UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS

Departamento de Bioquímica y Biología Molecular

TESIS DOCTORAL

Design and structural characterization of minimized systems to study biomolecular interactions through NMR

Diseño y caracterización estructural de sistemas minimizados para el estudio de interacciones biomoleculares mediante RMN

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Angélica Inés Partida Hanon

Directoras

Marta Bruix Bayés
María Ángeles Jiménez López

Madrid, 2019

DESIGN AND STRUCTURAL CHARACTERIZATION OF MINIMIZED SYSTEMS TO STUDY BIOMOLECULAR INTERACTIONS THROUGH NMR

DISEÑO Y CARACTERIZACIÓN ESTRUCTURAL DE SISTEMAS MINIMIZADOS PARA EL ESTUDIO DE INTERACCIONES BIOMOLECULARES MEDIANTE RMN

Memoria con la que **Angélica Inés Partida Hanon** aspira al Grado de Doctor en Bioquímica, Biología Molecular y Biomedicina

Tesis dirigida por

Dra. Marta Bruix Bayés

Dra. María Ángeles Jiménez López

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Instituto de Química Física "Rocasolano"

Departamento de Química Física Biológica

Tutor

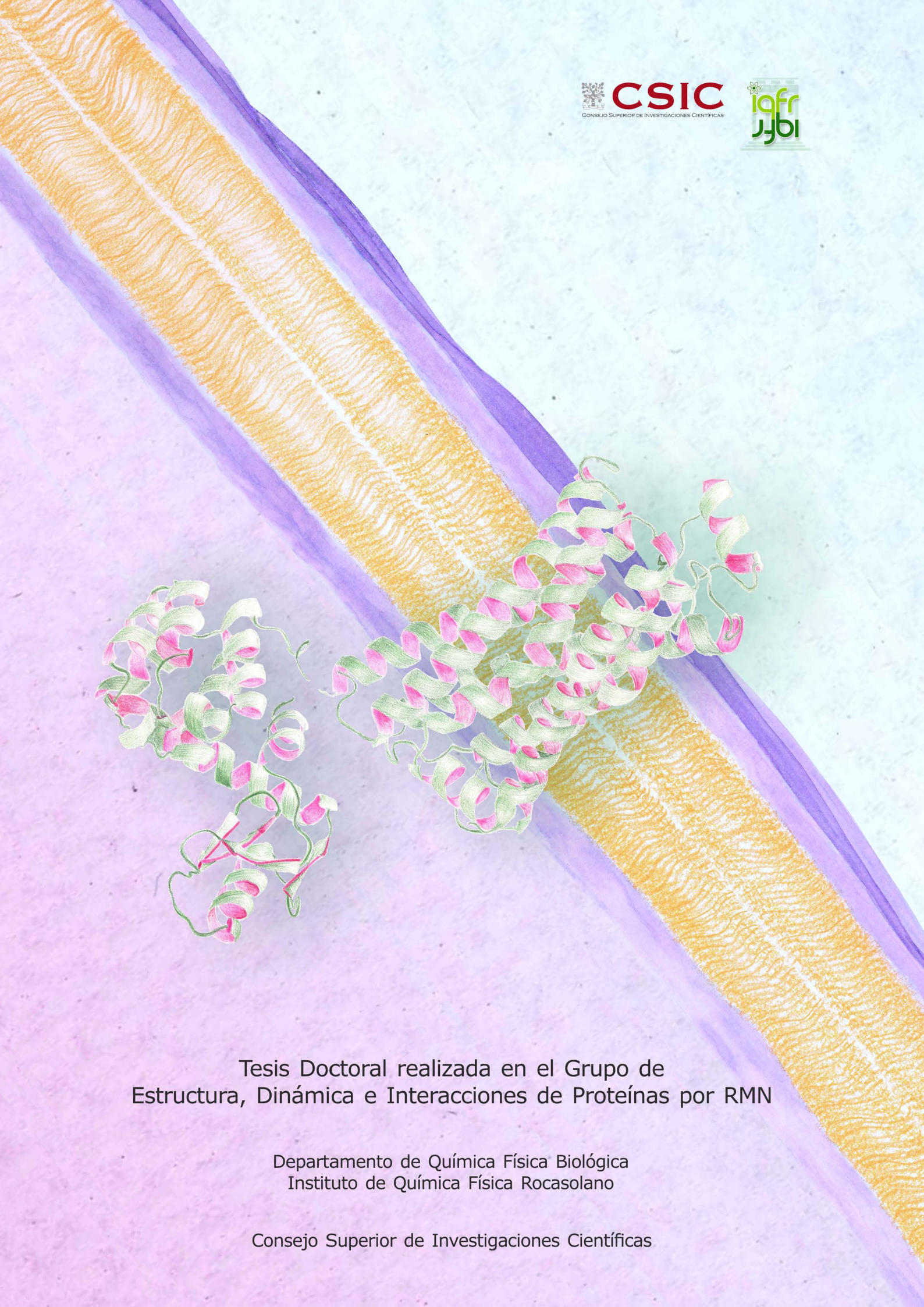
Dr. Álvaro Martínez del Pozo

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Químicas

Departamento de Bioquímica y Biología Molecular I

Madrid 2018



Tesis Doctoral realizada en el Grupo de
Estructura, Dinámica e Interacciones de Proteínas por RMN

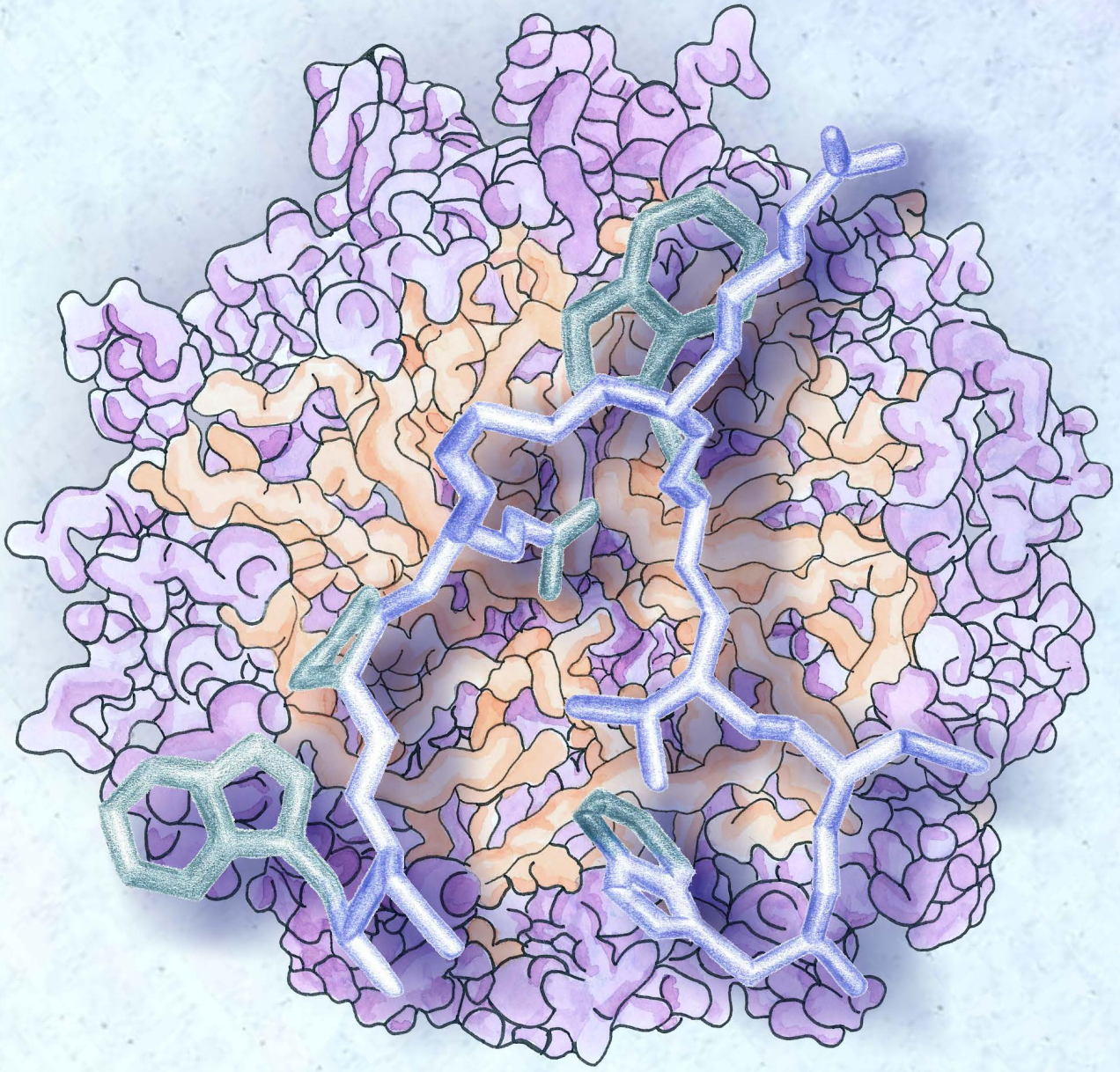
Departamento de Química Física Biológica
Instituto de Química Física Rocasolano

Consejo Superior de Investigaciones Científicas

ANGÉLICA INÉS PARTIDA HANON

Design and Structural Characterization of Minimized Systems to Study Biomolecular Interactions Through NMR

Diseño y Caracterización Estructural de Sistemas Minimizados para el Estudio de Interacciones Biomoleculares Mediante RMN



Tesis Doctoral dirigida por las Dras.
Marta Bruix Bayés y María Ángeles Jiménez López

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

Facultad de Ciencias Químicas
Universidad Complutense de Madrid

Cover: α-mating pheromone from the fungal pathogen *Fusarium oxysporum* within a dodecylphosphocholine micelle and its cognate transmembrane G protein-coupled receptor Ste2.
Technique: watercolor and colored pencils.

Angélica Partida Hanon

Agradecimientos

El trabajo de esta Tesis Doctoral se ha llevado a cabo en el grupo de Estructura, Dinámica e Interacciones de Proteínas por RMN, en el Departamento de Química Física Biológica del Instituto de Química Física "Rocasolano", en el Consejo Superior de Investigaciones Científicas durante el periodo de septiembre de 2014 a septiembre de 2018. La ayuda económica ha sido proporcionada por: un contrato a cargo del Proyecto CNB CS Interactomics S2010/BMD-2305 del Programa de Actividades de I+D de la Comunidad de Madrid, de los programas de becas del Consejo Nacional de Ciencia y Tecnología, y las ayudas para la formación de Doctores de Fundación Universia del Banco Santander.

Deseo expresar mi más sincero agradecimiento a:

Dra. Marta Bruix, codirectora de este trabajo, y quien ha sido mi madre científica y la persona con la que más he convivido de continuo en toda mi vida. Además, ha sido mi mayor referencia durante los últimos cuatro años de intenso trabajo contribuyendo a mi comprensión del mundo científico y su entorno.

Dra. María Ángeles Jiménez, codirectora de este trabajo, y quien ha sido una guía para comprender esta disciplina de la RMN y ha compartido sus conocimientos y experiencia de una manera tan amena y fácil.

Profesor Said Infante Gil por ser mi primer mentor científico y enseñarme el apasionante mundo de la Probabilidad y Estadística desde que soy niña y al Profesor Álvaro Martínez del Pozo, por ser mi primer profesor universitario en España y haber conseguido que hoy esté aquí.

También quiero reconocer con gratitud el trabajo y ayuda técnica de los doctores Douglas Laurents, David Pantoja, Miguel Ángel Treviño y José Manuel Pérez Cañadillas, por darme siempre seguridad para conseguir el experimento perfecto. Asímismo, quiero recordar aquí a mis compañeros de RMN: Sergio, es imposible tener un mejor compañero de despacho, Irene, Carlos, Israel, Miguel G., Miguel M., Héctor, Belén y Paula. No quiero olvidar a los

que han estado aquí de estancia y con los que he pasado grandes momentos: Federica, Santi y Vero.

A mis compañeros de la casita Mössbauer, que me han recibido como una más del equipo y con quienes he pasado mi año más estresante de escritura: Juan, Pepe, María y Guiomar. No me olvido de quienes también me han orientado en mi carrera científica: Josemi Mancheño, Ulises Acuña, Paloma Martínez y Mercedes Oñaderra: grandes del mundo de la cristalografía, la fluorescencia y las técnicas instrumentales Químicas y Bioquímicas.

Quiero también agradecer el apoyo continuado de mis compañeros informáticos y telecos: Antonio, Victoria y Ángel y la dedicación del equipo de Gerencia: Antonio, Sagrario, Mar y las dos Glorias. Estoy especialmente agradecida por la ayuda y exquisita amabilidad del personal de la Biblioteca del Instituto: Esperanza, Feli, Vicky y Samuel.

También ocupan un sitio especial en esta dedicatoria las personas que me han ayudado desde mi llegada a España en 2007, y a quienes considero mi familia: Maruja y Mey Fernández de Caleya, Raúl Sacristán, Pablo del Real, Félix Pantoja, Sagrario Fernández de Córdova, Pilar Goytia, Toñi Durán, David Chiner, Arancha Arroyo y Manuela de la Pedrosa. Agradezco también la fidelidad de mis amigas Raquel Fernández de Llano y Silvia Benito Cortés, a quienes he visto crecer profesionalmente.

Quiero recordar también con gratitud al equipo del Programa de Becas de la Fundación Universia del Banco Santander, entre quienes están Marta Gago, Sonia Viñas, Ramón Capdevila y Javier Roglá, personas que creyeron en mí y en mi proyecto desde que empecé mis estudios, y que nunca dejaron de apoyarme.

Finalmente, quiero agradecer la compañía de Antonio Tarruell Pellegrin, con el que he recorrido un camino paralelo durante los últimos años, compartiendo conmigo sus conocimientos e invaluable amistad.

Table of contents

List of figures **i**

List of tables **v**

Nomenclature **ix**

Resumen **1**

Abstract **5**

1 General introduction and objectives **9**

1.1 Background 10

1.2 Structural and functional insights in biomolecular recognition 10

1.2.1 Biological interactions and molecular recognition 13

1.2.1.1 Ionic interactions 13

1.2.1.2 Hydrogen bonds 13

1.2.1.3 Van der Waals interactions 14

1.2.1.4 The hydrophobic effect 14

1.3 The biological membrane 15

1.3.1 Properties and composition of the biological membrane 16

1.3.1.1 Lipids 16

1.3.1.2 Membrane proteins 18

1.3.1.3 Carbohydrates 18

1.4 Protein-protein and protein-peptide interactions (PPIs) 19

1.5 Protein and peptide interactions with membranes 20

1.6 Experimental methods to study biomolecular interactions 21
1.6.1 Mass spectrometry (MS) 22
1.6.2 Dynamic Light Scattering (DLS) 22
1.6.3 Circular Dichroism (CD) 23
1.6.4 Isothermal Titration Calorimetry (ITC) 24
1.6.5 Fluorescence assays 24
1.6.6 Small-angle X-ray scattering (SAXS) 24
1.6.7 X-ray crystallography 24
1.6.8 NMR spectroscopy 25
1.6.8.1 NMR as a tool to study protein interactions 25
1.6.8.2 Advantages and limitations of NMR 26
1.7 Theoretical methods to study biomolecular interactions 27
1.7.1 Molecular Docking 27
1.8 The use of minimized systems to surpass NMR limitations 28
1.8.1 Model peptides as protein mimics 28
1.8.1.1 Sequence-based peptide design 29
1.8.1.2 Structure-based peptide design 30
1.8.1.3 Protein function-based peptide design 30
1.8.2 Micelles as membrane mimetics 30
1.9 Global aims and structure of this Thesis 34
1.9.1 Organization of this Thesis 35

2 Structural study of the centrosomal TACC3-XMAP protein-protein interaction from a minimized system of peptides and cosolvents 39
2.1 Introduction and objectives 40
2.1.1 The TACC3-XMAP interaction promotes microtubule growth . . . 42
2.1.2 TACC3-XMAP interaction model 42
2.1.3 The advantages of the use of simplified systems 44
2.1.4 Objectives 46
2.2 Materials and methods 47
2.2.1 Chemicals 47
2.2.2 Peptide synthesis 47

2.2.3 CD spectroscopy . . . 48
2.2.4 NMR sample preparation . . . 49
2.2.5 NMR spectra acquisition . . . 49
2.2.6 NMR spectra assignment . . . 49
2.2.7 Structure calculation . . . 50
2.2.8 NMR estimation of helix populations and interaction mapping . . . 51
2.2.9 Computational modeling and NMR-driven docking . . . 51
2.3 Results . . . 53
2.3.1 Peptide design . . . 53
2.3.2 Circular dichroism of isolated peptides . . . 54
2.3.3 NMR characterization of isolated peptides . . . 56
2.3.4 CD and NMR characterization of peptide mixtures . . . 63
2.3.5 Molecular modeling of the TACC-α2 + Ct-TACC-α4 + XMAP-pCt assembly . . . 70
2.4 Discussion . . . 72
2.4.1 Short linear peptides are good models to mimic native protein structure 72
2.4.2 Short linear peptides are good models to study protein interactions . 73
2.4.3 Structure-function relationships . . . 74
2.5 Conclusions . . . 77

3 Structural analysis of protein-membrane interactions employing peptides derived from the HIV-1 ectodomain gp41 in cosolvents 81

3.1 Introduction and objectives . . . 82
3.1.1 Main structural components of the HIV viral particle . . . 82
3.1.2 The Env complex is essential for T-cell recognition and infection . . 83
3.1.3 Structural basis of the gp41 glycoprotein . . . 84
3.1.4 Broadly neutralizing antibodies against MPER of gp41 as therapeutic molecules . . . 86
3.1.5 Simplification of the Env complex system . . . 88
3.1.6 Objectives . . . 90
3.2 Materials and methods . . . 92
3.2.1 Chemicals . . . 92

3.2.2 Peptide synthesis . . . 92
3.2.3 NMR sample preparation . . . 92
3.2.4 NMR spectra acquisition, assignment and structure calculation . . . 93
3.2.5 NMR spectra assignment and interaction mapping . . . 93
3.2.6 NMR estimation of helix populations . . . 94
3.2.7 Structure calculation . . . 94
3.3 Results . . . 95
3.3.1 Design of optimized peptide epitopes for 10E8 . . . 95
3.3.2 Secondary structure of the MPERb and MPER-H2 peptides . . . 96
3.3.3 3D NMR structure of the MPER peptides . . . 100
3.3.4 MPER peptides interact with membrane mimetics . . . 104
3.4 Discussion . . . 106
3.4.1 H1 and H2 in MPER are structurally independent modules . . . 106
3.4.2 Bending in the helices can be found in different regions depending on the solvent . . . 107
3.4.3 Peptide-membrane interaction and insertion . . . 109
3.4.4 The structural flexibility found in the peptides might contribute to the membrane fusion mechanism . . . 110
3.4.5 Broadly neutralizing antibodies bind differently to distinct MPER epitopes . . . 110
3.5 Conclusions . . . 115

4 Structure and molecular association in sequences derived from the antimicrobial peptide protegrin-1 in the presence of membrane mimetics 117

4.1 Introduction and objectives . . . 118
4.1.1 The antibiotic crisis . . . 118
4.1.2 Antimicrobial peptides . . . 119
4.1.3 Classification . . . 121
4.1.4 Mechanisms of action of AMPs . . . 122
4.1.4.1 Membrane-active AMPs . . . 123
4.1.4.2 Intracellularly-active AMPs . . . 125
4.1.5 Important features of AMPs for structure-function relationships . . . 125
4.1.6 Protegrins: a promising family of antimicrobial peptides . . . 126

