H _N	8.17		¹³ C _α	59.6
	¹⁵ N _H	121.6		¹³ C _β	65.0
	¹³ C _α	59.9			
	¹³ C _β	37.9			
Thr448	H _N	9.39			
	¹⁵ N _H	122.8			
	¹³ C _α	58.7			
	¹³ C _β	70.6			
Ser449	H _N	8.44			
	¹⁵ N _H	119.9			
	¹³ C _α	58.3			
	¹³ C _β	64.3			

Table A15. Structural statistics parameters for the ensemble of the 20 lowest target function conformers calculated for peptide LytA₂₃₉₋₂₅₂ in aqueous solution and in micelles. ^a Close contacts are H atoms at distance < 1.6 Å and heavy atoms at distance < 2.2 Å. ^b χ_2 (χ_{21} in I244) range is also < 10°. ^c χ_2 (χ_{21} in I244) range is < 30°. ^d χ_3 and χ_4 ranges are also < 10°.

	Aqueous solution	DPC micelles	SDS micelles
Number of distance restraints			
Intraresidue ($i - j = 0$)	39	75	93
Sequential ($ i - j = 1$)	23	45	65
Medium range ($1 < i - j < 5$)	23	36	79
Long range ($ i - j \geq 5$)	87	0	0
Total number	172	156	237
Averaged total number per residue	12.3	11.1	16.9
Number of dihedral angle constraints			
Number of restricted ϕ angles	12	12	12
Number of restricted ψ angles	12	12	11
Total number	24	24	23
Maximum violations per structure			
Distance (Å)	0.12	0.00	0.00
Dihedral angle (°)	0.4	0.0	0.1
Number of close contacts ^a	0	0	0
Deviations from ideal geometry			
RMSD for Bond angles (°)	0.2	0.2	0.2
RMSD for Bond lengths (Å)	0.001	0.001	0.001
Pairwise RMSD (Å)			
Backbone atoms	0.3 ± 0.1	0.6 ± 0.2	0.2 ± 0.1
All heavy atoms	1.0 ± 0.1	1.4 ± 0.4	0.9 ± 0.2
Residues with ordered side chains			
χ_1 range < 10°	W241 ^b , K243 ^b , I244 ^b , K247, W248 ^b , Y249, Y250, F251	T239 ^b , W241 ^b , I244 ^b , K247 ^b , W248, Y249, F251	W241 ^b , K242 ^b , I244 ^c , K247 ^d , W248 ^b , Y249, F251
χ_1 range < 30°	K242 ^c , D246	K242, K243, Y250	K243, D246, N252
Ramachandran plot (%)			
Most favoured regions	72.7	100	90.9
Additionally allowed regions	27.3	0	9.1
Generously allowed regions	0	0	0
Disallowed regions	0	0	0

Table A16. Structural statistics parameters for the ensemble of the 20 lowest target function conformers calculated for peptide LytA₂₅₉₋₂₇₂ in aqueous solution and in micelles. ^a Close contacts are H atoms at distance < 1.6 Å and heavy atoms at distance < 2.2 Å. ^b χ_2 range is also < 10°. ^c χ_2 range is < 30°.

	Aqueous solution	DPC micelles	SDS micelles
Number of distance restraints			
Intraresidue ($i - j = 0$)	73	78	85
Sequential ($ i - j = 1$)	41	49	51
Medium range ($1 < i - j < 5$)	32	39	14
Long range ($ i - j \geq 5$)	57	0	0
Total number	203	166	150
Averaged total number per residue	14.5	11.9	10.7
Number of dihedral angle constraints			
Number of restricted ϕ angles	12	11	10
Number of restricted ψ angles	12	11	10
Total number	24	22	20
Maximum violations per structure			
Distance (Å)	0.00	0.49	0.00
Dihedral angle (°)	0.0	6.1	0.0
Number of close contacts ^a			
Deviations from ideal geometry			
RMSD for Bond angles (°)			
RMSD for Bond lengths (Å)			
Pairwise RMSD (Å)			
Backbone atoms	0.09 ± 0.04	0.3 ± 0.2	0.3 ± 0.2
All heavy atoms	0.8 ± 0.2	0.8 ± 0.2	1.6 ± 0.4
Residues with ordered side chains			
χ_1 range < 10°	W261 ^b , V262, K263 ^b , Y264, K265 ^b , T267, W268 ^b , Y269, Y270, L271 ^b	W261 ^b , V262, K263 ^b , Y264, K265 ^b , T267, W268 ^b , Y269, Y270, L271 ^b	K265 ^b , T267, W268 ^b , Y269, Y270
χ_1 range < 30°	-	-	Y264, L271
Ramachandran plot (%)			
Most favoured regions	72.7	63.6	94.1
Additionally allowed regions	27.3	36.4	5.9
Generously allowed regions	0	0	0
Disallowed regions	0	0	0

Table A17. Structural statistics parameters for the ensemble of the 20 lowest target function conformers calculated for peptide LytA₂₃₉₋₂₇₂ in aqueous solution and in SDS micelles. ^a Close contacts are H atoms at distance < 1.6 Å and heavy atoms at distance < 2.2 Å. ^b χ_2 (χ_{21} in I244, V262 and T267) range is also < 10°.