4.1.7 Previous structural studies on protegrins . . . 127
4.1.8 Objectives . . . 129
4.2 Materials and methods . . . 130
4.2.1 Chemicals . . . 130
4.2.2 Peptide synthesis . . . 130
4.2.3 CD spectroscopy . . . 131
4.2.4 NMR sample preparation . . . 131
4.2.5 NMR spectra acquisition . . . 132
4.2.6 NMR spectra assignment . . . 133
4.2.7 Structure calculation . . . 133
4.2.8 Dimer modeling and NMR-driven docking . . . 133
4.3 Results . . . 135
4.3.1 Peptide design . . . 135
4.3.2 CD data of PG-1 variant peptides . . . 135
4.3.3 NMR data reveals peptide β-hairpin tendency in all variants . . . 136
4.3.4 Structure calculation . . . 139
4.3.5 Peptide oligomerization . . . 143
4.4 Discussion . . . 146
4.4.1 Structures of PG-1 variants, the role of Cys residues . . . 146
4.4.2 Insights into the mechanism of action . . . 147
4.5 Conclusions . . . 150

5 Structure and molecular interactions of the α-mating pheromone from *Fusarium oxysporum* in the presence of membrane mimetics and other cosolvents 153
5.1 Introduction and objectives . . . 154
5.1.1 Morphological and physiological traits in ascomycetes . . . 154
5.1.2 General characteristics of *Fusarium* and filamentous fungi . . . 155
5.1.3 Economic and health impact of *Fusarium* species . . . 156
5.1.4 Fungal sex pheromones trigger hyphal growth . . . 158
5.1.5 Structural insights on the fungal α-pheromone of *S. cerevisiae* . . . 159
5.1.6 Surfactants as mimic molecules of the pheromone's receptor environment . . . 161

5.1.7 Objectives . . . 163
5.2 Materials and methods . . . 164
5.2.1 Chemicals . . . 164
5.2.2 Peptide synthesis . . . 164
5.2.3 Dynamic light scattering . . . 165
5.2.4 Mass spectrometry . . . 165
5.2.5 NMR sample preparation . . . 165
5.2.6 NMR spectra acquisition . . . 166
5.2.7 NMR spectra assignment . . . 167
5.2.8 Structure calculation based on NMR restraints . . . 167
5.2.9 Micelle interaction evaluation . . . 168
5.3 Results . . . 169
5.3.1 Peptide design and cosolvent selection . . . 169
5.3.2 NMR characterization of α-pheromone related peptides . . . 169
5.3.3 Ageing peptides yield different species depending on the environment 172
5.3.4 3D structure of α-pheromone related peptides . . . 179
5.3.5 Peptide interaction with micelles . . . 188
5.4 Discussion . . . 191
5.4.1 *Fusarium* α-pheromone's 3D structure . . . 191
5.4.2 *Fusarium* α-pheromone contains discrete structural and functional regions . . . 193
5.4.3 Additional insights into the pheromone's mechanisms of action, membrane interaction . . . 194
5.4.4 Model of the α-pheromone-membrane interaction . . . 194
5.5 Conclusions . . . 197

6 General conclusions **201**
6.1 General summary and general conclusions . . . 202
6.2 Publications derived from this Thesis . . . 204

7 Bibliography **207**

Appendices **237**

Appendix A Structural study of the centrosomal TACC3-XMAP protein-protein interaction from a minimized system of peptides and cosolvents 239

A.1 Chemical shift tables for peptide TACC-α2 240

A.2 Chemical shift tables for peptide Ct-TACC-α4 wt 243

A.3 Chemical shift tables for peptide Ct-TACC-α4 mut 246

A.4 Chemical shift tables for peptide XMAP-pCt 249

A.5 Chemical shift table for peptide mixtures 252

Appendix B Structural analysis of protein-membrane interactions employing peptides derived from the HIV-1 ectodomain gp41 in cosolvents 255

B.1 Chemical shift tables for peptide MPERb 256

B.2 Chemical shift tables for peptide MPER-H2 258

Appendix C Structure and molecular association in sequences derived from the antimicrobial peptide protegrin-1 in the presence of membrane mimetics 261

C.1 Chemical shift tables for peptide PG-T4 262

C.2 Chemical shift tables for peptide PG-C6C15 264

C.3 Chemical shift tables for peptide PG-C8C13 266

Appendix D Structure and molecular interactions of the α-mating pheromone from *Fusarium oxysporum* in the presence of membrane mimetics and other cosolvents 269

D.1 Chemical shift tables for peptide α-pheromone 270

D.2 Chemical shift tables for peptide scrambled 274

D.3 Chemical shift tables for peptide D-Ala^{1,2} 278

D.4 Chemical shift tables for peptide D-Ala^{6,7} 280

List of figures

1.1 Schematic diagram of the conformational selection model 12
1.2 Secondary structure stabilization through hydrogen bonds 14
1.3 Fluid mosaic membrane model 15
1.4 Main classification of the lipid components in the membrane 16
1.5 Most abundant phosphoglycerides 17
1.6 Different types of membrane proteins 18
1.7 Protein-protein interactions between subunits 19
1.8 Obligate and non-obligate protein-protein interactions 20
1.9 Biophysical techniques in biomolecular characterization 23
1.10 Design of peptide models as protein mimics. 29
1.11 Organic solvents and surfactants used as membrane mimetics in NMR . . . 31
1.12 Solubilization of integral membrane proteins 32

2.1 Schematic representation of the cell and centrosome cycles 40
2.2 Bipolar spindle assembly diagram 41
2.3 Cartoon representation of the XTACC3-XMAP interaction 43
2.4 Helical wheel representations for the structures formed by the peptides . . . 53
2.5 Far UV-CD spectra of individual peptides in different solvents 54
2.6 Far UV-CD spectra of individual peptides in different pH conditions 55
2.7 Far UV-CD spectra of individual peptides in different concentrations 56
2.8 Hα and Cα secondary chemical shifts observed for the peptides, pH 5.5 . . 58
2.9 Regions of the ^{1}H-^{1}H NOESY spectra 60
2.10 Ensemble of the 20 structures with the lowest target function values in 30 % TFE 62

2.11 Ensemble of the 20 structures with the lowest target function values in 20 mM DPC 63
2.12 Far UV-CD data of pair peptide mixtures in H_2O and 30 % TFE 65
2.13 Far UV-CD data of pair peptide mixtures in 20 mM DPC 66
2.14 Far UV-CD of trio peptide mixtures in different conditions. 67
2.15 Far UV-CD of trio peptide mixtures in H_2O and TFE at ratio 1:1:2 68
2.16 Superposition of the finger print region of the TOCSY NMR spectra 69
2.17 Weighted NMR chemical shift differences 69
2.18 Structural 3D computed model of TACC-α*2*, Ct-TACC-α4 wt and XMAP-pCt, showing residues affected by the interaction. 71
2.19 Alignment of sequences of the *Xenopus laevis* XMAP215 and XTACC3 with their human homologs. 75

3.1 Schematic image of the viral structure of HIV 83
3.2 Schematic representation of the cell recognition and fusion during infection 84
3.3 Domains and sequence of gp41 85
3.4 Structural elements of the peptides studied in this study. 87
3.5 α-helical wheel representations for MPERb and MPER-H2 peptides 95
3.6 Hα and Cα secondary chemical shifts observed for the peptides MPERb and MPER-H2 96
3.7 Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of MPERb 98
3.8 Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of MPER-H2 99
3.9 Ensemble of the 20 structures with the lowest target function values of MPERb and MPER-H2 in HFIP 102
3.10 Ensemble of the 20 structures with the lowest target function values of MPERb and MPER-H2 in DPC 103
3.11 Structures adopted by MPERb and MPER-H2 in HFIP and DPC. 104
3.12 Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of MPERb and MPER-H2 105
3.13 Comparison of the Hα secondary chemical shifts observed for the peptides 106
3.14 Comparison of the Hα secondary chemical shifts observed for all the peptides studied HFIP and DPC micelles 108
3.15 Models for gp41 MPER-TMD organization in HFIP and DPC micelles as inferred from the calculated NMR structures. 109

3.16 10E8, 4E10 and 2F5 epitope organization and putative mechanisms of antibody recognition. 112
3.17 Schematic representation of the Env complex and its antibody neutralization 113

4.1 Current AMPs statistical distribution based on number of residues and isoelectric point. 119
4.2 Examples of the different classes found in antimicrobial peptides 120
4.3 Cartoon representation of Gram-positive and Gram-negative bacterial cell membranes 122
4.4 Schematic representation of the most popular mechanisms of action of membrane-active AMPs. 124
4.5 Schematic representation of other mechanisms of action of membrane-active AMPs 124
4.6 Alignment of sequences of protegrins (PG) 1 to 5 126
4.7 90º rotated views in the horizontal axis of one conformer of PG-1. 127
4.8 CD spectra of 30 μM peptides PG-T4, PG-C6C15 and PG-C8C13 in H_2O and 30 mM DPC micelles, 5 ºC, pH 3.5. 135
4.9 Chemical shift deviations barplot of the peptides studied. 137
4.10 Barplot diagram of the Hα chemical shifts of the peptides in DPC and H_2O solvents 138
4.11 Comparison of TOCSY spectra under different conditions. 138
4.12 Ensemble of the 20 structures with the lowest target function values of PG-T4, PG-C6C15 and PG-C8C13 in H_2O 140
4.13 PG-T4, PG-C6C15 and PG-C8C13 structures calculated in 30 mM DPC . . 141
4.14 Structural 3D computed models. 144
4.15 Combined structural 3D computed models. 145
4.16 Models of possible interactions between PG-C6C15 peptides and biological membranes. 148
4.17 90º rotated views in the horizontal axis of the backbone superposition of one conformer for each peptide. 151

5.1 Biological cycle of ascomycetes fungi 155
5.2 Schematic diagram of the spores produced by ascomycetes fungi. 156
5.3 Chemotropic signaling routes in *Fusarium oxysporum* 159
5.4 Alignment of the sequences of *S. cerevisiae* and *F. oxysporum* α-pheromones 160

5.5 DPC and Gemini micelles as membrane mimetics. 162
5.6 Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of α-pheromone in H_2O and TFE 170
5.7 Selected region of the 2D ^{1}H-^{1}H NOESY spectra of α-pheromone in Gemini 171
5.8 Overlapping 1D NMR spectra of α-pheromone in H_2O 173
5.9 Selected regions of the 2D ^{1}H-^{1}H TOCSY spectra of α-pheromone in H_2O 174
5.10 Selected regions of the ^{1}H-^{13}C HSQC spectra of reduced and oxidized α-pheromone 175
5.11 Differences in the side chains of α-pheromone between reduced and oxidized states in DPC. 176
5.12 Proton chemical shift differences of the backbone HN and Hα and selected side chains 177
5.13 Mass spectrum of the 84 days old α-pheromone in H_2O 178
5.14 Ensemble of the 20 structures with the lowest target function values of α-pheromone, scrambled, D-Ala^{1,2} and D-Ala^{6,7} in H_2O. 184
5.15 Ensemble of the 20 structures with the lowest target function values of α-pheromone, scrambled, D-Ala^{1,2} and D-Ala^{6,7} in 30 % TFE 185
5.16 Ensemble of the 20 structures with the lowest target function values of α-pheromone and scrambled in DPC 186
5.17 Ensemble of the 20 structures with the lowest target function values of α-pheromone and scrambled in Gemini 187
5.18 Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of α-pheromone in 20 mM DPC/DPC-d38, 35 ºC, pH 5.0 189
5.19 Barplot representing the scaled intermolecular NOEs for each residue . . . 190
5.20 Model for the interaction between wt α-pheromone and the scrambled peptide with membranes. 195

List of tables

2.1 Sequences of the peptides studied 43
2.2 NMR spectra conditions of the peptides studied 50
2.3 Percentages of global helical populations 57
2.4 Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated in TFE and DPC 61

3.1 Sequences of the peptides studied 92
3.2 NMR conditions 93
3.3 Percentages of global helical populations 97
3.4 Main structural statistical parameters for the conformers calculated in HFIP and DPC 101

4.1 AMP Classification 121
4.2 Sequences of the peptides studied 130
4.3 NMR conditions 132
4.4 Main structural statistical parameters for the conformers calculated in H_2O and DPC 142
4.5 Intermolecular NOEs 143

5.1 Main toxins found in *Fusarium* spp. 157
5.2 Sequences of the peptides studied 164
5.3 NMR spectra conditions of the peptides studied 166
5.4 DLS conditions 178
5.5 Main structural statistical parameters for the conformers calculated in H_2O and TFE 182

5.6 Main structural statistical parameters for the conformers calculated in DPC and Gemini 183

A.1 Chemical shift list of TACC-α2 in H_2O, 5 ºC, pH 5.5 240
A.2 Chemical shift list of TACC-α2 in TFE, 25 ºC, pH 5.5. 241
A.3 Chemical shift list of TACC-α2 in DPC, 25 ºC, pH 5.5 242
A.4 Chemical shift list of Ct-TACC-α4 wt in H_2O, 5 ºC, pH 5.5 243
A.5 Chemical shift list of Ct-TACC-α4 wt in TFE, 25 ºC, pH 5.5 244
A.6 Chemical shift list of Ct-TACC-α4 wt in DPC, 25 ºC, pH 5.5 245
A.7 Chemical shift list of Ct-TACC-α4 mut in H_2O, 5 ºC, pH 5.5 246
A.8 Chemical shift list of Ct-TACC-α4 mut in TFE, pH 5 ºC, pH 5.5 247
A.9 Chemical shift list of Ct-TACC-α4 mut in DPC, 40 ºC, pH 5.5 248
A.10 Chemical shift list of XMAP-pCt in H_2O, 5 ºC, pH 5.5 249
A.11 Chemical shift list of XMAP-pCt in TFE, 25 ºC, pH 5.5 250
A.12 Chemical shift list of XMAP-pCt in DPC, 25 ºC, pH 5.5 251
A.13 Chemical shift list for peptides in mixture 252

B.1 Chemical shift list of MPERb in HFIP, 35 ºC, pH 7.0 256
B.2 Chemical shift list of MPERb in DPC, 35 ºC, pH 7.0 257
B.3 Chemical shift list of MPER-H2 in HFIP, 35 ºC, pH 7.0 258
B.4 Chemical shift list of MPER-H2 in DPC, 35 ºC, pH 7.0 259

C.1 Chemical shift list of PG-T4 in H_2O, 5 ºC, pH 3.5 262
C.2 Chemical shift list of PG-T4 in DPC, 35 º, pH 3.5C 263
C.3 Chemical shift list of PG-C6C15 in H_2O, 5 ºC, pH 3.5 264
C.4 Chemical shift list of PG-C6C15 in DPC, 35 ºC, pH 3.5 265
C.5 Chemical shift list of PG-C8C13 in H_2O, 5 ºC, pH 3.5 266
C.6 Chemical shift list of PG-C8C13 in DPC, 35 ºC, pH 3.5 267

D.1 Chemical shift list of Gln-Pro *trans* α-pheromone in H_2O, 5 ºC, pH 5.0 . . 270
D.2 Chemical shift list of Gln-Pro *cis* α-pheromone in H_2O, 5 ºC, pH 5.0 270
D.3 Chemical shift list of Gln-Pro *trans* α-pheromone in TFE, 5 ºC, pH 5.0 . . . 271
D.4 Chemical shift list of Gln-Pro *cis* α-pheromone in TFE, 5 ºC, pH 5.0 271

D.5 Chemical shift list of reduced Gln-Pro *trans* α-pheromone in DPC, 25 ºC, pH 5.0 272
D.6 Chemical shift list of oxidized Gln-Pro *trans* α-pheromone in DPC, 25 ºC, pH 5.0 272
D.7 Chemical shift list of reduced Gln-Pro *trans* α-pheromone in Gemini, 25 ºC, pH 5.0 273
D.8 Chemical shift list of oxidized Gln-Pro *trans* Gln-Pro *trans* α-pheromone in Gemini, 35 ºC, pH 5.0 273
D.9 Chemical shift list of Trp-Pro *trans* scrambled in H_2O, 5 ºC, pH 5.0 274
D.10 Chemical shift list of Trp-Pro *cis* scrambled in H_2O, 5 ºC, pH 5.0 274
D.11 Chemical shift list of Trp-Pro *trans* scrambled in TFE, 5 ºC, pH 5.0 275
D.12 Chemical shift list of Trp-Pro *cis* scrambled in TFE, 5 ºC, pH 5.0 275
D.13 Chemical shift list of Trp-Pro *trans* scrambled in DPC, 25 ºC, pH 5.0 276
D.14 Chemical shift list of Trp-Pro *cis* scrambled in Gemini, 25 ºC, pH 5.0 . . . 277
D.15 Chemical shift list of Gln-Pro *trans* D-Ala^{1,2} in H_2O, 5 ºC, pH 5.0 278
D.16 Chemical shift list of Gln-Pro *cis* D-Ala^{1,2} in H_2O, 5 ºC, pH 5.0 278
D.17 Chemical shift list of Gln-Pro *trans* D-Ala^{1,2} in TFE, 5 ºC, pH 5.0 279
D.18 Chemical shift list of Gln-Pro *cis* D-Ala^{1,2} in TFE, 5 ºC, pH 5.0 279
D.19 Chemical shift list of dAla-Pro *trans* D-Ala^{6,7} in H_2O, 5 ºC, pH 5.0 280
D.20 Chemical shift list of dAla-Pro *cis* D-Ala^{6,7} in H_2O, 5 ºC, pH 5.0 280
D.21 Chemical shift list of dAla-Pro *trans* D-Ala^{6,7} in TFE, 5 ºC, pH 5.0 281
D.22 Chemical shift list of dAla-Pro *cis* D-Ala^{6,7} in TFE, 5 ºC, pH 5.0 281

Nomenclature

Symbols

Å	Ångström
Da	Dalton
J	Scalar coupling
K_d	Dissociation constant
m/z	Mass-to-charge ratio
δ	Chemical shift
θ	Ellipticity
λ	Wavelength

Acronyms / Abbreviations

6-HB	Six-helix bundle
AIDS	Acquired immunodeficiency syndrome
AMP	Antimicrobial peptides
BMRB	Biological magnetic resonance bank
CCR5	C-C Chemokine receptor type 5
CD	Circular dichroism
CD4	Cluster of differentiation 4
CHR	C-terminus helix region
chTOG	Colonic, hepatic tumor over-expressed gene
CMC	Critical micelle concentration

COSY	Correlation spectroscopy
CSP	Chemical shift perturbations
CT	Cytoplasmic tail
DLS	Dynamic light scattering
DPC	Dodecylphosphocholine
DSC	Differential scanning calorimetry
DSS	2,2-dimethyl-2-silapentane-5-sulfonate sodium salt
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
Env	Envelope complex
Fab	Fragment antigen-binding
Fmoc	Fluorenylmethyloxycarbonyl
FP	Fusion peptide
gp120	Glycoprotein 120
gp41	Glycoprotein 41
GPCR	G protein-coupled receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFIP	Hexafluoroisopropanol
HIV	Human immunodeficiency virus
HNP-3	Defensin human neutrophil peptide-3
HPLC	High-performance liquid chromatography
HR1/2	Heptadrepeat regions 1 or 2
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence spectroscopy
ID	Immunodominant region
IDPs	Intrinsically disordered proteins

IR	Infrared
ITC	Isothermal titration calorimetry
MALDI	Matrix-assisted laser desorption/ionization
MAP	Microtubule-associated protein
MAPK	Mitogen-activated protein kinase
MPER	Membrane proximal external region
Mpk1	Mitogen-activated protein kinase 1
MRSA	Meticillin-resistant *Staphylococcus aureus*
MS	Mass spectrometry
MT	Microtubule
MTOC	Main microtubule organizing center
NHR	N-terminus helix region
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
PC	Phosphatidylcholine
PDB	Protein data bank
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PG	Protegrin
PI	Phosphatidylinositol
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
ppm	Parts per million
PR	Polar region

PS	Phosphatidylserine
RC	Random coil
RDC	Residual dipolar coupling
RMSD	Root mean square deviation
RP-HPLC	Reversed-phase high-performance liquid chromatography
rpm	Revolutions per minute
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulfate
ssNMR	Solid state NMR
TACC	Transforming acidic coiled-coil
TD4	TACC domain 4
TFA	Trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TMD	Transmembrane domain
TMS	Tetramethylsilane
TOCSY	Total correlation spectroscopy
TOF	Time of flight
tR	Retention time
XMAP	*Xenopus* microtubule-associated protein
XTACC	*Xenopus* transforming acidic coiled-coil

Resumen

Diseño y Caracterización Estructural de Sistemas Minimizados para el Estudio de Interacciones Biomoleculares Mediante RMN

Los procesos moleculares que tienen lugar en los organismos vivos se regulan a través de una enmarañada red de grandes complejos biológicos que interaccionan entre sí. En este contexto, los mecanismos de reconocimiento biomolecular que se producen en cada uno de ellos son esenciales para dar lugar al complejo activo y biológicamente funcional. Por su importancia, uno de los retos actuales en Biología es comprender a fondo los detalles específicos que definen estos procesos. Por ejemplo, necesitamos conocer cómo y cuándo ocurren, cuáles son los factores esenciales en cada tipo de reconocimiento, qué fuerzas están en juego, cómo se equilibran y dirigen estos mecanismos, etc. Una de las estrategias más apropiadas para esclarecer estas incógnitas es la caracterización de estos sistemas a escala atómica. Sin embargo, el estudio de estas asociaciones biomoleculares que forman grandes redes de interacciones, a este nivel de detalle, es extremadamente complicado. En este escenario, una aproximación realista para el estudio de un sistema biológico complejo consiste en llevar a cabo la caracterización experimental de versiones simplificadas del mismo en sistemas modelo.