	Aqueous solution	SDS micelles
Number of distance restraints		
Intraresidue ($i - j = 0$)	162	207
Sequential ($ i - j = 1$)	36	86
Medium range ($1 < i - j < 5$)	39	97
Long range ($ i - j \geq 5$)	41	0
Total number	278	390
Averaged total number per residue	8.2	11.5
Number of dihedral angle constraints		
Number of restricted ϕ angles	32	31
Number of restricted ψ angles	32	30
Total number	64	61
Maximum violations per structure		
Distance (Å)	0.00	0.00
Dihedral angle (°)	0.0	0.0
Number of close contacts ^a		
Deviations from ideal geometry		
RMSD for Bond angles (°)		
RMSD for Bond lengths (Å)		
Pairwise RMSD (Å)		
Backbone atoms		
Full length	-	3 ± 1
Residues 239–252*	0.3 ± 0.1	0.5 ± 0.3
Residues 259–272*	0.6 ± 0.4	0.7 ± 0.3
All heavy atoms		
Full length	-	4 ± 2
Residues 239–252*	1.5 ± 0.4	1.4 ± 0.5
Residues 259–272*	1.5 ± 0.5	1.5 ± 0.4
Residues with ordered side chains		
χ_1 range < 10°	W241 ^b , I244, W248 ^b , Y249, K258, W261 ^b , W268, Y269, Y270, L271	W241 ^b , I244, W248, Y269, Y270, L271 ^b
χ_1 range < 30°	K242, E253, K263, K265, T267	K242, K243, Y249, F251, E254, M257, W261, K263, Y264
Ramachandran plot (%)		
Most favoured regions	67.4	90.0
Additionally allowed regions	29.8	10.0
Generously allowed regions	2.8	0
Disallowed regions	0	0

FIGURES

Figure A1. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of LytA_{239–252} peptide sequence. Conditions: H₂O/D₂O 9:1 v/v (black bars), 30 mM [D₃₈]-DPC in H₂O/D₂O 9:1 v/v (grey bars), 30 mM [D₂₅]-SDS in H₂O/D₂O 9:1 v/v (white bars), and 30 % vol. TFE in H₂O/D₂O 9:1 v/v (stripped bars). All samples were at pH 3.0, and spectra were measured at 25 °C, except the sample in DPC micelles that was measured at 35 °C. $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

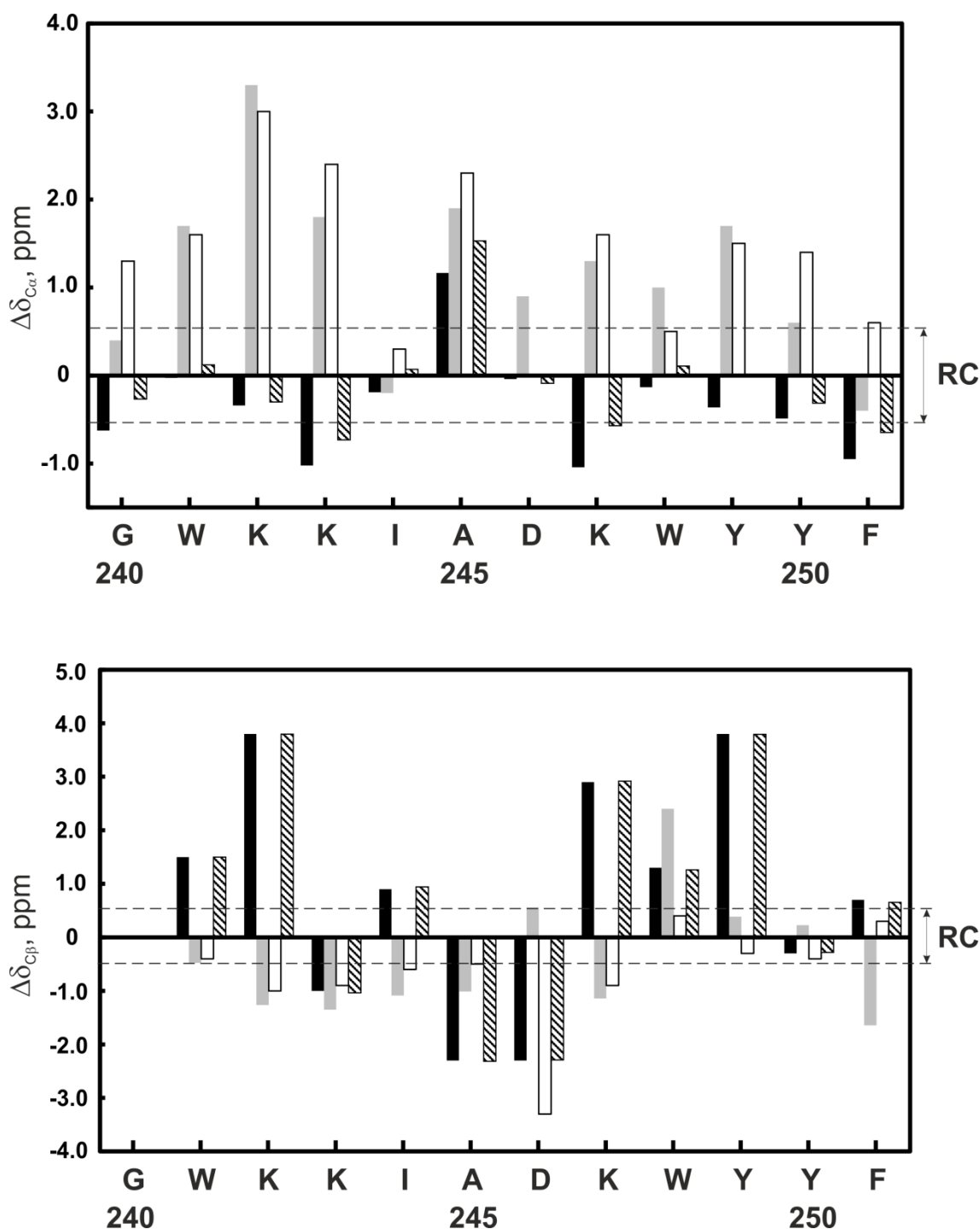


Figure A2. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of LytA_{259–272} peptide sequence. Conditions: H₂O/D₂O 9:1 v/v (black bars), 30 mM [D₃₈]-DPC in H₂O/D₂O 9:1 v/v (grey bars), 30 mM [D₂₅]-SDS in D₂O (white bars), and 30 % vol. TFE in H₂O/D₂O 9:1 v/v (stripped bars). All samples were at pH 3.0, and spectra were measured at 25 °, except the sample in DPC micelles that was measured at 35 °C. $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. *Not determined values.

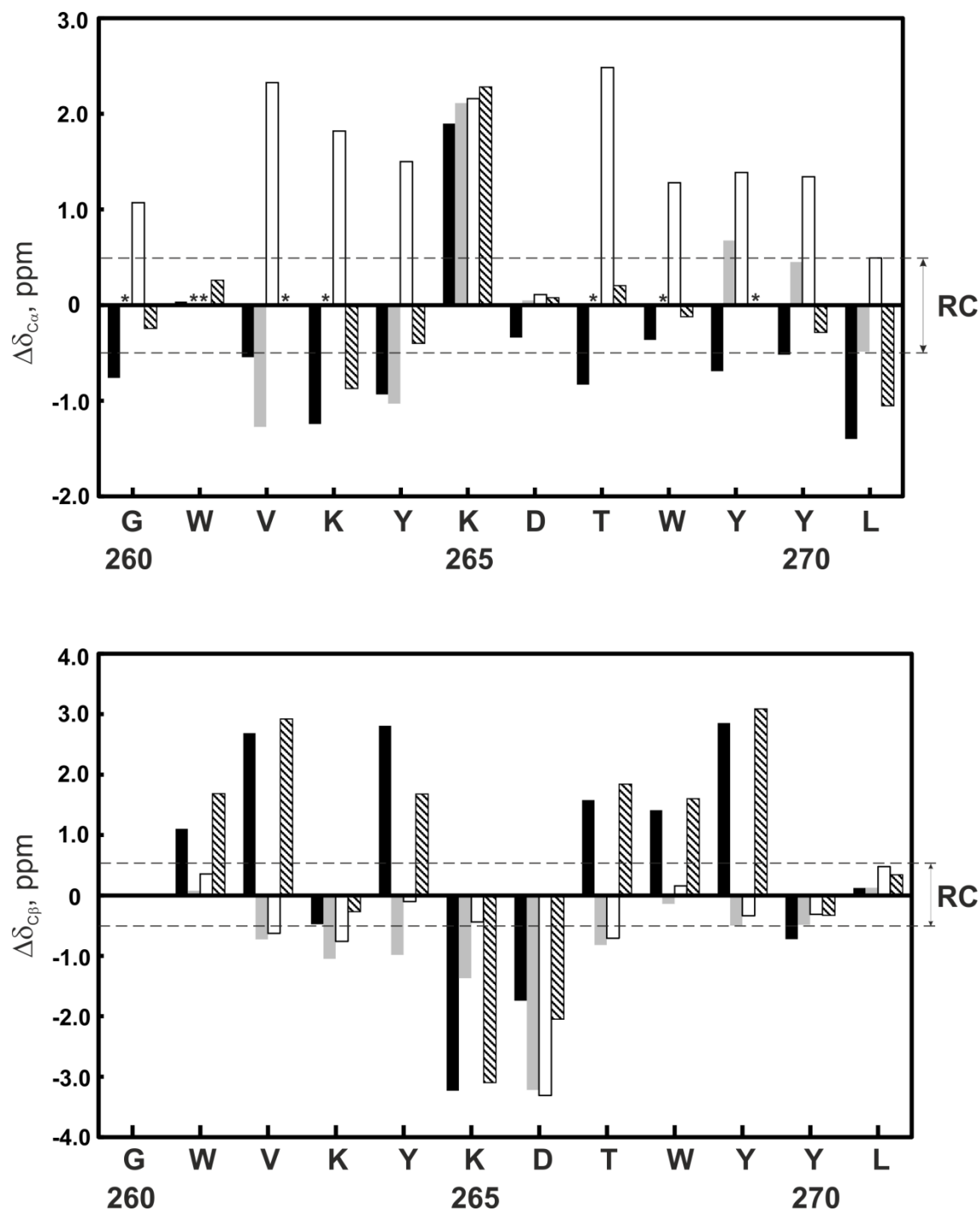


Figure A3. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of LytA_{239–272} peptide sequence. Conditions: D₂O (black bars), 30 mM [D₃₈]-DPC in D₂O (grey bars), 30 mM [D₂₅]-SDS in D₂O (white bars), and 30 % vol. TFE in H₂O/D₂O 9:1 v/v (stripped bars). All samples were at pH 3.0, and spectra were measured at 35 °. $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. *Not determined values.