Así pues, el objetivo planteado en esta Tesis ha sido la descripción de sistemas biológicos completos a partir de la información estructural y funcional de modelos minimizados diseñados al respecto. Para este fin se ha utilizado fundamentalmente la espectroscopía de Resonancia Magnética Nuclear (RMN), así como otras técnicas de caracterización biofísicas y computacionales. Tal como describimos más adelante, hemos aplicado este enfoque para el estudio de cuatro sistemas concretos. En estos casos, las interacciones péptido-péptido y péptido-micela de los modelos simplificados se han utilizado para describir los mecanismos de reconocimiento de sus correspondientes sistemas proteína-proteína y proteína-membrana. Nuestros resultados confirman que la espectroscopía de RMN es una herramienta excelente para describir a escala atómica las interacciones moleculares débiles y en ocasiones transito-

rias que se dan en estos complejos, las cuales son fisiológicamente las más relevantes para la vida.

Debido a su intrínseco interés biológico, nos hemos centrado en la identificación de las bases estructurales de: i) el reconocimiento de proteínas relacionadas con el ciclo celular, en concreto, TACC3 y XMAP, ii) la glicoproteína gp41 presente en la envuelta del virus VIH, la cual participa en los procesos iniciales de infección celular, iii) algunas variantes del péptido antimicrobiano protegrina-1, y iv) la α-feromona del hongo patógeno *Fusarium oxysporum*, que es un péptido corto implicado directamente en procesos de infección y propagación. En todos estos casos, y en base a los datos experimentales obtenidos en esta Tesis, hemos propuesto los correspondientes modelos moleculares que dan cuenta de las interacciones proteína-proteína o proteína-membrana objeto de estudio.

Con respecto al sistema TACC3 y XMAP, concluimos que para que la interacción suceda, las estructuras de los péptidos modelo de TACC3 y XMAP deben estar preformadas, y que la interacción ocurre sin necesidad de alcanzar la proporción molecular de la forma nativa. Este es un nuevo modelo de interacción que podría describir las etapas iniciales en los eventos de reconocimiento que ocurren durante la formación del complejo activo de TACC3-XMAP.

En cuanto a los péptidos derivados de la glicoproteína gp41, los resultados descritos en esta Tesis ofrecen información estructural y dinámica concerniente a los mecanismos de fusión entre las membranas del virus del VIH y la célula. La información recogida de este análisis puede resultar de interés para el diseño de nuevos péptidos inmunogénicos como dianas en el desarrollo de vacunas.

En el caso de los modelos derivados del péptido antimicrobiano protegrina-1, los datos estructurales brindan información sobre el mecanismo de interacción de la protegrina-1 con la membrana biológica. Este conocimiento puede ser relevante para el diseño racional de nuevos fármacos que sean potencialmente capaces de solventar los problemas debidos a la resistencia a antibióticos de determinados microrganismos.

Finalmente, los resultados descritos en esta Tesis sobre la α-feromona del hongo patógeno *F. oxysporum* son los únicos ensayos experimentales realizados hasta la fecha sobre este péptido. Los datos obtenidos han permitido elaborar un modelo de la posible interacción entre la α-feromona con la bicapa lipídica, y proponer además las distintas funciones biológicas de la molécula en su forma oxidada o reducida. Esta información puede ser de utilidad en el futuro para el control de plagas y el tratamiento médico de infecciones en personas y animales.

Abstract

Design and Structural Characterization of Minimized Systems to Study Biomolecular Interactions Through NMR

The molecular processes that take place in living organisms are regulated by an intricate network of large complexes interacting among them. In this context, the biomolecular recognition processes involving each component are essential to produce the fully biologically active complex. Understanding the fine details on how these processes occur, which are the important players in each step of the recognition, which forces are involved, how they are balanced and drive the process, etc., is the current challenge in Biology. One approximation to solve these open questions is the characterization of these systems at atomic resolution. However, the study of the native biomolecular associations in the form of multiple interacting networks, at this resolution level, is extremely difficult. Given these premises, the best approach is to perform an experimental analysis of the biological systems of interest by using simplified versions with the help of model systems.

In this regard, this Thesis is focused on the applications of Nuclear Magnetic Resonance (NMR) spectroscopy and other biophysical and computational methods to obtain structural and functional information of designed models, as a first step towards the complete understanding of the full and active biological systems. Here, we illustrate this approach by applying it to four particular systems of biological relevance (see below). Peptide-peptide and peptide-micelle recognition are examined as simplified models of their corresponding protein-protein and protein-membrane interactions. Our results corroborate that NMR spectroscopy is an excellent tool to describe, at atomic resolution, the weak, and sometimes transient, molecular interactions, which are physiologically the most relevant in life.

For their intrinsic biological interest, we aim to get insights into the structural features responsible for: i) the recognition of proteins related with the cell cycle, in particular, TACC3 and XMAP, ii) the gp41 glycoprotein present in the envelope of HIV virus, which participates in the initial events of cell infection, iii) some variants of the antimicrobial peptide protegrin-

1, and iv) the α-mating pheromone from the pathogenic fungus *Fusarium oxysporum*, which is a short peptide directly implicated in infective and dissemination processes. In all cases, molecular models of the protein-protein or protein-membrane interactions have been proposed on the basis of the data obtained in this Thesis.

In the case of the TACC3 and XMAP system, we conclude that both preformed structures of TACC3 and XMAP peptides should be present simultaneously to trigger the interaction, but not necessarily in the native molecular ratio. This new model of interaction could describe the initial steps of the recognition event in the formation of the fully active XTACC3-XMAP215 complex.

Concerning the peptides derived from the gp41 glycoprotein, the results described in this Thesis bring structural and dynamic information about the mechanism of membrane fusion between the HIV virus and the cell, which takes place at the first stages of the infective process. Interestingly, the information gathered can be of interest towards the design of new immunogenic peptides as targets for vaccine development.

Regarding the peptides derived from the antimicrobial protegrin-1, the obtained structural data provide insights into the way of interaction of these antimicrobial peptides (protegrin-1) with the cell membrane. This knowledge is of relevance towards a rational design of new drugs able to overcome the current antibiotic resistance scenario.

Finally, the results described in this Thesis for the α-mating pheromone from the pathogenic fungus *F. oxysporum* are the first and only existent structural experimental data for this peptide. These data allow us to build a model of the possible interaction of the α-pheromone with the lipid bilayer and propose different biological roles of the oxidized and reduced forms of the molecule. We suggest that this knowledge can be applied in the future in disease control and further medical treatments for infected humans and animals.

1

General introduction and objectives

1.1 Background

The molecular biology of the cell is an intricate network in which different kinds of molecules establish contacts that may vary from weak to strong; from transient, to permanent; from close to long distance, and so on. In this context, the interactions that rule these contacts are essential to maintain the biochemical pathways that regulate the organization of the cells, tissues, organs, systems, and consequently, life itself.

In living organisms, the cellular membrane is mainly composed by a bilayer of phospholipid molecules. This membrane acts as a compartment that allows the cell to dispose of various environments and perform diverse biochemical processes. These individual cells can be capable to connect or communicate with other cells employing different molecules. In this way, large polymeric systems are usually found to act as scaffolds of tissues and cells, such as cytoskeletal molecules (microfilaments, microtubules and intermediate filaments) or the basal lamina. Between these polymeric systems, structural proteins are important elements that can be associated with other molecules to achieve different degrees of organization, flexibility and function.

In the cell, each cellular compartment contains a large variety of different molecules that have to unambiguously recognize their partners in order to perform a specific biological action. These molecular recognition processes consist of specific interactions between the involved molecules and occur fundamentally throughout non-covalent forces. These interactions have to be regulated, usually by other molecules, to ensure the correct cell signaling, enzyme catalysis, ligand-receptor binding, regulation routes, cell transport, cell cycle arrest, among others. When the homeostasis of the system in which these interactions are taking place is deregulated, different health issues may appear.

1.2 Structural and functional insights in biomolecular recognition

Structural studies of biomolecules offer information about the spatial location of each atom or regions in the space and sometimes these data can be analyzed in a time dependent matter to obtain additional information. Regarding polypeptides, there is an established consensus that defines four main levels of protein organization: primary, secondary, tertiary and quaternary structure. The primary structure refers strictly to the sequence of amino acids in the polypeptide chain and contains all the information required for its unique three-

dimensional final structure, while the secondary structure denotes to the 3D disposition of local arrangements of a polypeptide chain such as an α-helix or a β-strand. The tertiary structure refers to the whole 3D structure of a single polypeptide chain and usually contains several secondary structures. Finally, the quaternary structure regards to the organization of multiple polypeptide chains composed by two or more monomers.

It has been widely proposed that polypeptides that have elements of secondary structure, but do not have a globular 3D structure, such as keratins, are usually found as structural proteins and act as scaffolds for the cells and organs. On the other way, proteins with globular conformations are usually found to catalyze biochemical reactions. Concerning these catalytic proteins or enzymes, the German chemist H. E. Fisher proposed in 1894 a mechanistic model of molecular recognition known as the "lock and key model" and founded a dogma based on the idea in which a substrate accommodates specifically into a determined enzyme (like a key into its corresponding lock), and only one precise and rigid conformation is able to fit into the enzyme's active site to trigger the reaction [1]. However, five decades later, Koshland suggested that the interaction between the ligand and the receptor induces a conformational change in the enzyme, acting as an "induced fit"; therefore, the binding process is the motor that forces the changes observed in the complex [2]. Though, both models consider that proteins hold single and stable conformations under specific conditions. Later on, as the dynamic and flexible nature of the complexes has gained importance, additional models have been proposed. Currently, it is well known that proteins can adopt several different conformations and besides its native state, proteins can be capable to exist in other conformational states [3, 4].

In this regard, the conformational selection model proposes that weakly populated states are capable to recognize and bind different molecules. As suggested by the energy landscape theory of protein dynamics proposed in 1991 [3], proteins can assume several diverse and nearly isoenergetic conformations (conformational sub-states), and the binding of a certain partner changes the free energy landscape of the protein towards one of the possible conformations.

For example, it has been found by nuclear magnetic resonance (NMR) that each ensemble obtained from the ligand-free ubiquitin corresponds with a different bound conformation (**Figure 1.1**). This means that in solution and in the absence of any of its possible ligands, the protein adopts a variety of conformations, and when bound to a determined molecule, one specific conformation is selected from the overall set of conformers present in the free state [5, 6].

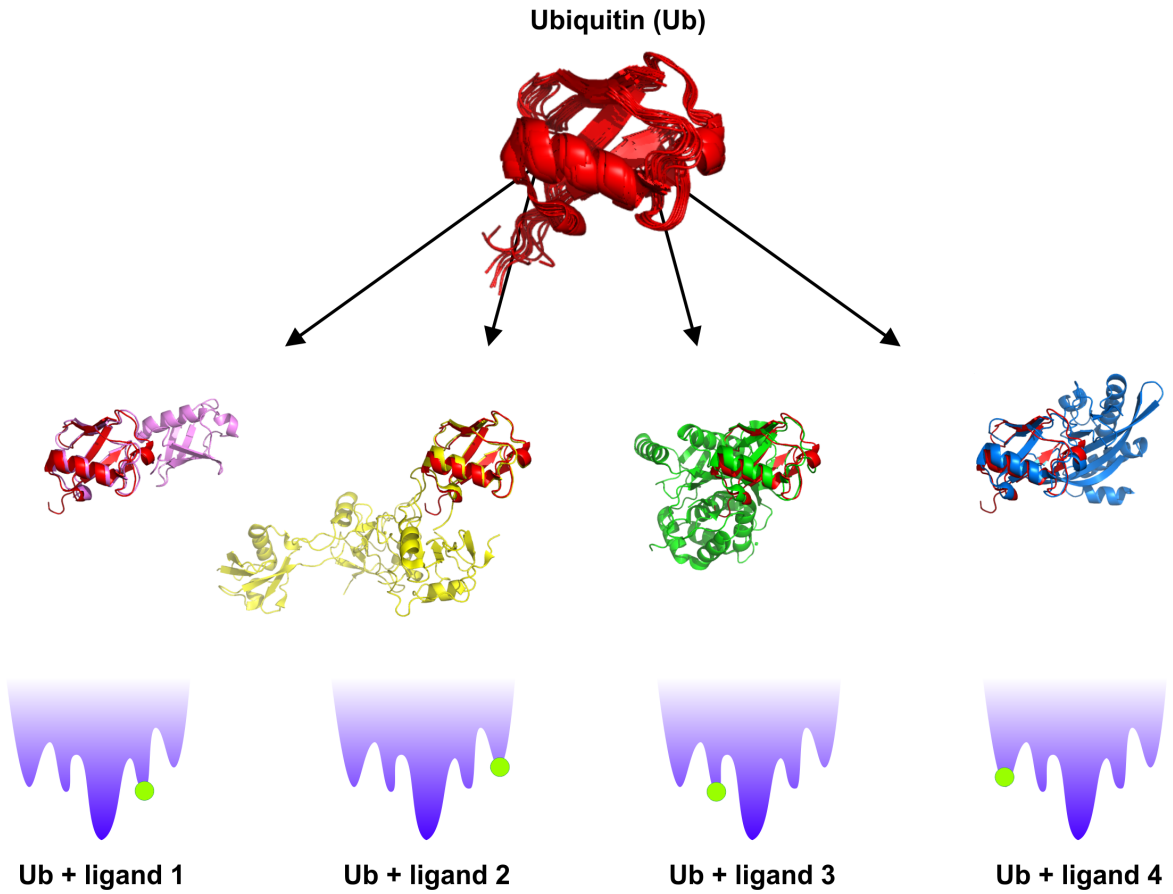


Figure 1.1. Ubiquitin as an example of the conformational selection model. Within the protein folding thermodynamic funnel, several conformations are possible for the free protein (in red). Once the ligand (pink, yellow, green and blue) binds to one of the possible conformations over the overall conformers within the thermodynamic funnel (green dots), no additional conformational possibilities can be selected.

Following this line, intrinsically disordered proteins and regions (IDPs and IDRs, respectively) are proteins or portions of a protein that lack a defined 3D structure. These proteins are highly flexible and despite the absence of ordered structure, they are still functional and can become structured upon specific interactions with other macromolecules. This challenges the classical structure-function paradigm and makes their study interesting as they have been found to be in a high proportion in superior organisms to function as signal pathways regulators, to work as adaptors and to regulate post-translational modifications [7].

1.2.1 Biological interactions and molecular recognition

As mentioned above, molecular recognitions are mainly due to weak non-covalent interactions. These interactions are considered as "weak" forces that determine the properties and diverse functions of biomolecules. The energy required to break this kind of interactions is lower than in covalent bonds (around 1 - 5 kcal/mol); therefore, they can be constantly formed and broken at physiologic conditions.

The classification of non-covalent interactions can be diffuse and differs between authors, however, in this Introduction chapter we are going to focus on the four main widely-accepted types of non-covalent interactions: ionic bonds or interactions, hydrogen bonds, van der Waals interactions and the hydrophobic effect.

1.2.1.1 Ionic interactions

In this type of interaction, a positively charged ion (cation) and a negatively charged one (anion) are mutually attracted; as well, repulsive interactions can also take place between ions with the same charge. The strength of this interaction is determined by the distance between the charges and the concentration of other ions in solution. In consequence, an increase in the salt concentration in a biological sample could weaken the ionic forces that are maintaining the structural properties of the biomolecule.

In polypeptides, this kind of interactions can take place between a positively charged side chain (Lys, His, Arg) and a negatively charged one (Asp, Glu, Tyr, Cys). Since these side chains get protonated at certain pH values, these interactions are pH dependent. In the same way, the C- (anionic) and N-terminus (cationic) ends of the polypeptide can participate in ionic interactions depending on their charge.

1.2.1.2 Hydrogen bonds

Hydrogen bonds are formed when a partially positively charged hydrogen atom from a dipole interacts with unpaired electrons of another atom. The hydrogen atom may acquire a partially positive charge when attached to a highly electronegative atom. This kind of interaction does not involve full charges and its strength is generally less than charge-charge interactions.

Hydrogen bonds determine the water solubility of the biological molecules in an aqueous solution as it depends largely on their capabilities to interact with the dipoles of the water molecules. In the same way, they can also be found stabilizing secondary structures in

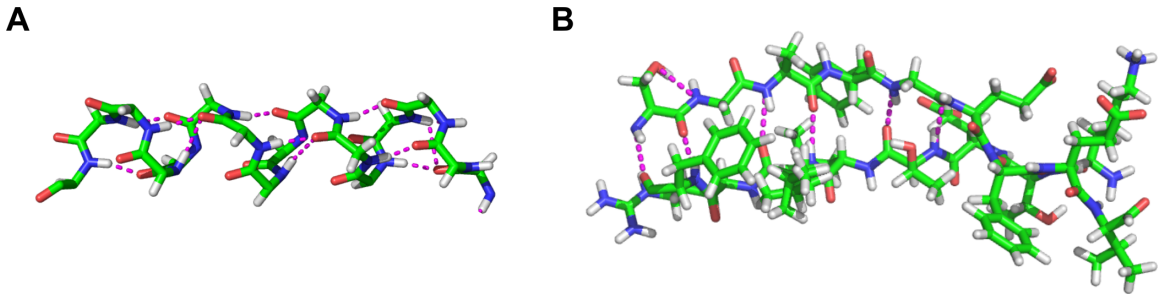


Figure 1.2. Secondary structure stabilization through hydrogen bonds. Hydrogen bonds (magenta dashes) between the HN and the O atoms of the backbone stabilize the α-helix (A) and β-sheet (B) secondary structures.

polypeptide chains, such as α-helices and β-sheets (**Figure 1.2**). Indeed, one of the characteristics used to classify the type of secondary structure are the residues participating in backbone hydrogen bonds. Also, a variety of chemical groups (amine, carboxyl, hydroxyl, amide, guanidinium) from the backbone or the side chains can interact with other residues through this type of chemical bond.

1.2.1.3 Van der Waals interactions

This is a weak attraction due to transitory dipoles in the electron clouds of nearby atoms (1 kcal/mol) and non-specific force created by the close approaching of permanent or transient dipoles that can be attractive or repulsive. They are distance dependent and non-directional forces that occur in all types of molecules, either polar or non-polar. Even if this is a short-ranged interaction, when the atoms become too close, their negatively charged electrons create a repulsive force between them and the balance created between the repulsive and the attractive forces results in a van der Waals contact.

1.2.1.4 The hydrophobic effect

The hydrophobic effect is especially relevant in biomolecules, which acquire their native structure within an aqueous physiological environment. During folding, the hydrophobic regions of proteins are usually oriented to the protein core in order to avoid their non-favorable contacts with the polar solvent. However, this is not always the case and sometimes, large non-polar regions remain oriented towards the exterior. These hydrophobic patches lack a dipole moment that is essential to interact with water molecules and therefore, be solubilized. In these cases, the hydrophobic effect can promote recognition processes but may spur harmful aggregation.

In the case of biological membranes, the hydrophobic chains of the lipids associate between them, while their polar heads are in contact with water molecules. Concerning transmembrane proteins, the hydrophobic regions can be in contact with the non-polar chains of the lipids that compose the bilayer and therefore, be integrated as transmembrane proteins.