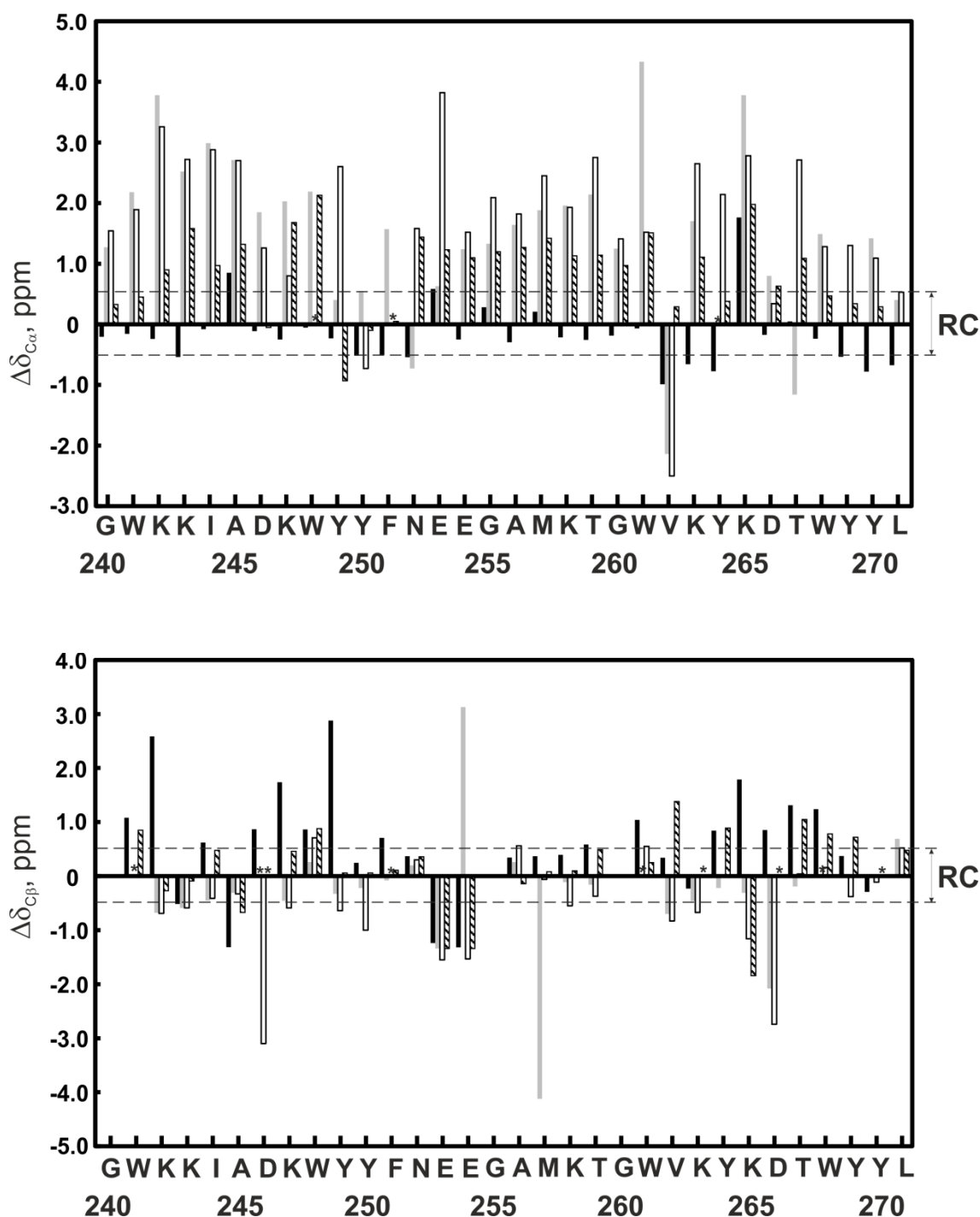


Figure A4. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of SESYV11 peptide sequence. Conditions: D₂O at 5 °C (black bars), and 30 mM [D₃₈]-DPC in D₂O at 25 °C (grey bars). All samples were at pH 3.0. $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. *Not determined values.

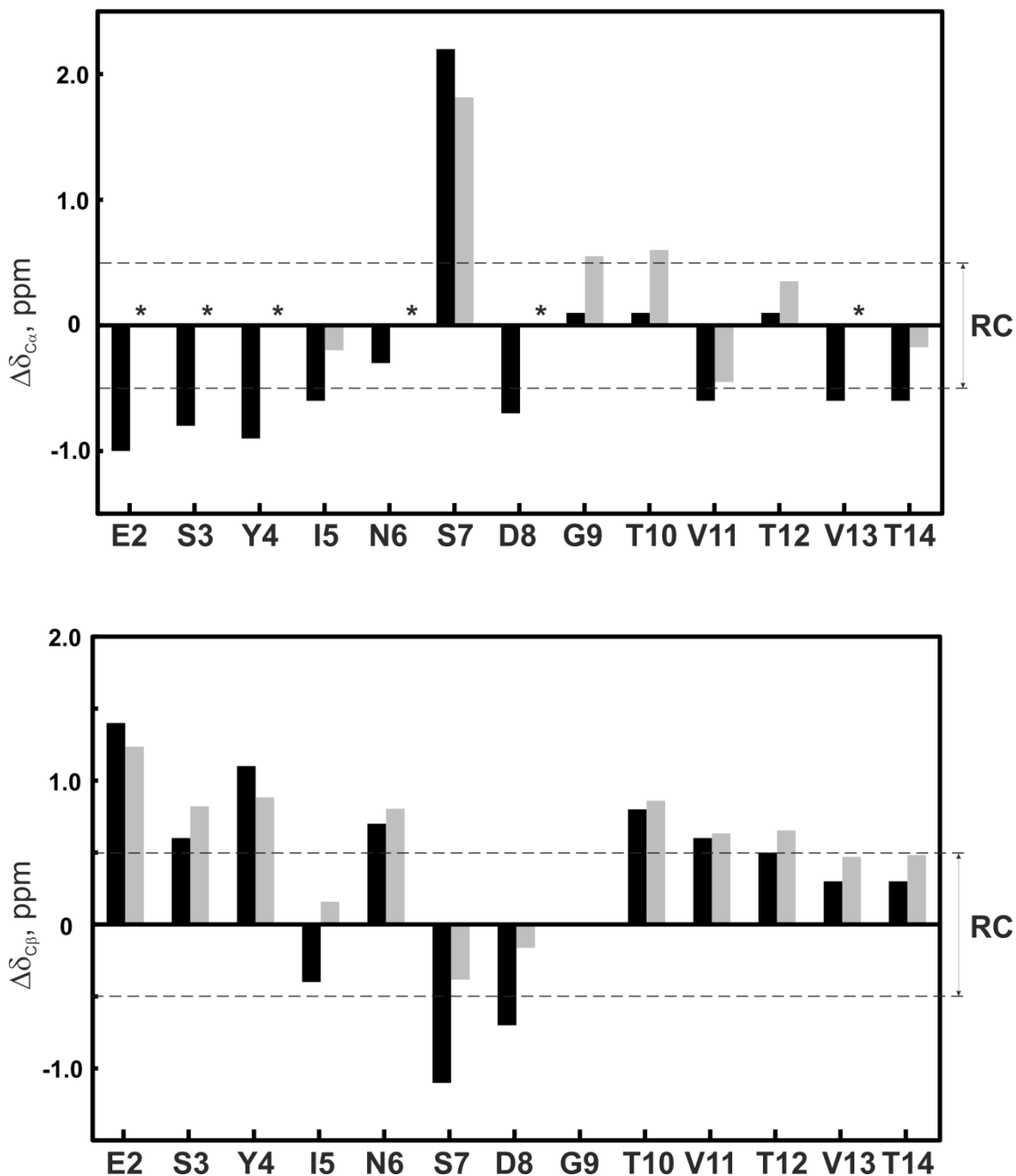


Figure A5. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of SESYW11 peptide sequence. Conditions: D₂O at 5 °C (black bars), and 30 mM [D₃₈]-DPC in D₂O at 25 °C (grey bars). All samples were at pH 3.0. $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. *Not determined values.