1.3 The biological membrane

Biological membranes share the same general architecture across the different taxonomic domains (Archaea, Bacteria and Eukarya), which consist of a phospholipid bilayer with embedded proteins, sterols and carbohydrates. Biomembranes are flexible and thin surfaces that create sealed and separated compartments, in which the flow of molecules and ions between the inner and outer faces is controlled, including the water molecule transport. In 1972, Singer and Nicolson proposed a model for the structure of biomembranes known as the fluid mosaic model (**Figure 1.3**), they defined the membrane as a two-dimensional fluid and dynamic mosaic of proteins embedded within two layers of lipids [8]. Although this model is still valid, a more complete characterization of the membrane properties emerging from recent studies indicates the presence of defined domains known as lipid rafts, differences in the lipid and protein mobility, and the presence of molecular associations and localized functions therein [9, 10].

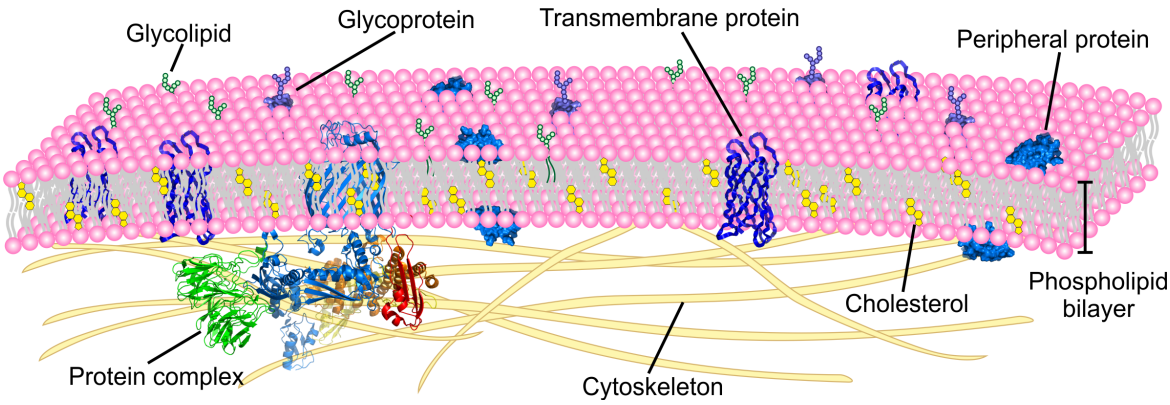


Figure 1.3. Fluid mosaic membrane model. The membrane consists of a lipid bilayer with embedded additional constituents. The main components of the biological membrane are phospholipids, glycolipids, sterols, transmembrane proteins, peripheral proteins and membrane associated proteins. Depending on the face, the membrane contains different proportion of lipids, such as phosphatidylcholine (PC), phosphatidylserine (PS), phosphatidylethanolamine (PE) and phosphatidylinositol (PI).

1.3.1 Properties and composition of the biological membrane

Membrane composition varies in the type and relative abundance of its components. This differentiated ratio contributes to the specialization of certain types of cells even within the same organism.

1.3.1.1 Lipids

Lipids are the main component of the biological membrane; their non-polar nature provides to the membrane its impermeable properties. Alongside their non-polar properties, they can still be dissolved with organic solvents such as chloroform, benzene or acetone. Lipids can be classified as well into three primary types: phospholipids, glycolipids and sterols (**Figure 1.4**), being the phospholipids the most abundant ones.

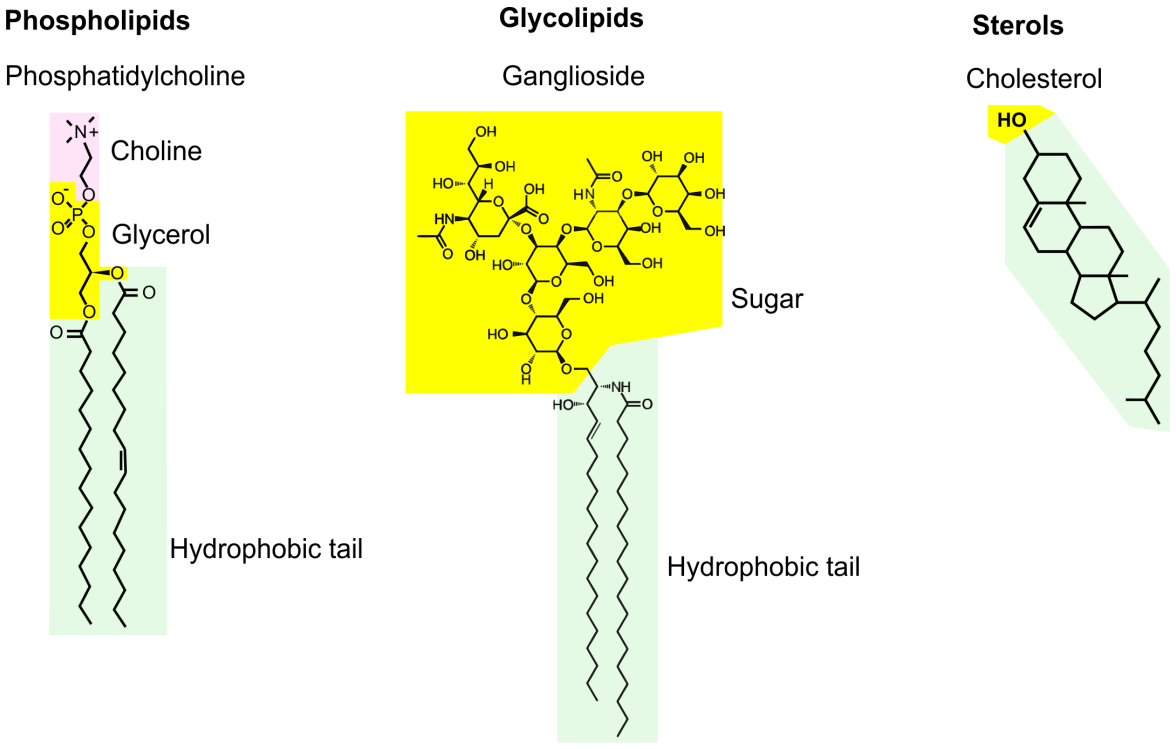


Figure 1.4. Main classification of the lipid components in the membrane. Three primary types of lipids can be classified as phospholipids, glycolipids and sterols, and are composed by a polar region (yellow and pink) and a hydrophobic tail (green). Sterols are mainly fully hydrophobic and present a hydroxyl group.

Phospholipids. They consist of amphipathic molecules that can form bilayers spontaneously. Their amphipathic nature makes them the key components of the biological membranes, and when isolated in an aqueous solution, they are capable of aggregating into three primary different structures: micelles, liposomes and bilayers (**Figure 1.4**).

Phospholipids can also be subclassified as phosphoglycerides and sphingolipids (**Figure 1.5**). Phosphoglycerides are the most abundant lipids found in membranes and are composed by two long hydrophobic fatty acid tails (hydrophobic tail) esterified in the two hydroxyl groups in glycerol phosphate, with this portion bound to a variable polar region (polar head). The most abundant phosphoglycerides are phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS) and phosphatidylinositol (PI).

Glycolipids. These are amphipathic molecules with attached sugar groups to the lipid structure and constitute around 2 to 10 percent of the total content of the plasma membrane.

Sterols: cholesterol and analogs. This is the third most important class of membrane lipids. Sterols are composed by a four-ring isoprenoid-based hydrocarbon (**Figure 1.4**). Given their extremely hydrophobic nature, sterols are found intercalated between phospholipid molecules and are not capable to form bilayers spontaneously. When they are incorporated into the membranes, sterols offer structural support and provides a more gel-like consistency to the bilayer.

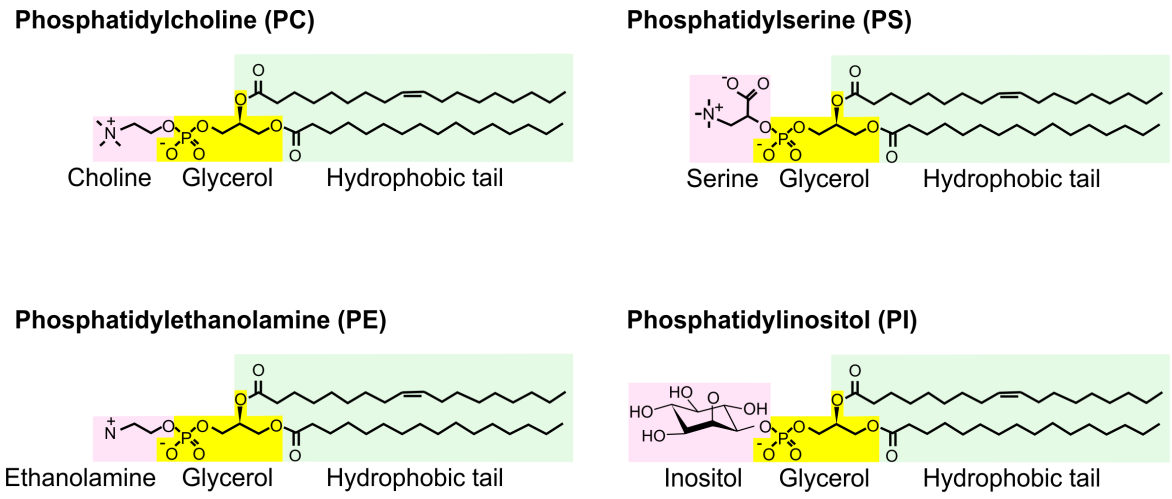


Figure 1.5. Most abundant phosphoglycerides. Phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine and phosphatidylinositol are composed by a variable polar head (pink), a glycerol group (yellow) and a hydrophobic tail (green). The relative abundance of a defined lipid varies on the species and the cell type.

1.3.1.2 Membrane proteins

Proteins are the second major component of cell membranes and are the main generators of biological reactions [11]. According to their degree of association within or around the biological membrane, three classes can be defined: integral, lipid-anchored and peripheral proteins (**Figure 1.6**).

Integral membrane proteins. Integral or transmembrane proteins span the phospholipid bilayer and present three main domains: the region that faces to the cytoplasm, the transmembrane domain and the exoplasmic domain. (**Figure 1.6 A**).

Membrane-anchored proteins. These proteins are covalently bound to one or more lipids inserted into the membrane. (**Figure 1.6 B**).

Peripheral membrane proteins. This kind of proteins usually perform enzymatic reactions and their binding with the membrane is indirect. (**Figure 1.6 C**).

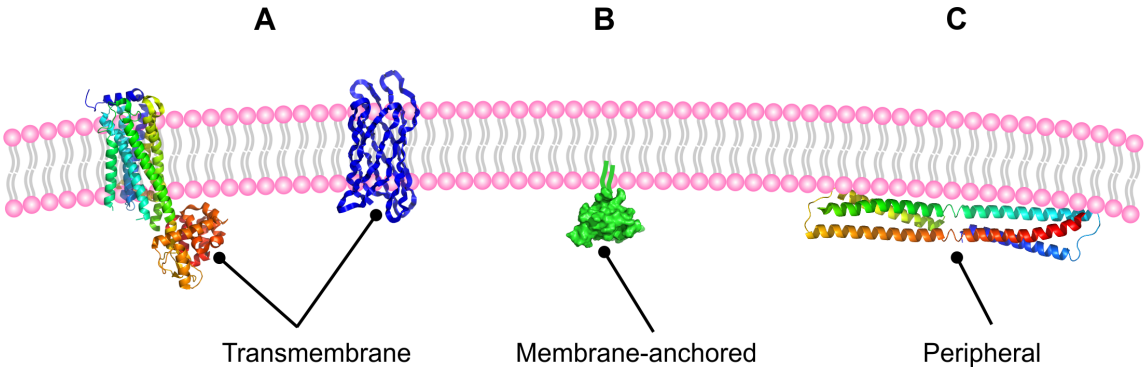


Figure 1.6. Different types of membrane proteins. Transmembrane proteins (A) span the full bilayer, while membrane-anchored proteins (B) are proteins with a lipid portion that face to one side of the membrane. Peripheral proteins (C) adhere temporally to the membrane.

1.3.1.3 Carbohydrates

Finally, carbohydrates are part of the membrane and can be found associated in lipids (glycolipids) or proteins (glycoproteins). Carbohydrates usually represent no more than the 10 % of the membrane total weight and are more usually found as glycoproteins rather than as glycolipids. In the same way, there is an asymmetry regarding the location of the sugar groups, being more abundant in the external portion of the plasma membrane and making

them capable of interacting with components of the extracellular matrix, lectins, growth factors and antibodies.

1.4 Protein-protein and protein-peptide interactions (PPIs)

Although most of the interactions in nature involve proteins, there are also short peptides found in the cells that are prone to interact with other proteins and biological membranes (see below) [12], also there is a growing recognition taht many peptides are expressed and are putatively functional in cells. Thus, protein-protein or protein-peptide interactions are key to understand the biological functions that occur in the organisms and regulate their vital functions. As mentioned before, formation of protein complexes is ruled by non-covalent and reversible contacts and the forces involved in recognition and association are of the same nature that those responsible for stabilization of protein tertiary structure [6].

Protein complexes can composed by either identical (homo-oligomers) or different (hetero-oligomers) polypeptide units. Among the association possibilities, the monomers can be assembled through identical regions and create symmetrical motifs or be associated in a non-symmetrical way (**Figure 1.7**) [13–15].

One interesting question is that the contacts established between different polypeptide chains should be highly specific in order to trigger a determined biological function. In this regard, there can be two types of protein-protein interactions: obligate and non-obligate (**Figure 1.8**). An obligate interaction is one in which their individual constituents are unstable

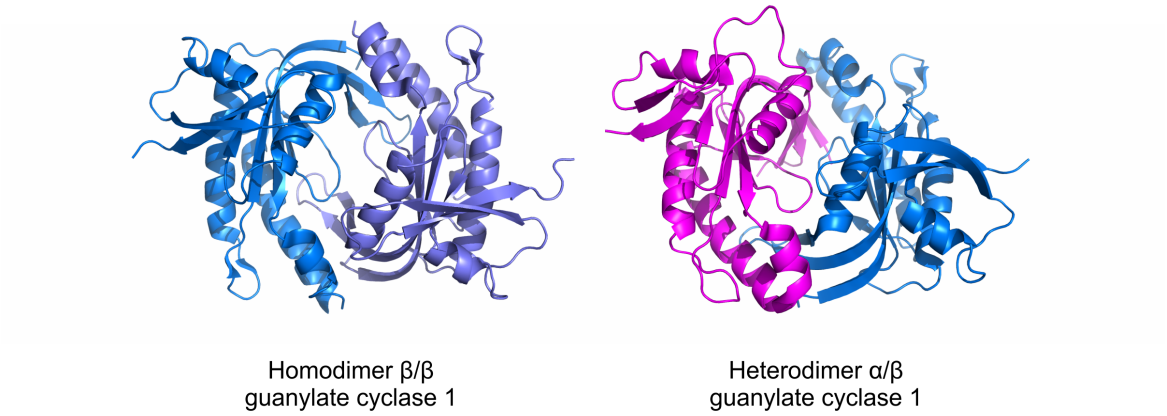


Figure 1.7. Protein-protein interactions between subunits. Dimeric interactions of the guanylate cyclase 1 between identical subunits (β, in light and dark blue; PDB code 2wz1) and association of different subunits (α and β, in magenta and light blue, respectively; PDB code 3uvj).

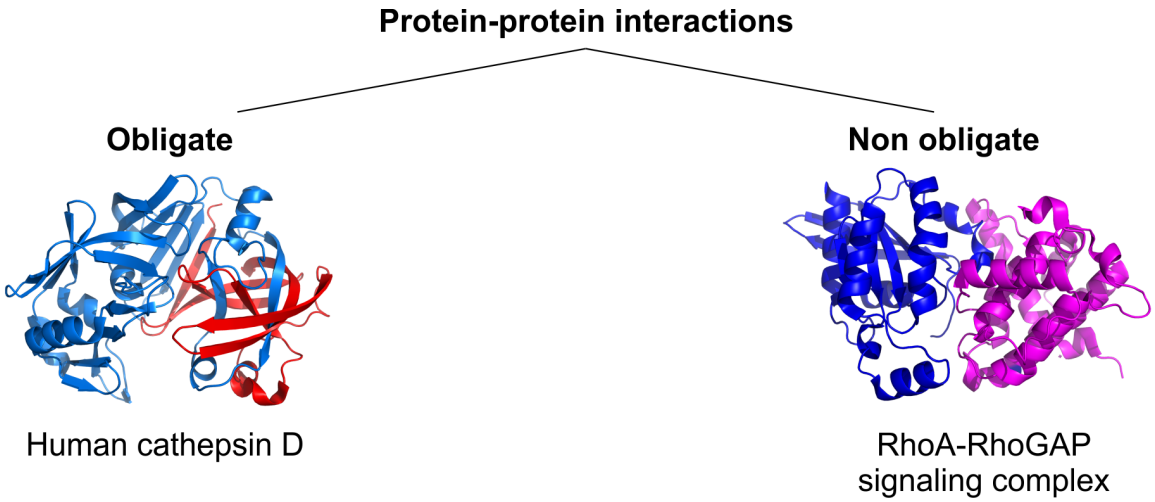


Figure 1.8. Obligate and non-obligate protein-protein interactions. In the obligate interaction of the human cathepsin D (PDB code 1lyb) the simultaneous presence of the units is necessary towards the stabilization of both constituents. In a non-obligate interaction (PDB code 5irc), both constituents are stable independently.

while not bound and are forced to form a complex in order to gain stability, while in a non-obligate association, the individual components are stable independently [14].

As mentioned before, biomolecular associations can also be described in a time-dependent manner. In this way, when the PPIs are strong and irreversible, the association is very stable and the complex is known as permanent. Transient complexes on the contrary, are reversible and can be further classified as weak or strong. Weak transient interactions show a variety of dynamic oligomeric states *in vivo*, while in strong transient complexes only another molecule (e.g. a ligand) is needed to trigger the change of the quaternary state [14, 16].

1.5 Protein and peptide interactions with membranes

Lipid-protein and lipid-peptide interactions are highly important during the regulation of biological processes. Their relevance for the system arrangement is evident, since one third of the genome of certain organisms, including humans and yeasts has been found to code for membrane proteins [11, 17]. Lipid-protein interactions are indispensable in some cellular mechanisms such as cell signaling and transport of molecules and ions into the cell. Concerning these tasks, it is known that several cytoplasmic and membrane proteins, as well as lipids and sterols, can be enrolled into the cell membrane and arrange localized complexes in form of lipid rafts [18].

In order to experimentally study this relevant kind of interactions, there are several concerns to take into account. i) The transmembrane regions of the proteins are rich in hydrophobic amino acids, this makes them unstable when exposed to water and other polar solvents and consequently, tend to aggregate. This give rises to difficulties during protein expression and purification. ii) The majority of protein-lipid interactions are not highly specific. iii) The highly dynamic nature of membranes contributes to practical limitations towards to obtain a complete map of these interactions. iv) The mechanisms of lipid recognition can involve multiple steps, such as membrane binding and insertion in which proteins undergo conformational changes, therefore, the characterization of these interactions requires very detailed structural and dynamic information.

Nowadays, there is still not a complete understanding about the structural nature of the interactions between proteins and membranes. In order to get new insights of these biomolecular interactions and to clarify the molecular bases that govern these processes, several techniques can be applied for their characterization.

1.6 Experimental methods to study biomolecular interactions

Currently, several methods can be applied to study these interactions (protein-protein and protein-membrane), each of which has advantages and limitations, especially in terms of the sensitivity and specificity. These procedures can be divided into two major groups: biochemical and biophysical techniques, and theoretical methods.

At the same time, intermolecular interactions of biomolecules can be described with two levels of resolution: atomically or globally. Likewise, they can be studied from a kinetic point of view through reaction rates, chemical equilibrium or by thermodynamic analysis. Thermodynamic analysis can be done by isothermal titration calorimetry (ITC) and differential scanning calorimetry (DSC), as well as for spectroscopic techniques, such as circular dichroism (CD) spectroscopy and fluorescence. These last two techniques, as well as IR and Raman spectroscopy, mass spectrometry and atomic force microscopy, can allow us to get structural information on protein interactions. But, undoubtedly, the techniques that provide the most information for the study and structural analysis of complexes with atomic resolution are electron microscopy (EM), X-ray crystallography and Nuclear Magnetic Resonance (NMR).