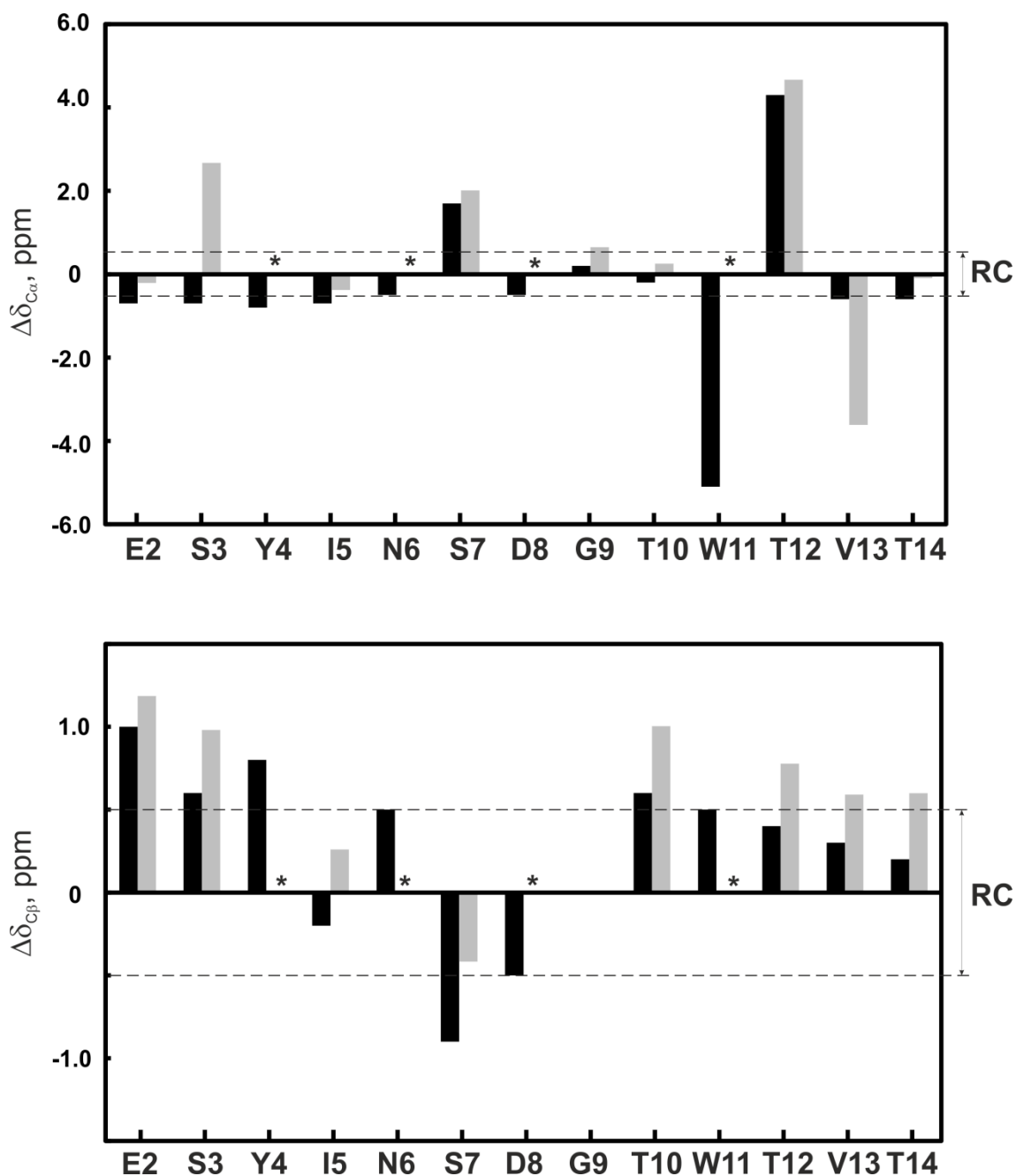


Figure A6. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of peptide sequence in aqueous solution (H₂O/D₂O 9:1 v/v, or D₂O) at 25 °C and pH 3.0. Peptides shown are K3W5–LytA_{239–252} (black bars) and W5K10–LytA_{239–252} (grey bars). $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