The intricacy of the study of complete systems is a barrier to overcome and sometimes it may be necessary to apply a reductionist approach. This means that occasionally, the complex needs to be simplified and its components have to be studied separately. At the same time, different methods should be used simultaneously in order to obtain complementary information.

In the following paragraphs the biochemical and biophysical techniques most widely used for the study of interactions and structural properties of biomolecules are described. The main techniques employed during this Thesis are among them.

1.6.1 Mass spectrometry (MS)

MS is an analytical and destructive spectroscopic technique that ionizes chemical species and ranks them based on their mass/charge (M/Z) ratio. The ions within the sample are separated depending on their charge and size and once detected, a plot is generated in which different isotopic forms can be visualized. MS is usually employed at the final stages of protein and peptide purification to determine the purity of the sample. In our studies, MS has also been used to determine the presence of reduced or oxidized states of Cys residues based on the proton contents [19] (**Figure 1.9**). Given that the main components are separated, it can also be used to determine the presence of other components and, therefore, the presence of protein interactions [20, 21].

1.6.2 Dynamic Light Scattering (DLS)

DLS is a non-destructive spectroscopic technique employed to determine the distribution and diffusion behavior of suspended macromolecules based on their size. DLS measures the Brownian motion of macromolecules, relates it with the size and shape of the particles and their radii can be inferred and plotted as a histogram. The sample is subjected to a monochromatic beam of light; this disperses or scatters the light in different directions depending on the size and shape of the particles, as well as, on the temperature and solvent viscosity. [22].

During DLS analysis, different kind of results can be obtained. A monomodal scattering represents that the sample contains molecules with one size distribution. A multimodal size distribution indicates the presence of various groups of particles or aggregates with differentiated sizes. Also, the polydispersity index can be obtained (**Figure 1.9**).

DLS can also be employed towards the identification of protein-protein interactions. By means of the ionic strength and the concentration dependence of the diffusion coefficients, different interaction parameters can be measured [23, 24].

1.6.3 Circular Dichroism (CD)

CD is another non-destructive spectroscopic technique that employs circularly polarized light. UV-CD is a commonly used technique to obtain information about the secondary structure of proteins, peptides and DNA. It is also widely used to characterize folding processes and monitor conformational changes and the presence or not of interactions between molecules.

Far-UV CD spectra can reveal secondary structure properties as it is widely used to estimate the fraction of α-helix, β-sheet and random coil conformations present in proteins and peptides in solution (**Figure 1.9**). Usually, mathematical deconvolution algorithms are used to quantify the CD data and derive the percentage of the secondary structure of these macromolecules. In this regard, the comparison of the CD data before and after the complex formation can reveal the establishment of macromolecular interactions, as well as the change of the secondary structure induced during the process [25].

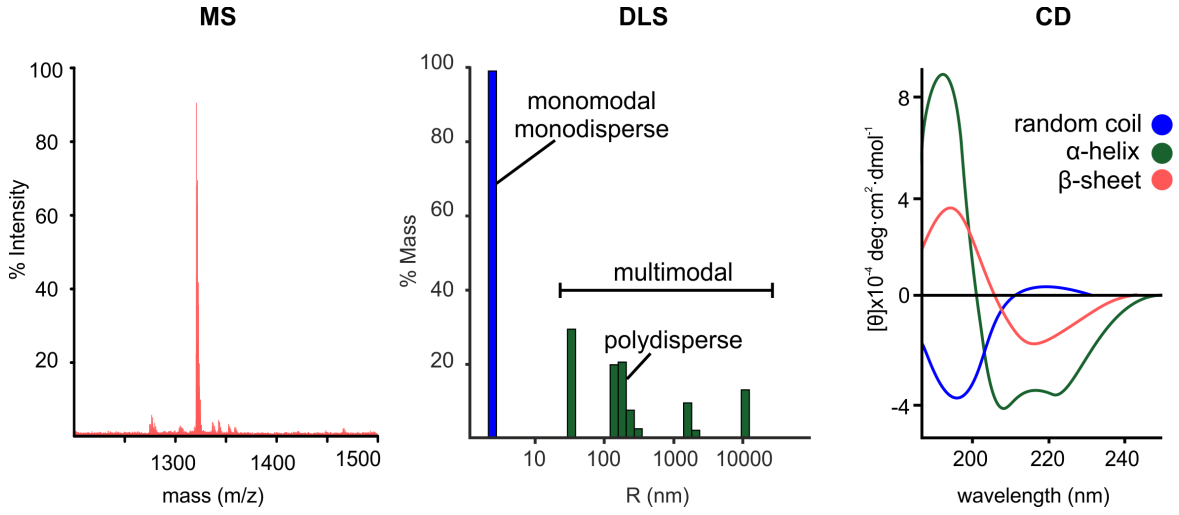


Figure 1.9. Biophysical techniques in biomolecular characterization. In mass spectrometry (MS), the protein can be identified in the basis of its mass; the spectrum plot generated shows the different isotopic forms. In dynamic light scattering (DLS), the presence of single particles or aggregates can be determined by means of the profile generated. A multimodal size distribution might indicate the presence of aggregates, while a monomodal bar is related with particles with one size distribution. By circular dichroism (CD), the secondary structure and folding properties of a protein can be analyzed depending on the range of the wavelength measured. Far-UV spectra can show the contribution of α-helix, β-strand and random coil conformations of the sample.

1.6.4 Isothermal Titration Calorimetry (ITC)

This technique allows to determine the binding affinity and the enthalpic and entropic components of biomolecular interactions quantitatively. These thermodynamic parameters are related with the association constant and these data are obtained through titration assays at varying temperatures [26, 27].

1.6.5 Fluorescence assays

In this technique, a fluorophore-labeled or Trp-containing protein or peptide is excited with a polarized light and the produced fluorescence signal is recorded, both in the free state and in the complex. The signal intensity changes due to the different correlation time shown by the protein and the complex, allowing the characterization of the corresponding interaction [28, 29].

1.6.6 Small-angle X-ray scattering (SAXS)

This is a powerful tool to analyze weak interactions. It gives information about the stoichiometry of binding, interaction potentials, the equilibrium states, and association and dissociation processes. It also give low-resolution assemblies that can be later used in combination with other high-resolution techniques (NMR and X-ray crystallography) to obtain a more complete model of the complex [30].

1.6.7 X-ray crystallography

X-ray is a high-resolution technique that offers spatial information at atomic level and it is widely used in biomolecular studies. It can be used to determine the structure of a single protein or nucleic acid, or an interacting complex. Despite in most cases being the technique of choice due to its multiple advantages, it is not always possible to crystalize complexes due to the presence of highly dynamic domains or protein instability [31].

1.6.8 NMR spectroscopy

1.6.8.1 NMR as a tool to study protein interactions

NMR is one of the best tools to obtain highly detailed information concerning biomolecular interactions, including weak (K_d > 100 μM) and highly weak interactions. One advantage of NMR is that it can be used under a variety of biologically relevant conditions. However, solution NMR struggles with poorly soluble molecules or with large and complex systems like biomembranes. In these cases, the application of solid-state NMR spectroscopy can be an alternative.

The traditional approaches towards the characterization of biomolecular interactions is through the nuclear Overhauser effect (NOE), the analysis of the chemical shift perturbations (CSP) and the scalar and residual dipolar couplings (RDCs) [32].

Nuclear Overhauser Effect (NOE) The Nuclear Overhauser Effect (NOE) is the most important NMR parameter from a structural point of view. During the NOE effect, the magnetization is transferred between protons that are close in space and therefore, the intensity of the NOE signals depends on the distance between these nuclei in an approximate relation of $1/r^6$ or up to 5-6 Å.

NOEs are also useful towards the determination of protein interactions. Given that this parameter gives information regarding the spatially closed atoms, the interacting elements of two or more molecules can be inferred. This approximation is more effective in high-affinity complexes (K_d < 100 nM). In low-affinity interactions, the concentration of the bound state in the equilibrium should be sufficiently high to be detected by NMR and this might be an issue if the predominant species in solution are on their free states. In the case of intermediate affinity complexes, their kinetic nature results in a significant signal broadening that can impede to observe the NOEs between nuclei at the interface [32].

Chemical Shift Perturbations (CSP) The chemical shift (δ) is a parameter that relates the resonance frequency of a nucleus with respect to a standard, and depends on its chemical environment. The different factors determining the chemical shifts are: local diamagnetism, local paramagnetism, neighbor anisotropy, ring-currents, the electric field, solvent contribution and the presence of paramagnetic species.

CSP is the most widely used parameter in NMR for the study of interacting molecules. In this application, fingerprint regions of NMR spectra obtained for a molecule on its free and

bound state are compared. It is known that the addition of an interacting molecule causes a change in the electronic environment of the biomolecule and therefore, the chemical shifts within the interface region vary. On the basis of these changes, the interacting region can be defined and mapped on the molecular surfaces. However, not all interfacial residues present significant chemical shift changes and as well, some proteins can undergo into greater structural rearrangements and give large CSP values not related with the interface region [32]. Due to these factors, a careful analysis should be addressed depending on each case.

Scalar (J) and Residual Dipolar Couplings (RDCs) Scalar coupling constants, also known as J couplings, are a measure of the interaction between two nuclear spins, where the interaction is mediated by the electrons participating in the bond(s) connecting these nuclei. The values of the scalar couplings, measured in Hz, depend on the molecular geometry and particularly, on the angles defining the molecular conformation. When an interaction between molecules occurs, structural changes can take place and modifications in the values of the scalar couplings are measured. This is how this parameter can be used to determine the presence of interactions.

On the other side, in anisotropic conditions, dipolar interaction between nuclei established through bonds, are not averaged and a new parameter can be measured and analyzed. In these conditions, molecules in solution align weakly relatively to the field, resulting in the appearance of residual dipolar couplings (RDC). These interactions are measured also in Hz and their absolute value depends on the degree of molecule alignment, the distance and the relative orientation of the pair of involved nuclei. The difference between the signals splitting within an anisotropic media (which is the sum of J and RDC), *versus* the isotropic one, gives the RDC values. In this regard, the measurement of the difference obtained between the free and the complex states provides information on the relative orientation between the molecules present in the interacting state [33, 32].

1.6.8.2 Advantages and limitations of NMR

As mentioned before, NMR spectroscopy is a non-destructive technique that offers large advantages compared with other biophysical techniques. This property makes it possible to study the same process or molecule at different conditions such as pH variations, temperature, cosolvents, in the presence of diverse ligands and repeat these experiments several times. Additionally, NMR spectroscopy can be used in solution, solid and gaseous phases, being the solution phase the most widely used in biomolecular NMR. NMR makes it possible to

analyze the molecule or complex under the most biologically relevant state. Also, molecules which can not be crystalized, can be subjected to study using this technique.

If the sample concentration is high, small polypeptides can be studied using not only ^{1}H, which is the most naturally abundant and sensitive NMR-active nuclei, but also ^{13}C, and ^{15}N heteronuclei at their natural abundance. This enlarges the amount of obtainable information and yields a more complete structural characterization. In the case of polypeptides greater than 15 kDa, the isotopic labeling and multidimensional NMR spectra recording is mandatory since the natural abundance of ^{13}C and ^{15}N is low.

Beside these advantages, the main limitation of NMR spectroscopy lie on its low sensitivity. As well, proteins are large molecules and this leads to a high number of resonance signals that might overlap and make spectral assignment difficult. Concerning the overlapping signals, ^{13}C and ^{15}N isotopic labeling might be a solution, but yet expensive. Regarding the low sensitivity of the technique, to overcome this issue, high concentrations of the sample are needed; this increases the cost and might lead to its precipitation, while its soluble proportion falls. Also, the implementation of a cryoprobe and a low salt media improve the quality of the spectra.

For large biomolecules, the correlation time increases as the molecule size does. This means that these molecules experiment slow tumbling, which results in NMR signal broadening. In these cases, the conventional experiments are not appropriate and more sophisticated approaches are needed [34, 35].

1.7 Theoretical methods to study biomolecular interactions

Theoretical methods to study biomolecular interactions are fundamentally focused to predict or infer structural domains, dynamical properties and possible interactions between molecules. They are usually computer-based algorithms that can collect big datasets that might derive from experimental outputs. One of the main theoretical approaches regarding this subject is molecular docking.

1.7.1 Molecular Docking

This method can employ experimental data as input to predict the preferred conformation adopted by a molecule when bound to another one in order to form a stable complex.

Docking algorithms are data driven approaches that integrate information from other biophysical methods such as NMR spectroscopy, X-ray crystallography, X-ray scattering and cryo-electron microscopy. Depending on the complexity of the program and the interface, the user can define sets of active and passive residues that were previously identified as relevant for the interaction. The program tests different docking combinations by exploring the space of all possible orientations of the molecules, this gives an output that is evaluated in the basis of favorable interactions. Finally, the best scoring structures calculated are selected and analyzed by the researcher [36–38].

1.8 The use of minimized systems to surpass NMR limitations

As already mentioned, larger entities (including large proteins, big protein-protein complexes or membranes) are not as feasible to be fully and successfully characterized through standard solution NMR. Apart from the specific experiments mentioned before, which have limited applications, other experimental approaches like the use of minimized systems might be the best choice to address the study of certain systems [39].

1.8.1 Model peptides as protein mimics

One of the minimalist approaches is the design and characterization of model peptides that mimic the secondary structure believed to be adopted by a portion of the whole protein. The use of synthetic peptides offers a great advantage as they can be exact copies of protein regions and amino acids can also be easily replaced. This minimalistic tactic towards the study of the structure and function of bigger polypeptides and proteins usually starts with the formulation of a hypothesis that proposes which regions are potentially essential for a determined interaction. Then these portions can be analyzed under diverse experimental conditions. However, it has to be taken into account that regions not included in the designed peptides, even if they are sequentially far, might still exert weak interactions that would contribute to the behavior of the full system. Thus, unfortunately the effects of this type of interactions can not be examined by this approach [40].

It is important to remember that small peptides are more likely to be highly mobile in solution and, in the case of having a preferred conformation, it is usually in fast exchange with the random coil distribution in the angular space. In this regard, in order to establish a

well-founded conclusion, it is necessary to evaluate together as many information as possible and analyze additional data derived from other biophysical approaches [41, 42].

Considering these premises, the design of the correct peptides that mimic the portion of interest cannot always be straightforward; and three main different approaches can be considered depending on the information available concerning to the system of study. As well, more than one approach can be used based on: i) the polypeptide sequence, ii) the protein structure and iii) the protein function [40].

1.8.1.1 Sequence-based peptide design

This peptide scanning strategy is based on the synthesis of short overlapping peptides carefully selected that belong to an entire or a large section of a protein (**Figure 1.10**). Each individual peptide is studied on its free state and in the presence of the partner targeted by the full protein. In this way, it is possible to identify which peptides are affected by the presence of the protein partner, and so the regions involved in the interaction are mapped [40]. However, and as mentioned before, this approach should take into account that some protein binding sites might be formed by regions far apart in the sequence, and so they can not be mimicked by overlapping peptides. Consequently, it can be relevant to make use of more than one strategy depending on the system [40].

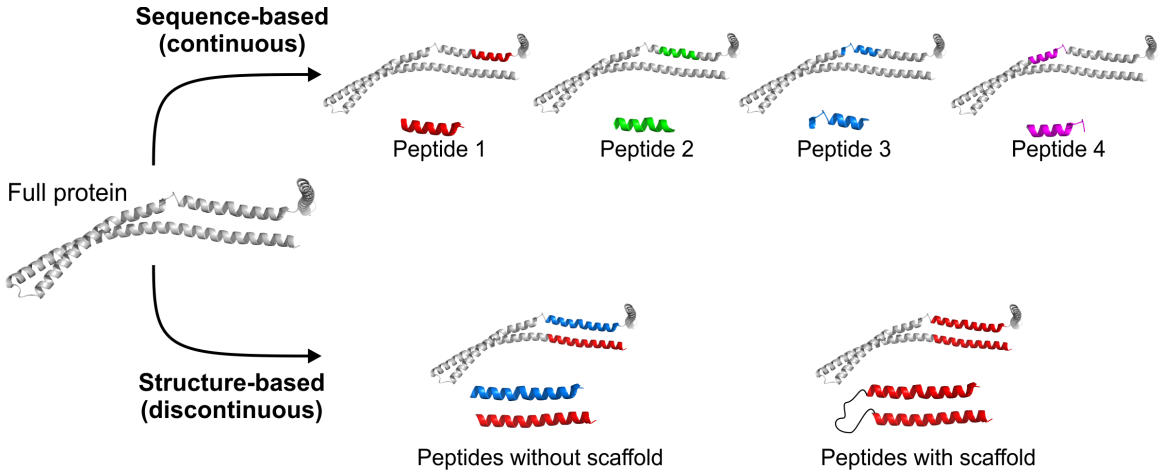


Figure 1.10. Design of peptide models as protein mimics. In the sequence-based (upper panel) peptide design, short overlapping peptides are selected to be studied mainly separately. In the structure-based peptide design, discontinuous segments believed to be involved in interactions are selected and studied principally together, a scaffold approaching the segments might also be necessary.

1.8.1.2 Structure-based peptide design

The design of peptides based on the protein structure consists of the selection and identification of the regions involved in the interactions within the full protein and synthesize a mimetic version of this binding site (**Figure 1.10**). These regions may be far apart from a sequential point of view and as the mimetic should comprise all the sequences involved, a molecular scaffold might be needed to function as a linker between the different peptide regions [40] [43].

1.8.1.3 Protein function-based peptide design

Finally, this kind of design is based on the known protein function. This is fundamentally a randomized computer-based strategy in which combinational peptide libraries are employed to perform a screening procedure that explores the possible binders for a determined protein. The possible peptides combinations obtained from the screening might result in inhibitors or sequences that activates the protein [40, 44, 45], and depending on the objectives of each project, different peptide combinations can be studied.

1.8.2 Micelles as membrane mimetics

As biomembranes are complex and large systems, they constitute a challenge for their study by NMR. For years, a simplified approach based on the use of different surfactants or detergents has been widely adopted. However, in order to choose the correct mimetic and keep the conditions closest to the native environment, the variety of the lipids that constitute the biological membranes, their composition and properties has to be taken into account. However, this is not always feasible due to the membrane complexity.

One traditional approach is to study the membrane protein in the presence of organic solvents of different polarity. The employment of non-polar cosolvens (**Figure 1.11**) such as trifluoroethanol (TFE) and hexafluoroisopropanol (HFIP) is very popular in the study of peptides by NMR as they mimic non-polar environments and promote the stabilization of the secondary structure in a similar way as they might adopt in the biological membrane [46].