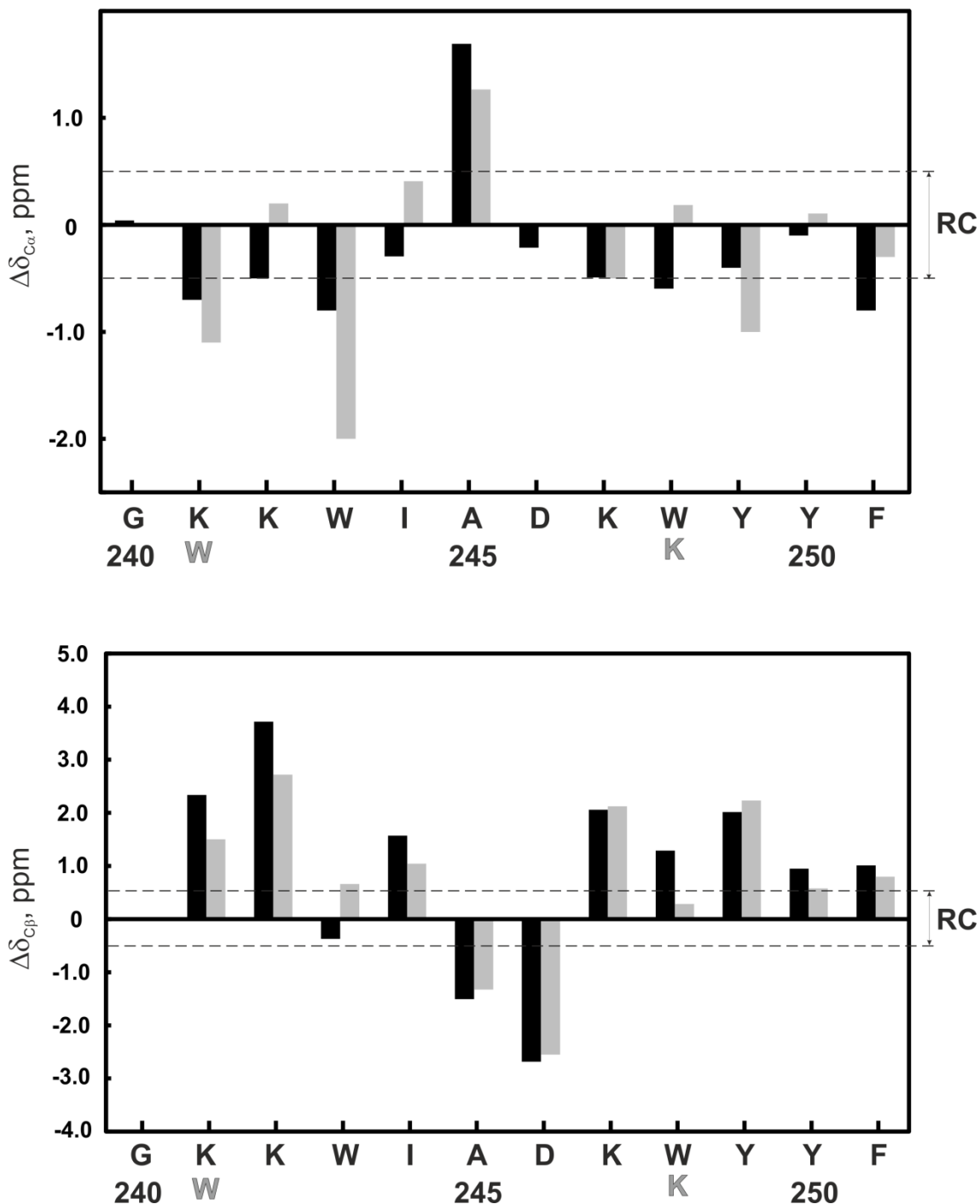


Figure A7. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of peptide sequence in 30 mM $[D_{38}]$ -DPC (in H_2O/D_2O 9:1 v/v or D_2O) at 25 °C and pH 3.0. Peptides shown are K3W5-LytA₂₃₉₋₂₅₂ (black bars) and W5K10-LytA₂₃₉₋₂₅₂ (grey bars). $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. *Not determined values.

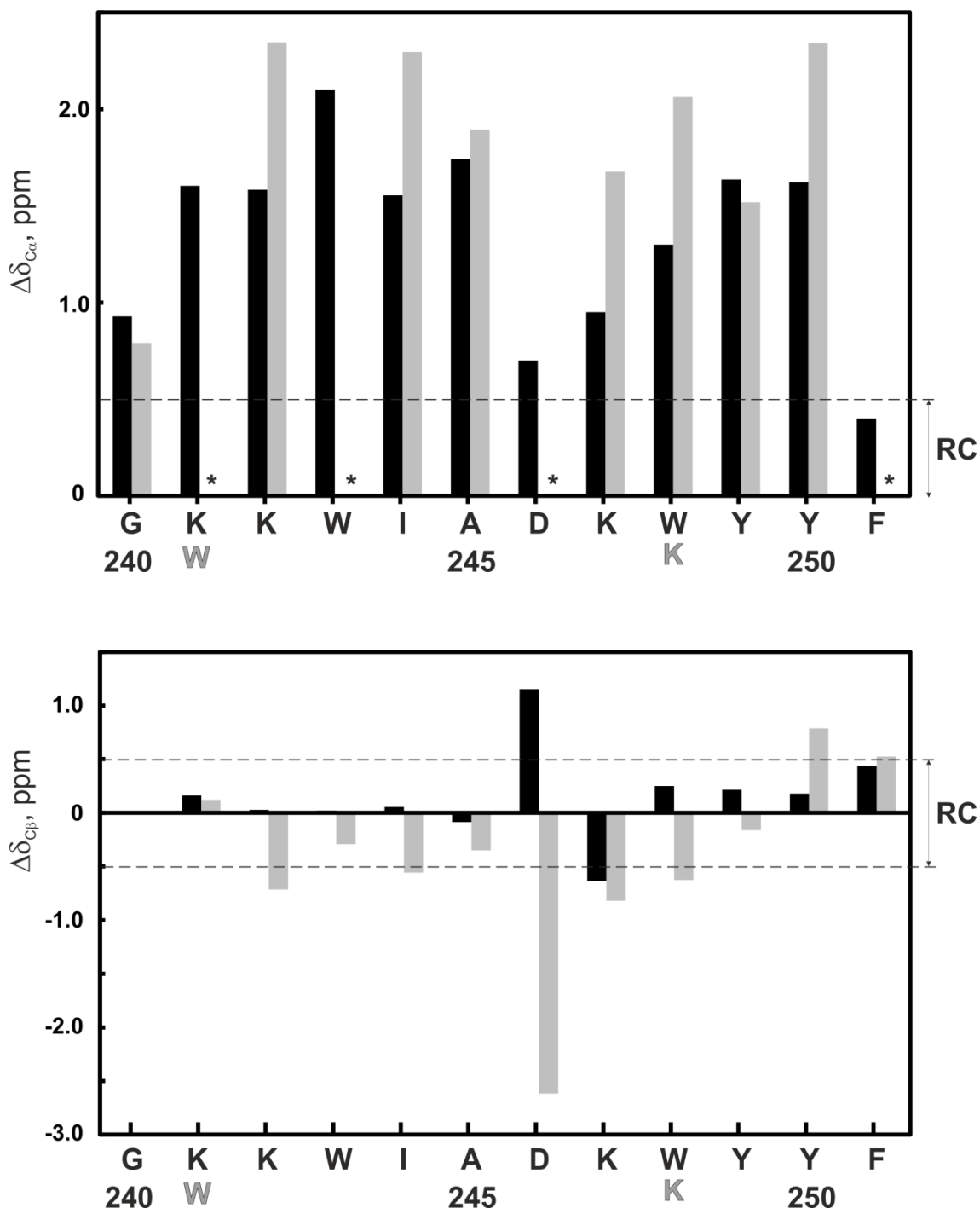


Figure A8. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of peptide sequence in H₂O/D₂O 9:1 v/v, at 25 °C and pH 3.0. Peptides shown are S3S10–LytA_{239–252} (black bars), I3V10–LytA_{239–252} (grey bars), and I5Y6T11T13–LytA_{239–252} (white bars). $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range.