Another widely adopted simplified membrane model system in NMR consists of the use of surfactants or detergents that spontaneously form micelles (**Figure 1.11**) and due to their amphiphilic nature, micelles are better membrane mimetics than TFE or HFIP. The selection of the surfactant should be based on the shape, curvature, thickness, lateral pressure, dielectric

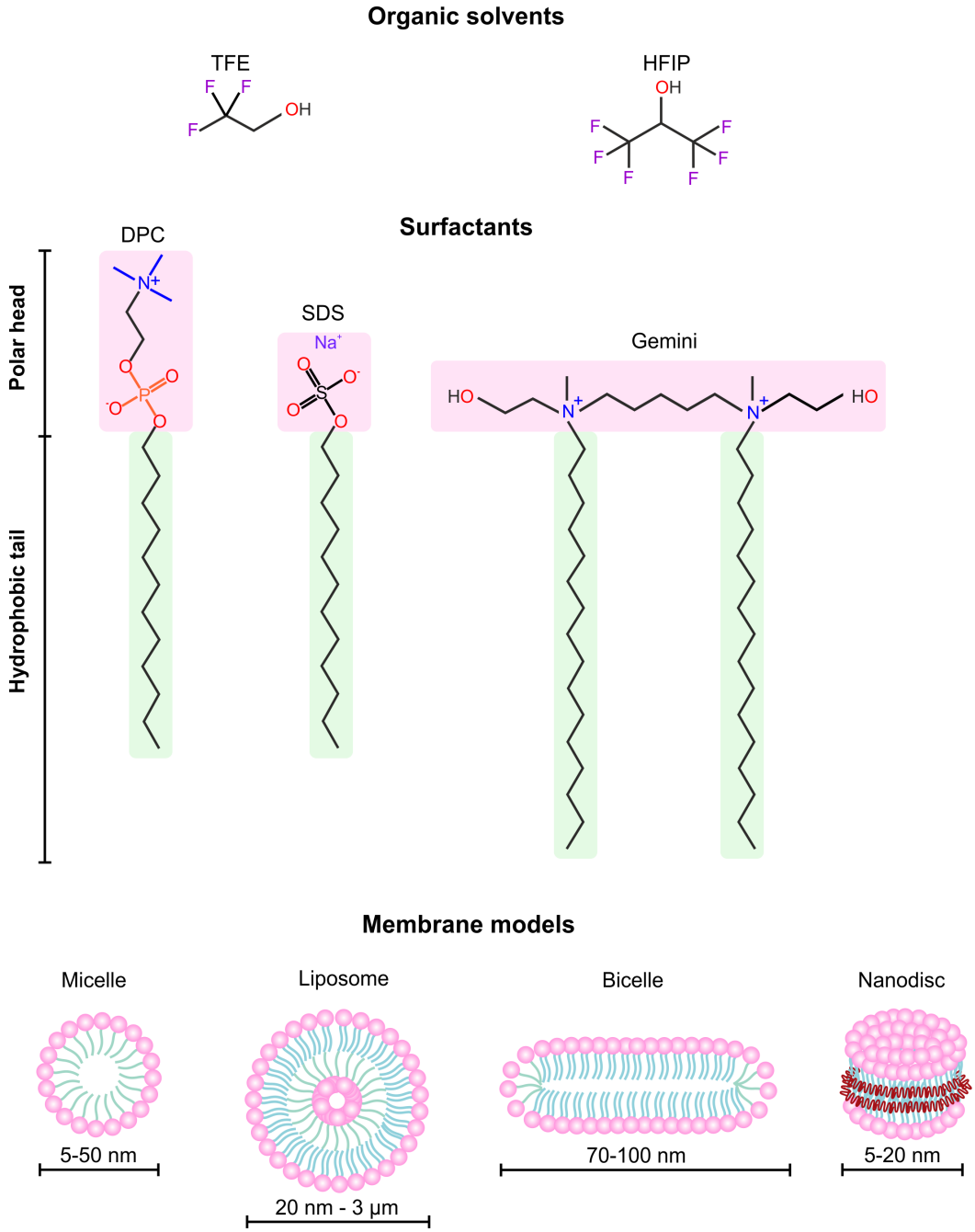


Figure 1.11. Organic solvents and surfactants used as membrane mimetics in NMR. Fluorinated alcohols such as trifluoroethanol (TFE) and hexafluoroisopropanol (HFIP) are used as cosolvents to maintain the biologically relevant structure of a membrane protein. Surfactants such as sodium dodecyl sulfate (SDS), dodecylphosphocholine (DPC) and Gemini associate and form micelles and embed membrane proteins. Due to their size, micelles are the most widely used membrane models in solution NMR.

constant, hydration and charge of the native membrane [47] and taking into account these properties, a variety of detergents have been characterized and selected to be the preferred for NMR analysis, such as sodium dodecyl sulfate (SDS) and dodecylphosphocholine (DPC).

Membrane proteins can keep their structure and native states in the presence of membrane mimetics, and can be solubilized in the presence of surfactants or detergents. If the detergent used is at a concentration higher than its critical micelle concentration (CMC), the detergent can solubilize both membrane lipids and integral proteins, keeping the protein within a micelle composed by a mixture of detergent and the native membrane lipids. On the contrary, if the concentration employed is below its CMC, the protein gets dissolved and separated from the membrane without forming micelles as surfactant molecules coat the hydrophobic regions of the protein (**Figure 1.12**).

Detergents that are often used to solubilize proteins can be ionic, non-ionic and zwitterionic [48]. Ionic detergents such as SDS are composed by a hydrophobic tail and a polar head with either a positive or negative charge. Ionic detergents are widely used for NMR spectroscopy, as they afford good quality NMR spectra, but they have to be used with caution as they might also denature proteins [48].

On the contrary, non-ionic detergents such as Triton X-100® or Tween® are composed with a hydrophobic region and a non-charged polar head. However, due to their monodisperse nature, they do not tend to form stabilized micelles at the concentrations needed for NMR experiments and can bring bad quality spectra [48].

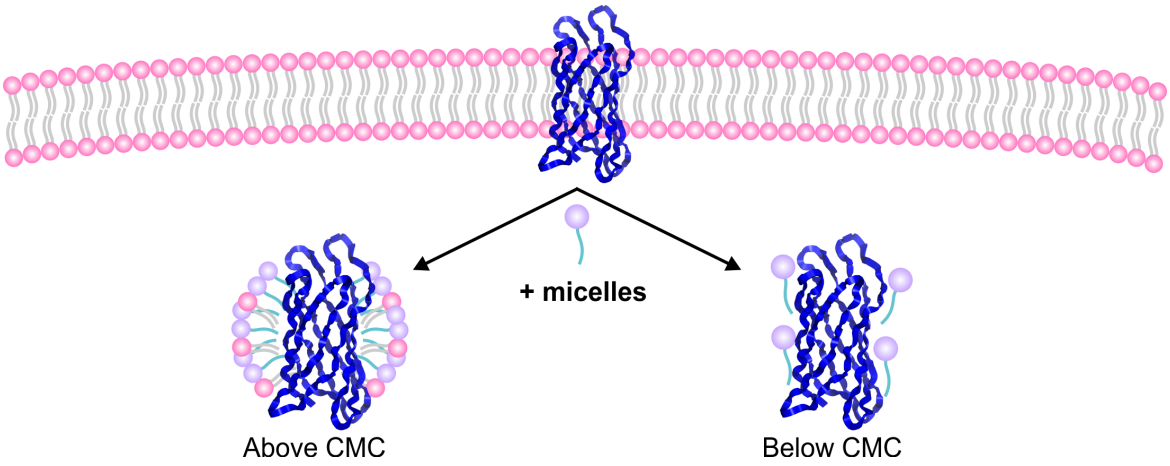


Figure 1.12. Solubilization of integral membrane proteins. The addition of micelles at a concentration higher than the critic micelle concentration (CMC) creates a micelle with an embedded protein that can be studied as a mimic of its native membrane environment, while at a lower CMC values, the protein gets dissolved and might be denatured.

Zwitterionic detergents such as DPC, have a net neutral charged polar head and a hydrophobic portion. Similar to ionic detergents, they can solubilize proteins, but their aggressiveness is lower, thus maintaining the native state and charge of the embedded proteins [48].

A final consideration towards the use of detergents as membrane mimetics is their critical micelle concentration (CMC). Above this concentration, surfactants tend to form micelles. Common surfactants used in NMR have high CMC values, thus, they have to be employed in their deuterated forms. Signal broadening is also a problem aroused from micelle formation and the use of combined experiments in the presence of non-polar environments (TFE, HFIP) might help to complement the spectra analysis [49, 50].

Regarding the high CMC values of popular detergents for NMR, the use of the not so common Gemini surfactants can be also interesting. Their low CMC, slow millisecond monomer to micelle kinetics and smaller micelle sizes make them suitable as membrane mimetics for NMR study. Hence, their low CMC value allows using them at a low concentration without being deuterated and the interactions found are due to the polypeptide being associated with the micelle and not with the monomeric form of the surfactant [49].

Additionally, there are other membrane mimetics that can be employed for solution NMR studies, such as bicelles, liposomes or nanodiscs. Bicelles are planar surfaces usually composed by a mixture of lipids and detergents that mimic the lipid bilayer; however, they are less stable than micelles and their use is not very widespread in solution NMR. Similarly, liposomes are also less stable, and their larger sizes contribute to the signal broadening a fact that deteriorates the quality of the NMR spectra. Finally, nanodiscs are composed by a phospholipid bilayer stabilized by helical proteins that act as scaffolds, and given their stability and bilayer nature, they are considered better membrane models than micelles and liposomes and are gaining popularity in solution NMR spectroscopy [51, 52].

1.9 Global aims and structure of this Thesis

Throughout this Introduction chapter, it has been highlighted the importance of the biomolecular interactions, particularly those involving protein-protein and protein-membrane, and the fundamental role played by NMR in their characterization. In this context, this Thesis presents the design of experimental approaches intended to analyze and characterize profoundly the molecular basis of these kinds of interactions that are essential in the establishment and regulation of certain biological processes. This has been done for proteins and complexes found in diverse processes, which are implicated in a wide variety of organisms.

In order to obtain deep information of these complex processes, solution NMR results have been analyzed in combination with those obtained with other biophysical and theoretical methodologies. In this regard, the proper use and selection of the experimental conditions and the different methodologies can make this analysis biologically relevant and shed light on the comprehension of the complex biological processes in which the selected proteins are involved.

The general objectives of this Thesis are:

- To perform a structural analysis of the centrosomal TACC3-XMAP protein-protein interaction through the minimization of the complex employing peptides and cosolvents.
- To characterize the glycoprotein gp41 and its interactions through the structural study of a simplified version of peptides and membrane mimetics.
- To determine and characterize the interactions of the peptides derived from the antimicrobial peptide protegrin-1 in the presence and absence of membrane mimetics.
- To obtain the structure and characterize the molecular interactions with membranes of the fungal sex α-pheromone of *Fusarium oxysporum* in the presence of membrane mimetics.

On the basis of these general objectives, the organization of the present Thesis is explained in the following section.

1.9.1 Organization of this Thesis

This work has been structured in six chapters, four chapters based on the original results of this work, and another two separated chapters dedicated to the general introduction and the overall conclusions. Chapter 1 is an introduction in which the relevance of the biomolecular complexes and the methodology used towards their characterization is described. The last chapter, chapter 6, includes a summary of all the results obtained and the general conclusions of this Thesis.

The original sections comprise four experimental chapters (chapters 2 to 5) composed by the same structure and contain a specific introduction of the system studied, detailed objectives, experimental section, discussion of the results, and their specific conclusions.

Chapter 2 is focused on the study of the XTACC3-XMAP215 protein-protein interaction. This interaction has been found to be relevant in cell division and cancer-related processes and its study might help to understand the specific molecular recognition and biological function of these proteins. In this part, a structure-based peptide approach was chosen to design the peptide models for this kind of interaction; as well, NMR, circular dichroism and computational simulation methods were employed to characterize these peptides.

Chapter 3 contains the study of a biologically relevant protein-membrane interaction comprising different regions of the glycoprotein gp41 of HIV-1. Simplified versions of the interaction were characterized employing peptide models derived from a sequence-based approach in the presence of membrane mimetics. The pursued goal is to understand how the virus can be neutralized prior to its binding with the host cell, thus, it might yield relevant information towards the design of new anti HIV drugs.

In chapter 4, an exhaustive characterization of the peptide-peptide and peptide-membrane interaction of three variants of the antimicrobial peptide protegrin-1 has been performed through NMR, circular dichroism and computational docking simulations. We aimed to have new insights about the structure and function relationships of this member of the protegrin family that has been proposed as a good candidate for the design of novel antimicrobial agents.

Chapter 5 is dedicated to the structural and functional description through NMR and other experimental techniques of a peptide pheromone implicated on the reproductive cycle of the highly destructive fungus *Fusarium oxysporum*. The main objectives consist of the determination of the 3D structure and the delineation of the peptide regions that are crucial for specific functions such as its chemoattractant and growth inhibitory activities.

Finally, at the end of this Thesis, the Bibliography and the Appendices containing additional quantitative data used along all chapters can be found.

2

Structural study of the centrosomal TACC3-XMAP protein-protein interaction from a minimized system of peptides and cosolvents

2.1 Introduction and objectives

The cell cycle is a vital process in life, it comprises several biochemical routes along two main states: interphase and mitosis. The interphase is comprised by four different phases: G_1, G_0 (a quiescent state), S (synthesis) and G_2 (**Figure 2.1**). During these stages, different metabolic pathways take place, such as cell growth at G_1, DNA replication at the S phase, and the preparation for cell division at G_2. During the M phase (mitosis), the cell division occurs and two daughter cells are created. Therefore, all along this process it is crucial to distribute the genetic material properly into the new daughter cells. Defects on this process may result in important consequences like chromosomal instability, aneuploidy, apoptotic cell death, premature senescence and cancer [53–55].

One fundamental player in this process of cell division found in metazoans is the centrosome. The centrosome is an organelle that acts as the main microtubule-organizing center,

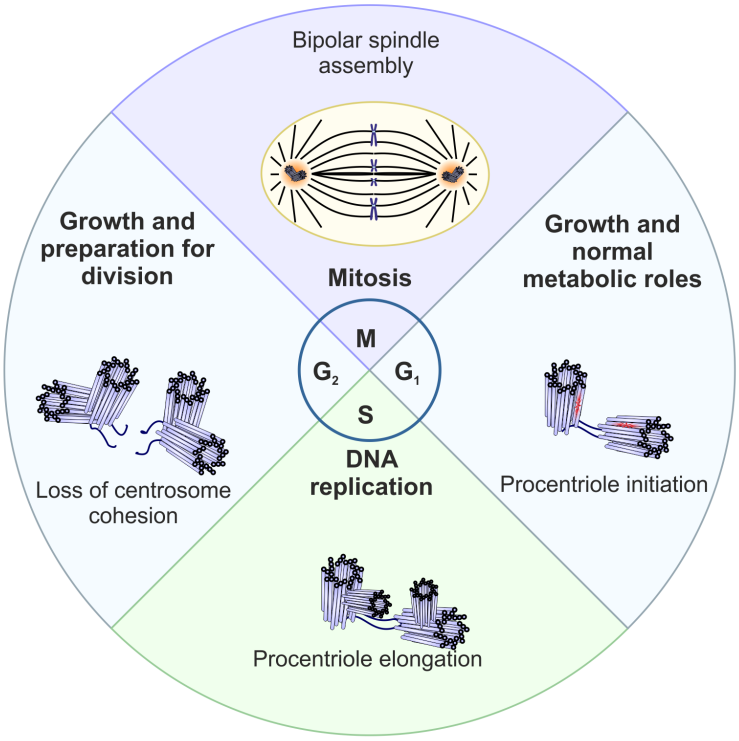


Figure 2.1. Schematic representation of the cell and centrosome cycles. At phase G_1, the cell performs its usual metabolic roles. At synthesis phase (S), the DNA replicates and the procentrioles start to elongate. At G_2 phase, the centrosomes breaks their cohesion and at mitosis (M) phase, the centrosomes locate in the opposite sides of the cell in order to arrange the genetic material prior to division.

and it is responsible for the distribution of the genetic material to the daughter cells. To do this essential job, centrosomes experience a division and maturation process during the so-called centrosome cycle (**Figure 2.1**) [56–58]. The structure and composition of the centrosomes is complex, transient and diverse. During the mitotic phase, they consist of one pair of centrioles surrounded by the pericentriolar matrix; as well, several proteins participate on their assembly and organization through different interactions and signaling routes [59–61, 57], see **Figure 2.2**.

During cell division, the nuclear membrane gets disassembled, an association of tubulin αβ-heterodimers occurs, in which randomized phases of polymerization and dissociation take place at the microtubules (MT) limits; and finally, the important spindle fibers are created [62]. The mitotic spindles are complex structures formed by MTs arranged as strong fibers that are part of the cell skeleton. They exert the forces required to align and segregate the genetic material precisely by controlling the motor and non-motor microtubule-associated proteins (MAPs), which define and organize the dynamic properties of MTs [63]. Therefore, MAPs are critical in this process in which self-correction mechanisms should occur to avoid errors during the chromosomal distribution that can contribute to tumor development and diverse cellular failures [63].

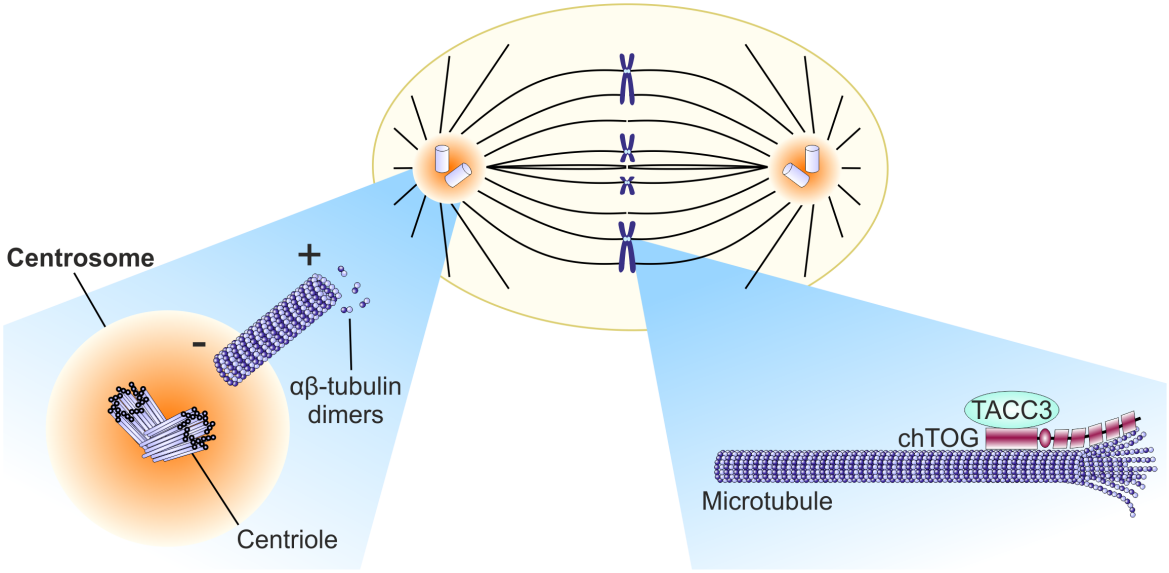


Figure 2.2. Bipolar spindle assembly diagram. During cell division, the centrioles replicate and migrate to an opposite side of the cell. Throughout the formation of the spindle poles, maturation of the centrosomes takes place, the pericentriolar material (orange) reorganizes and the microtubules get assembled. TACC3 binds to chTOG and promotes microtubule growth by the addition of αβ-tubulin heterodimers.

2.1.1 The TACC3-XMAP interaction promotes microtubule growth

Considering the above scenario towards the understanding on how these MAPs work and interact, and to achieve a description of the corresponding structure-function relationships, it is mandatory to obtain information at atomic level. In this context, we focus on an interaction that promotes MT growth: the interaction between one of these MAPs (XMAP215) and the centrosomal adaptor protein TACC3.

The colonic hepatic Tumor Over-expressed Gene (chTOG) (first reported in colonic and hepatic tumors [64]) is the human homolog of the *Xenopus laevis* microtubule-associated protein (XMAP215) [65]. XMAP215 is a potent polymerase, a promoter of plus end microtubule assembly through the addition of $\alpha\beta$-tubulin heterodimers [66, 67]. chTOG/XMAP215 is a highly conserved protein family that shares a common C-terminus coiled-coil domain and a variable number of TOG (from two to five) domains within its N-terminus region. Each TOG domain is capable on binding tubulin dimers at the microtubule end [68, 69].

TACC (transforming acidic coiled-coil) proteins are centrosomal adaptors that play an important role during the mitotic spindle assembly [70–72]. TACC proteins, specifically TACC3, were first found in a search for genomic regions amplified in breast cancer [70]. It is known that TACC3 interacts with chTOG/XMAP215 and regulates MT assembly [73, 74, 71]. A reduction of the protein levels of TACC3 disrupts the proper localization of the chTOG/XMAP215 association at the mitotic centrosome [75, 76], while increased levels of TACC3 enhance the recruitment of the complex at the spindle poles [70, 73].