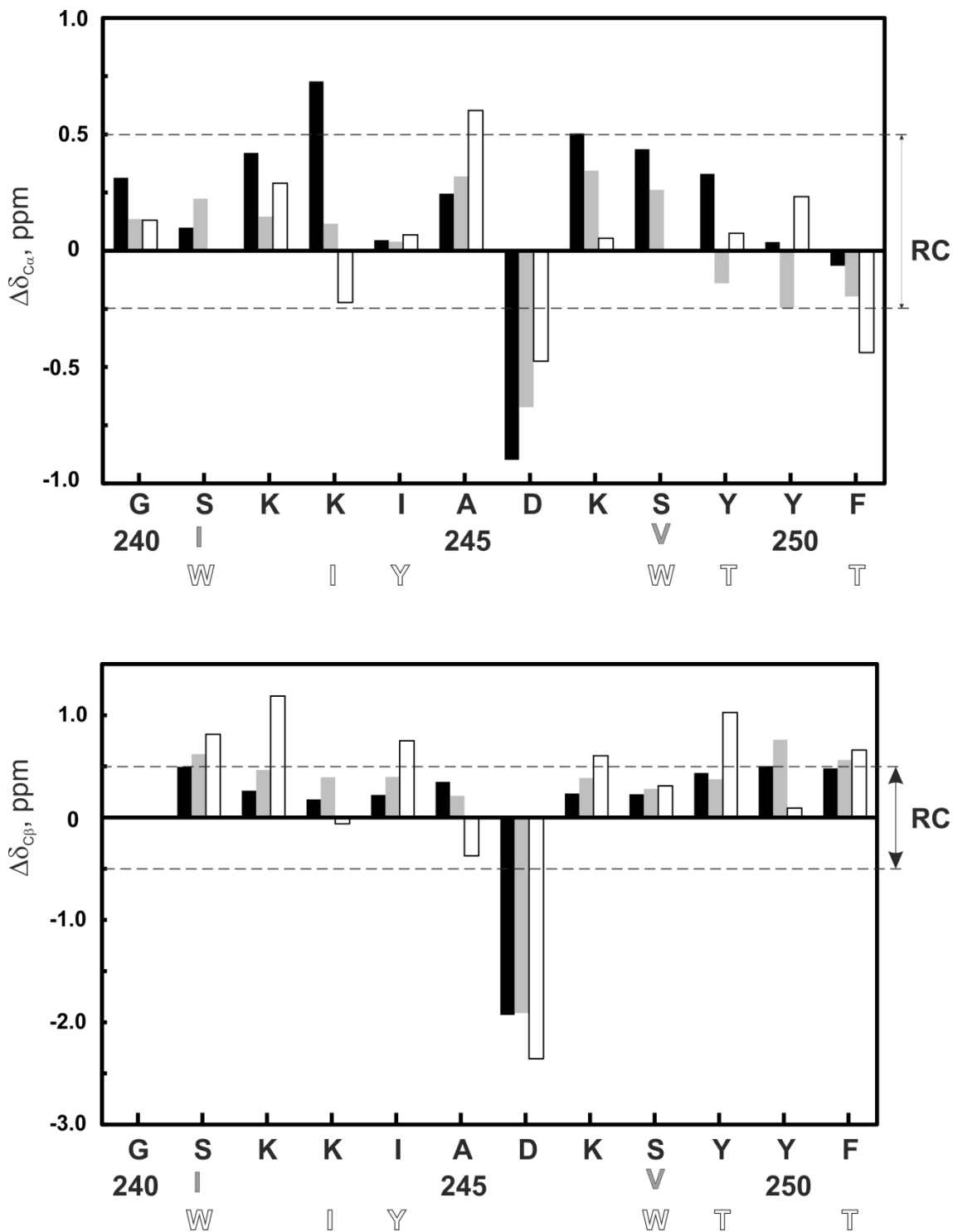


Figure A9. Bar plots of $\Delta\delta_{C\alpha}$ ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{\text{observed}} - \delta_{C\alpha}^{\text{RC}}$) (**top**) and $\Delta\delta_{C\beta}$ ($\Delta\delta_{C\beta} = \delta_{C\beta}^{\text{observed}} - \delta_{C\beta}^{\text{RC}}$) (**bottom**) as a function of peptide sequence in 30 mM $[D_{38}]$ -DPC (in H_2O/D_2O 9:1 v/v or D_2O), at 25 °C and pH 3.0. Peptides shown are S3S10-LytA₂₃₉₋₂₅₂ (black bars), I3V10-LytA₂₃₉₋₂₅₂ (grey bars), and I5Y6T11T13-LytA₂₃₉₋₂₅₂ (white bars). $\delta_{C\alpha}^{\text{RC}}$ and $\delta_{C\beta}^{\text{RC}}$ values were taken from (Wishart et al., 1995). The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. *Not determined values.