2.1.2 TACC3-XMAP interaction model

Quite recently, the interaction between TACC3 and XMAP was described based on biochemical and biological data by Mortuza and coworkers [77]. They defined a minimal active binding domain (TD4, 26 kDa) for the protein XTACC3 (the *Xenopus* homolog of the human TACC3, with a 79 % of sequence identity) and also for XMAP215 (C-t, 7 kDa) [77]. It is well known that human centrosomal proteins show a significant tendency to adopt both unstructured and coiled-coil conformations [78]. This complicates the determination of their 3D structure at atomic resolution. In this regard, the high flexibility and oligomerization propensity of TD4 thwarts the resolution of its 3D structure by either NMR or X-ray crystallography. For this reason, an experimental structural model of this complex is still lacking.

Despite these limitations, a medium resolution 3D model of TD4 is available based on small-angle X-ray scattering (SAXS) data. This model consists of an elongated helical

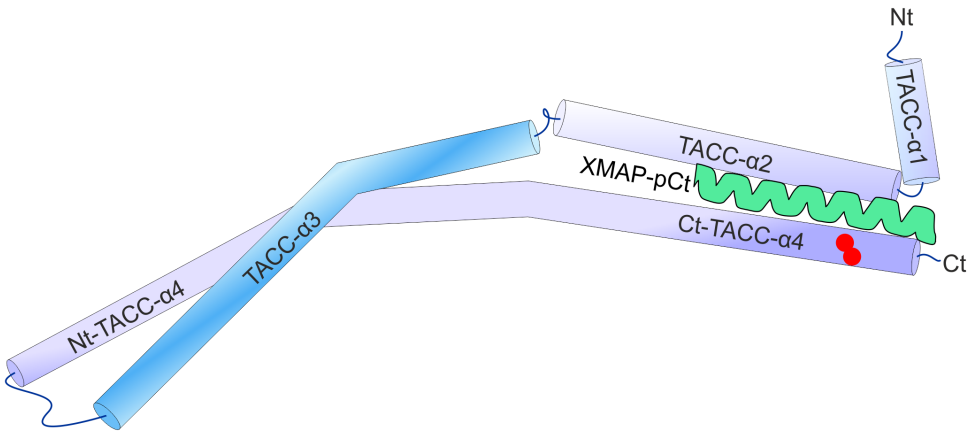


Figure 2.3. Cartoon representation of the XTACC3-XMAP interaction through the XMAP C-terminus domain (adapted from [77]). XTACC3 aspartic residues that are important for the activity (Asp921 and Asp922) are shown in red, XMAP-pCt two helices are superimposed. The N- and C-termini are also indicated.

structure, in which the biologically relevant helix α2 (TACC-α2) and the C-terminus of the helix α4 (Ct- TACC-α4) face each other in an antiparallel orientation.

On the other hand, the C-terminus binding region of XMAP is also predicted to be a coiled-coil. When isolated, XMAP tends to populate helical structures and the short peptide XMAP-pCt (**Table 2.1**) has been described to have similar structural propensities. It has also been shown that the pCt peptide has a good affinity with TD4. Therefore, it is an appropriate model to mimic the intermolecular interactions of the full-length protein [77].

Likewise, a structural model for the XTACC3/XMAP215 (1:2 ratio) complex has been previously proposed [77]. In this model (**Figure 2.3**), the helical structure of XMAP-pCt lies in the grooves between the helix α2 and the C-terminus region of the helix α4 of TD4. It was also shown that TD4's helices α2 and α4 contain several key residues for binding. Interestingly, two single point mutations in the C-terminus of the helix α4 (D921A and

Table 2.1. Sequences of the peptides studied, aspartic residues of Ct-TACC-α4 wt are shown in red, mutated residues in Ct-TACC-α4 mut are underlined. The sequence number of these residues in the full length protein are also indicated.

Peptide (protein sequence #)	Sequence	Theoretical pI
TACC-α2 (823-856)	VEMGKIIAEFEGTITQILEDSQRQKETAKLELNK	4.82
Ct-TACC-α4 wt (897-930)	MKIQSLERSLEQKSKENDELTKISDDLILKMEKI	5.18
Ct-TACC-α4 mut (897-930)	MKIQSLERSLEQKSKENDELTKISAALILKMEKI	8.1
XMAP-pCt (2047-2066)	STNIDDKLKKRLERIKSSRK	10.55

D922A) are sufficient to impair the molecular interaction and protein function [77]. Although the proposed SAXS model could explain some biological observations regarding these proteins' association and microtubule elongation, many important questions and the atomic structural details are still lacking.

2.1.3 The advantages of the use of simplified systems

The best way to extract accurate and biologically relevant information regarding the structure and function relationships is to study the biological system with the conditions that better reproduce its native environment in the cell. However, this is not usually experimentally easy. In our case, we are interested on obtaining information at atomic level of a large complex of centrosomal proteins, which have a tendency to aggregate and interact in the pericentriolar media. Both system and media are in this case hard to reproduce and study employing biophysical methods. Consequently, no models of atomic structures under these conditions have been reported so far. To overcome these problems, some simplifications were proposed in different examples [40, 79, 80]. A strategy widely validated in recent decades is the use of linear peptides as simple models to reproduce the globular structures of proteins, and among these studies there are important contributions of our group [81–83]. Peptides are also used routinely for the characterization of protein-protein interactions, whose study is simplified by using protein-peptide complexes, in which the peptide contains the recognition site for binding [84, 85]. As well, peptide-peptide interactions have been recently used to reproduce protein-protein interactions [86]. This approximation is valid given that the peptides are generally capable to reproduce the conformational tendencies and the behavior of the selected sequences in the full proteins [40].

Linear peptides show many benefits in structural studies. They are easy to obtain in large amounts by standard peptide synthesis; in the case of problems, the addition of different N- or C-tags can increase the solubility; they are easily studied by homonuclear NMR if they are not too large; and a variety of different NMR parameters can bring us information about their structure and interactions. The most powerful ones are chemical shifts, coupling constants and intermolecular NOEs. All of them vary depending on structural modifications and intermolecular associations. As a main drawback of this approximation, we can mention the peptide conformational average in solution. In consequence, the discussion of some results is not always an easy task, but conclusions about the most populated conformers can be generally extracted [41]. Finally, all these details have been discussed in the introductory chapter of this Thesis.

On the other side, physiological conditions can be mimicked by varying the pH, ion concentration, or by adding cosolvents or crowding agents. In this study, TFE and DPC micelles were selected as cosolvents since they mimic low dielectric environments, such as those found within high concentrations of macromolecules and other entities encountered in cells [87]. These cosolvents promote intramolecular hydrogen bonds in peptides and proteins and are known to stabilize pre-existing secondary structures in short linear peptides [88, 89].

In summary, by using mimic systems, we can study some relevant complexes that would be unattainable by other methods.

2.1.4 Objectives

The main objective of this study is to obtain new insights concerning the interaction between XTACC3 and XMAP215; and consequently, to fully understand the functional implications of this complex and the regulation of the microtubule dynamics during mitosis. Given the current limitations on gaining structural data at atomic scale for this large system, in this study we make use of a simplified version, which consists of: i) peptides derived from XTACC3 and XMAP215 and ii) the addition of cosolvents. Four different peptides corresponding to TACC-α2 and Ct-TACC-α4 domains, and pCt from XMAP215 (**Table 2.1**) were designed and characterized by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopies in water, TFE mixtures and DPC micelles. TACC3 and XMAP215 peptides were analyzed both alone and as part of all possible interacting mixtures.

The specific objectives proposed are:

- To define the regions and residues that can be crucial for the interaction between TACC3 and XMAP215 by means of a careful analysis of the interaction model proposed with SAXS data.
- To design peptides able to mimic the conformation of those sequences in the native full length proteins.
- To fully characterize these designed peptides in different media by CD and NMR, and examine whether their structures in the free state and within the available SAXS model are the same.
- To determine the structural consequences of key aspartic mutations. Particularly, the two aspartic acids located at the C-terminus of the helix α4 of TACC3; based on biological assays, were found to be important for the interaction and function and for which no information has been reported regarding to possible structural changes given by these mutations.
- To characterize at atomic level how the selected peptides interact and how these interactions are compatible with the SAXS model.
- To provide a structural model based on the experimental atomic data derived from the study of the designed linear peptides.
- To derive the structure-function relationships related with the TACC3-XMAP215 complex.

2.2 Materials and methods

2.2.1 Chemicals

The deuterated compounds [D_{38}] DPC (dodecylphosphocholine) (98 %), [D_3] TFE (2,2,2-trifluoroethanol) (99 %) and D_2O (99.9 %) were obtained from Cambridge Isotope Laboratories (USA). The percentages of deuteration are indicated in parentheses.

2.2.2 Peptide synthesis

TACC-α2, Ct-TACC-α4 wt, Ct-TACC-α4 mut and XMAP-pCt peptides were synthesized by Caslo Aps (Lyngbym, Denmark), peptide sequences are listed in (**Table 2.1**). Fmoc-solid phase synthesis protocols were followed and purification was performed by reverse-phase HPLC using a C18 column until reaching 95 % or more purity. The identity of the peptides was confirmed by MALDI-TOF MS. To better mimic the behavior of the peptides in the protein, peptides were acetylated and amidated at their N- and C-termini, respectively.

TACC-α2: RP-HPLC: tR = 12.13 min, purity 98.58 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 40-60 % B buffer gradient in 20 min). HRMS: Theoretical molecular mass: 3932.54 (m/z); $[M+H]^+$ found: 3932.73.

Ct-TACC-α4 wt: RP-HPLC: tR = 9.89 min, purity 95.38 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 40-60 % B buffer gradient in 20 min). HRMS: Theoretical molecular mass: 4075.79 (m/z); $[M+H]^+$ found: 4076.97.

Ct-TACC-α4 mut: RP-HPLC: tR = 10.60 min, purity 98.86 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 32-47 % B buffer gradient in 15 min). HRMS: Theoretical molecular mass: 3987.77 (m/z); $[M+H]^+$ found: 3989.48.

XMAP-pCt: RP-HPLC: tR = 11.15 min, purity 97.15 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 15-33 % B buffer gradient in 18 min). HRMS: Theoretical molecular mass: 2456.90 (m/z); $[M+H]^+$ found: 2457.52.

2.2.3 CD spectroscopy

CD spectra were recorded with a Jasco J-815 spectropolarimeter (Tokyo, Japan) equipped with a Peltier PTC-423S system. Stocks of known peptide concentration (about 1 mM) were prepared by weighing the lyophilized material. Aliquots of these stock solutions were taken and diluted into the appropriate volume of buffer to yield the peptide samples at the concentration required for each spectroscopic characterization. Samples were centrifuged 5 min, 13000 rpm prior to CD measurement.

All measurements were carried out by triplicate at 5 ºC, the cuvette path lengths were 0.1 cm in the far-UV region (250-195 nm). Individual peptide samples were prepared at a concentration of 30 μM and characterized under three different solvent conditions: i) 20 mM sodium phosphate, 150 mM NaCl, pH 5.5; ii) 30 % TFE, 20 mM sodium phosphate, 150 mM NaCl, pH 5.5; and iii) 20 mM DPC micelles, pH 3.0, 5.5 and 10.0. To test aggregation tendencies, spectra at 30 μM, 45 μM and 60 μM peptide concentrations were also recorded in H_2O. For the mixtures, samples containing 30 μM of each peptide in 1:1, 1:1:1 or 1:1:2 ratios were used.

Isothermal spectra for these samples were acquired at a scan speed of 50 nm·min^{-1} with a response time of 4 s and averaged over four scans for each sample. The same parameters were used to record the reference spectra with their respective buffer/cosolvent solutions. After subtracting the appropriate reference spectra from the sample spectra, CD data were processed with the adaptive smoothing method of the Jasco Spectra Analysis software. CD data are given in molar ellipticity units ([θ], deg·cm^2·dmol^{-1}, **Equation 2.1**) for the isolated peptides and ellipticity units (θ, mdeg) for the mixtures. The helical population of the different peptides was determined by: i) the method described by Rohl & Baldwin [90] (which considers peptide length and temperature corrections) on the basis of the mean-residue ellipticity experimentally observed at 222 nm; and ii) the on-line analysis for protein circular dichroism spectra (Dichroweb) server using the K2D Analysis programme [91].

$$\theta(deg \cdot cm^2 \cdot dmol^{-1}) = \frac{100 \cdot (CD)}{C \cdot n \cdot l} \qquad \textbf{(Equation 2.1)}$$

In which CD is the signal in mdeg, C is the sample concentration in mM, n is the number of amino acids and l is the path length in cm.

2.2.4 NMR sample preparation

NMR samples were prepared by dissolving the lyophilized peptide (1-2 mg) in 0.5 mL of solvent; i.e., H_2O/D_2O (9:1 v/v), pure D_2O, 20 mM [D_{38}] DPC in H_2O/D_2O (9:1 v/v), 20 mM [D_{38}] DPC in D_2O, 30 % [D_3] TFE in H_2O/D_2O (9:1 v/v), and 30 % [D_3] TFE in D_2O. The peptide concentrations were about 1.0 mM. The pH was adjusted to 5.5 by adding minimal amounts of NaOD or DCl, measured with a Hamilton glass micro-electrode. For the trio mixture, the sample was prepared with 0.5 mM of each peptide with a 1:1:1 peptide ratio. All samples were placed in 5 mm NMR tubes, and contained DSS as the internal reference for ^{1}H chemical shifts.

2.2.5 NMR spectra acquisition

NMR spectra were recorded with Bruker Avance spectrometers, operating at 600 and 800 MHz (^{1}H), both equipped with triple resonance, z-field gradient cryoprobes. Phase-sensitive two-dimensional correlated spectroscopy (COSY), total correlated spectroscopy (TOCSY) and nuclear Overhauser enhancement spectroscopy (NOESY) spectra were recorded with standard techniques. Water signal was suppressed by either presaturation or by using a Watergate pulse sequence [92].

TOCSY spectra were obtained by using 20 ms and 60 ms mixing times. The NOESY mixing time was 150 ms. ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra were recorded at ^{13}C natural abundance. ^{13}C δ-values were indirectly referenced by using the IUPAC-IUB recommended ^{1}H/^{13}C chemical shift ratio (0.25144953) [93]. Depending on peptide solubility, different temperatures were tested (5, 25 and 40 ºC). Data were processed with the standard TOPSPIN program (version 2.1) (Bruker Biospin, Karlsruhe, Germany). NMR and sample conditions are summarized in **Table 2.2**.

2.2.6 NMR spectra assignment

The NMR spectral assignment of each peptide in the different conditions was performed by following the well-established sequential-specific methodology based on homonuclear spectra [94] with the help of the SPARKY software [95]. The ^{13}C resonances were identified on the basis of the correlations between the hydrogen and the bound carbon atom present in the ^{1}H-^{13}C-HSQC spectra. All assigned chemical shifts are listed in **Appendix A**.

Table 2.2. NMR spectra conditions of the peptides studied.

Peptide	Solvent	pH	Temperature (ºC)
TACC-α2	H_2O	5.5	5
	D_2O	5.5	5
	30 % TFE/H_2O	5.5	25
	30 % TFE/D_2O	5.5	25
	20 mM DPC/H_2O	5.5	25
	20 mM DPC/D_2O	5.5	25
Ct-TACC-α4 wt	H_2O	5.5	5
	D_2O	5.5	5
	30 % TFE/H_2O	5.5	25
	30 % TFE/D_2O	5.5	25
	20 mM DPC/H_2O	5.5	25
	20 mM DPC/D_2O	5.5	25
Ct-TACC-α4 mut	H_2O	5.5	5
	D_2O	5.5	5
	30 % TFE/H_2O	5.5	25
	30 % TFE/D_2O	5.5	25
	20 mM DPC/H_2O	5.5	40
	20 mM DPC/D_2O	5.5	40
XMAP-pCt	H_2O	5.5	5
	D_2O	5.5	5
	30 % TFE/H_2O	5.5	5, 25
	30 % TFE/D_2O	5.5	5, 25
	20 mM DPC/H_2O	5.5	25
	20 mM DPC/D_2O	5.5	25
TACC-α2 + α4 wt + XMAP-pCt	H_2O	5.5	5
	30 % TFE/H_2O	5.5	5, 25

2.2.7 Structure calculation

Structure calculation of the peptides was done with the program CYANA 2.1 [96]. Small peptides do not adopt a unique, highly stable structure; instead, they adopt an ensemble of preferred, similar and modestly stable conformers, which are in equilibrium with low-populated conformers. To characterize the preferred conformers, medium and long range NOEs were selected. Intraresidue and sequential NOEs (except for HN-HN, which are characteristics of helices) were excluded given that their intensities are the most affected by the conformational averaging, as random coil conformations contribute to their intensities [41]. It is important to bear in mind that the presence of the helical conformers and their high populations are established independently on the basis of CD spectroscopy, as well as conformational chemical shifts. The selected NOE cross-peaks were integrated and translated

into distance restraints, and the Φ and Ψ dihedral angle restraints were obtained using the TALOS-N webserver [97]. 100 structures were computed, from which 20 structures with the lowest target functions were minimized following the CYANA protocol. The structural statistics of the calculations are summarized in **Table 2.4** (see Results **Section 2.3.3**). These minimized structures were analyzed and examined with MOLMOL [98] and PyMOL [99].

2.2.8 NMR estimation of helix populations and interaction mapping

Helix populations were quantified from the $^{13}C\alpha$ or $^{1}H\alpha$ secondary chemical shift values taken from the 2D-NMR spectra and applying the formula for either the carbon or proton [89]:

$$\delta\Delta\, H\alpha = \delta H\alpha_{observed} - \delta H\alpha_{random\ coil}$$
$$\delta\Delta\, C\alpha = \delta C\alpha_{observed} - \delta C\alpha_{random\ coil} \qquad \textbf{(Equation 2.2)}$$

Δδ Hα and Δδ Cα values were averaged for all residues, and the resulting mean divided by the Δδ Hα or Δδ Cα value corresponding with the 100 % helix formation (-0.39 ppm for Hα and 2.6 for Cα) [100] and multiplied by 100. Random coil values were taken from Wishart et al. [101]. Errors are estimated to be about 3 % when populations are obtained from the ^{1}H chemical shifts, and about 7 % if obtained from the $^{13}C\alpha$ data [102].

To evaluate the interaction between peptide pairs and trios, a weighting factor Δδ w (which accounts for Hα and HN variations) was calculated from the chemical shifts of assigned resonances of the peptides when alone (free) and in pairs or trios (mixture):

$$\Delta\delta w = \sqrt{(\Delta\delta\ H\alpha_{interaction})^2 + (\Delta\delta\ HN_{interaction})^2} \qquad \textbf{(Equation 2.3)}$$

in which:

$$\Delta\delta\ H\alpha_{interaction} = \Delta\delta\ H\alpha_{free} - \Delta\delta\ H\alpha_{mixture}$$
$$\Delta\delta\ HN_{interaction} = \Delta\delta\ HN_{free} - \Delta\delta\ HN_{mixture} \qquad \textbf{(Equation 2.4)}$$

2.2.9 Computational modeling and NMR-driven docking

* Molecular modeling was performed by Dr. Miguel Mompeán, from the Faculty of Sciences and Technology, IRICA University of Castilla-La Mancha, Ciudad Real, Spain, employing the experimental data obtained in this PhD Thesis.

To detect possible helix-helix interactions, a set of simulated annealing calculations in the torsion angle space were performed using CYANA with default parameters [96]. Hexa- and dodeca-Gly linkers were used to favor spatial proximity between TACC-α2 and Ct-TACC-α4. The salt bridge between Asp921 (TACC-α4) and Lys827 (TACC-α2) previously inferred from SAXS data [77] was detected in some of the runs, and subsequently used as a restraint for docking studies. The docking of TACC-α2 and Ct-TACC-α4 included the XMAP-pCt peptide to match our experimental conditions, and was performed using the HADDOCK program [36]. Residues whose resonances show changes in the trio mixture were chosen as active in the docking interface. Following a rigid body-energy minimization procedure, the 200 structures with the lowest intermolecular energy values were selected for further semi-flexible docking and explicit water solvation. The structures with the best Haddock docking scores were selected as representative conformers.

2.3 Results

2.3.1 Peptide design

The peptides studied here were designed on the basis of the proposed model for the TACC and XMAP interaction [77]. As described in the study, the model indicates that the helix of XMAP-pCt rests near the groove between the helix α2 and the C-terminus of helix α4 of TACC3 (**Figure 2.3**). For this reason, the peptide sequences were designed to span the possible interacting regions; specifically, helix α2 and the C-terminus of helix α4 of TACC3, and the previously described C-terminus peptide from XMAP. The peptides should be long enough to be relevant in this analysis, but sufficiently short to allow a full and unambiguous ^{1}H NMR assignment and obtain adequate information in order to generate a 3D structure.

The peptide sequences listed in (**Table 2.1**) correspond with the residues Val823-Lys857 (TACC-α2), Met897-Ile930 (Ct-TACC-α4) and Ser2047-Lys2066 (XMAP-pCt) in the full-length protein or domain, respectively. The wild-type sequence of Ct-TACC-α4 (Ct-TACC-α4 wt) was studied together and compared with a mutant (Ct-TACC-α4 mut), in which two residues important for function and binding (Asp922 and Asp923) were replaced by alanines. An analysis of the sequences shows that all α-helical structures have certain amphipathic character in all peptides, as shown in **Figure 2.4**. Thus, hydrophobic residues are clustered in specific helix faces, suggesting the role of these surfaces in intramolecular oligomerization or coiled-coil interactions. Also, as mentioned before, the analysis of these sequences showed that all peptides are prone to form coiled-coils [78].

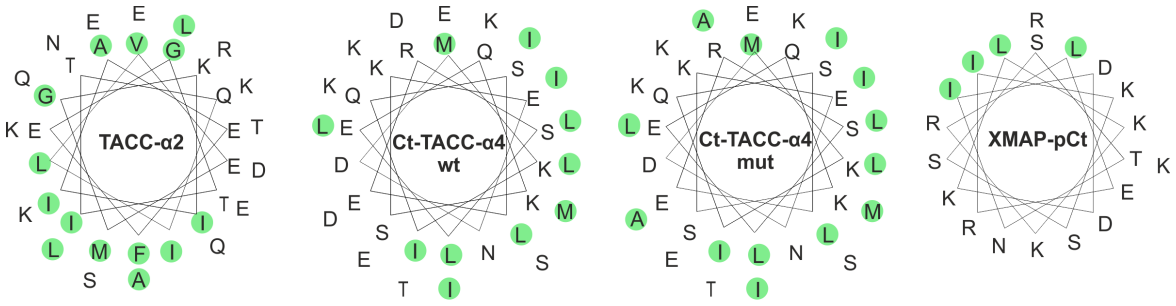


Figure 2.4. Helical wheel representations for the structures formed by the peptides derived from TACC and XMAP, showing their amphiphilic character in certain regions. Hydrophobic residues are marked in green.

2.3.2 Circular dichroism of isolated peptides

The structural behavior for each peptide was characterized by CD and varying different conditions, such as solvents, pH, and concentration (see **Figure 2.5**, **Figure 2.6** and **Figure 2.7**), so every possible combination can be analyzed. As mentioned before, TFE and DPC were selected as cosolvents to mimic cellular conditions.

As shown in **Table 2.3** (see **Section 2.3.3**), the CD spectra indicated that all peptides adopt mainly random conformations in H_2O; nevertheless, in the presence of TFE and DPC micelles they strongly tend to form helices (indicated by the characteristic minima at 208 and 222 nm; **Figure 2.5**). However, as the ratios of the signals found at 222 nm *versus* 208 nm were below 1.1 [103], no tendencies for coiled-coil motifs in the isolated peptides could be observed. Two different methods were used to evaluate the helix population from the CD data (see Method's **Section 2.2.3**). Upon comparison, the results obtained by the deconvolution analysis generally yield higher helical populations than the Rohl & Baldwin approach in the presence of cosolvents. In contrast, the Rohl & Baldwin method gives higher helical populations in aqueous solutions. In TFE and DPC at pH 5.5, the helical populations are high for TACC-α2, Ct-TACC-α4 wt and Ct-TACC-α4 mut (60-70 %), and moderate for XMAP-pCt (25 %). Since the charged state of the titratable side-chains, which can affect their contribution to the stability of the peptide structure, depends on the pH value, we decided to examine the influence of pH variation. Hence, taking into account the peptide's isoelectric

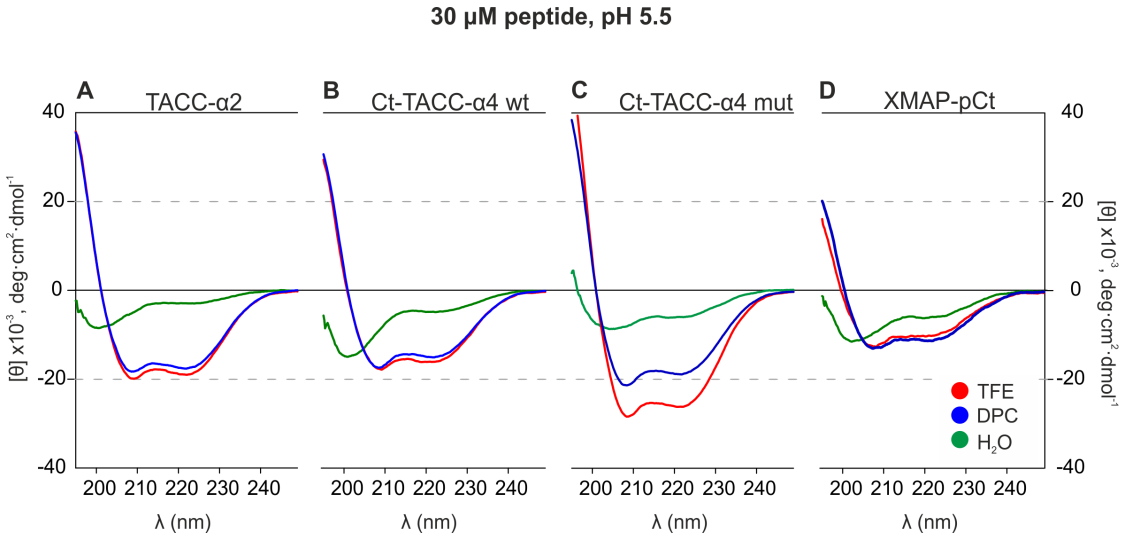


Figure 2.5. Far UV-CD spectra of individual peptides in different solvents. 30 µM TACC-α2 (A), Ct-TACC-α4 wt (B), Ct-TACC-α4 mut (C) and XMAP-pCt (D), pH 5.5. Solvents: 30 % TFE (red), 20 mM DPC (blue), H_2O (green).

point, we recorded CD spectra of the peptides in the presence of micelles at pH 3.0 and 10.0 (**Figure 2.6**).

At pH 10.0, the helix conformation seems to be destabilized in TACC-α2, Ct-TACC-α4 wt and Ct-TACC-α4 mut peptides. However, XMAP-pCt did not show important differences over all the pH range studied, which could be attributed to a compensation between the loss of favorable electrostatic interactions at pH 3.0, where Asp and Glu side chains are predominantly neutral, and a reduced unfavorable electrostatic interactions among Lys and Arg side chains at pH 10.0 (where a fraction of Lys side chains will be neutral). Interestingly, under all the conditions studied here, Ct-TACC-α4 mut populates more helical conformations than the wild type sequence despite the loss of favorable salt bridge interactions due to the substitution of the aspartics by alanines. This can be explained by the positive contribution of the alanine residues to helix formation [104] and the elimination of the charge-charge repulsion between two consecutive aspartic acids in the sequence [105].

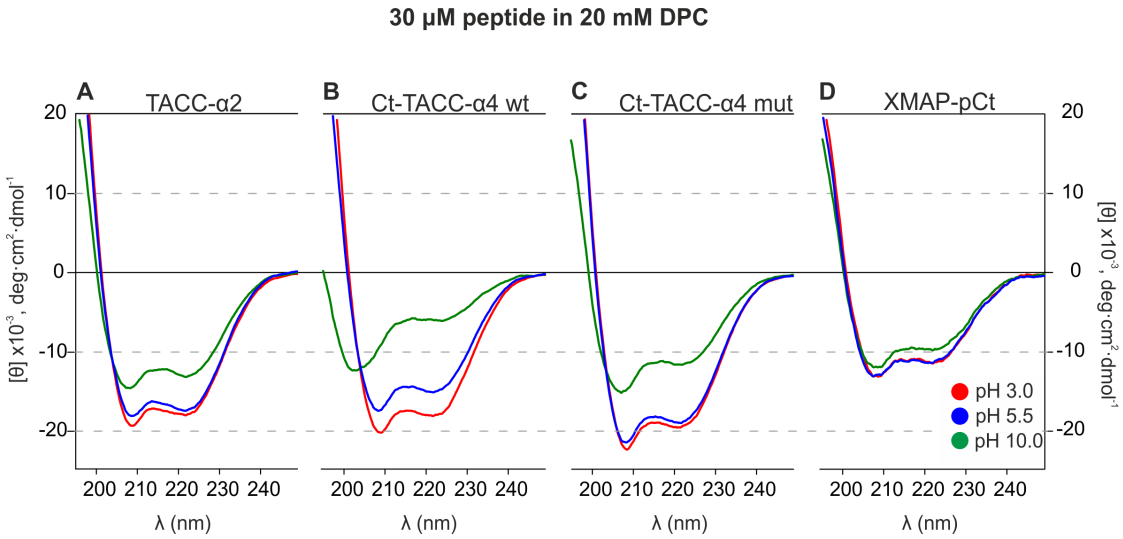


Figure 2.6. Far UV-CD spectra of individual peptides in different pH conditions. 30 µM TACC-α2 (A), Ct-TACC-α4 wt (B), Ct-TACC-α4 mut (C) and XMAP-pCt (D) in 20 mM DPC micelles. Conditions: pH 3.0 (red), pH 5.5 (blue) and pH 10.0 (green).

It was shown previously that peptides from many centrosomal proteins tend to aggregate [78, 106]. According to the CD experiments; in which the averaged structural behavior of the peptides is observed, the analysis of individual peptides over the concentration range from 30 to 60 µM at pH 5.5 in H_2O revealed that all peptides adopt mainly random coil conformations and no cases of concentration dependent self-associative processes were observed (**Figure 2.7**), at least under the experimental conditions of this study.

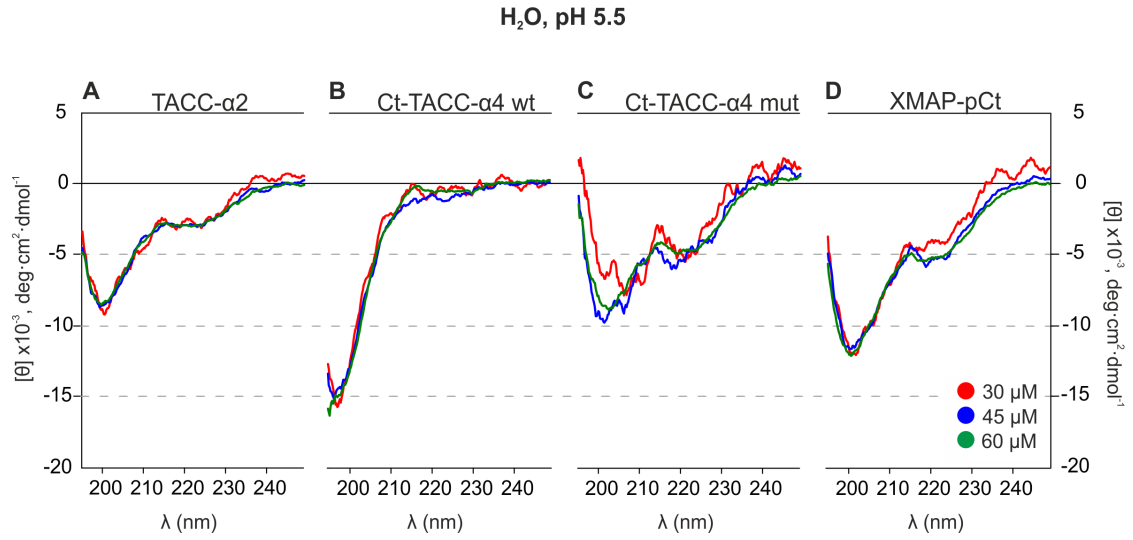


Figure 2.7. Far UV-CD spectra of individual peptides in different concentrations. TACC-α2 (A), Ct-TACC-α4 wt (B), Ct-TACC-α4 mut (C) and XMAP-pCt (D) in H_2O, pH 5.5. Peptide concentrations: 30 μM (red), 45 μM (blue) and 60 μM (green).

One explanation for the absence of noticeable coiled-coil structures might be that they are already present, but at a low population and in equilibrium with monomeric helices. This prevents their identification by CD. Likewise, it might be plausible that longer sequences are needed in order to form the coiled-coil structure and the peptides did not span a wide enough sequence. Finally, the presence of intermolecular interactions might be necessary to promote a structural change to drive coiled-coil formation.

2.3.3 NMR characterization of isolated peptides

In order to get more detailed structural information, TACC-α2, Ct-TACC-α4 wt, Ct-TACC-α4 mut and XMAP-pCt peptides were characterized by NMR in H_2O, 30 % TFE and DPC micelles. The NMR spectra in water were recorded at 5 ºC given that the population of structured conformers decreases at higher temperatures. However, because of the excessive broadness of the NMR lines at low temperatures in the presence of TFE and DPC cosolvents, the NMR spectra in these cosolvents were acquired at 25 ºC or 40 ºC (**Table 2.2**). Ct-TACC-α4 mut was poorly soluble in H_2O. For this reason, its NMR spectra were very poor and only the ^1H assignments were achieved. The NMR spectra for this peptide in the presence of TFE and micelles, and for the other three peptides within all solvent conditions were completely assigned, and the chemical shifts are reported in **Appendix A**. Negative Δδ Hα and positive Δδ Cα values (**Figure 2.8**) were observed in almost all residues, indicating the formation of helical structures in agreement with the CD data. The helix populations calculated from the

^{1}H and ^{13}C chemical shifts (**Table 2.3**) are congruent and similar to those derived from the CD data, and particularly to those obtained with Rohl & Baldwin's method [90]. However, the helix percentage obtained by NMR for the peptide XMAP-pCt is lower than the one previously reported [77]. The distinct peptide design (different N- and C-terminus groups), buffer and pH conditions may account for this difference.

Table 2.3. Percentages of global helical populations for TACC-α2, Ct-TACC-α4 wt, Ct-TACC-α4 mut, and XMAP-pCt peptides from CD (5 ºC) and NMR (5 ºC in H_2O, and 25 ºC in TFE and DPC), data under different conditions.

			Helix population (%)			
					NMR	
Peptide	**Solvent**	**pH**	**CD [a]**	**CD [b]**	**Δδ Hα**	**Δδ Cα**
TACC-α2	H_2O	5.5	8	12	23	20
	TFE	5.5	64	51	57	64
	DPC	5.5	61	48	52	67
	DPC	3	63	49	n.d.	n.d.
	DPC	10	38	37	n.d.	n.d.
Ct-TACC-α4 wt	H_2O	5.5	8	17	12	20
	TFE	5.5	59	44	33	49
	DPC	5.5	57	42	46	51
	DPC	3	67	59	n.d.	n.d.
	DPC	10	11	20	n.d.	n.d.
Ct-TACC-α4 mut	H_2O	5.5	13	20	23	- - [c]
	TFE	5.5	94	69	38	62
	DPC	5.5	69	51	37[d]	31[d]
	DPC	3	73	52	n.d.	n.d.
	DPC	10	28	33	n.d.	n.d.
XMAP-pCt	H_2O	5.5	8	21	23	21
	TFE	5.5	24	32	40	24
	DPC	5.5	30	35	40	63
	DPC	3	30	35	n.d.	n.d.
	DPC	10	24	31	n.d.	n.d.

[a] Data obtained by the Dichroweb server, using the K2D Analysis programme [91].
[b] Data from Rohl & Baldwin's method [90].
[c] Due to its low solubility, ^{13}Cδ data could not be obtained.
[d] Temperature was increased to 40 ºC due to the broad NMR signals at lower temperatures.
n.d., not determined.

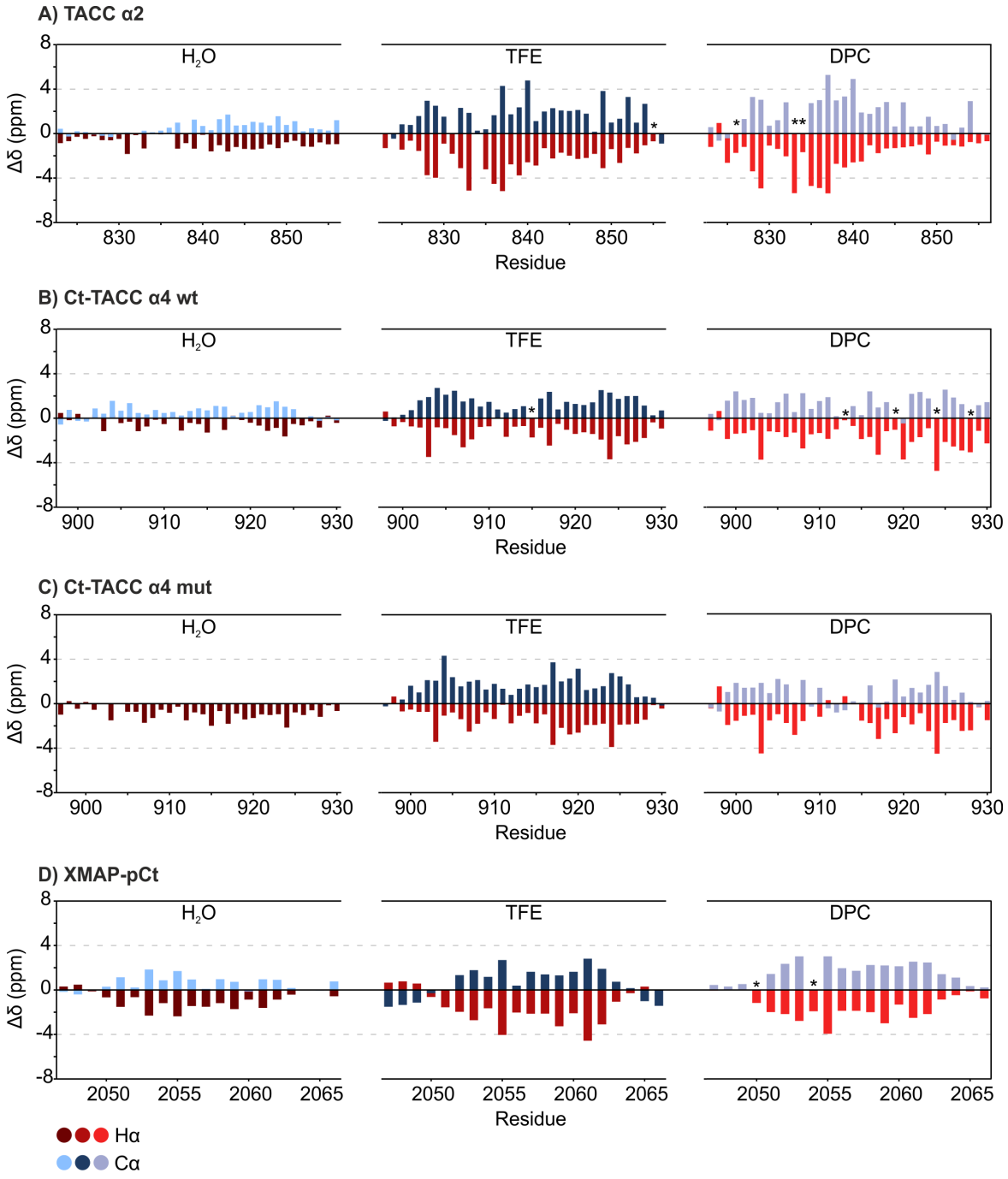


Figure 2.8. Hα and Cα secondary chemical shifts observed for the peptides, pH 5.5. Δδ Hα Values were multiplied by 10 (for clarity). TACC-α2 (A), Ct-TACC-α4 wt (B), Ct-TACC-α4 mut (C), and XMAP-pCt (D) peptides in H_2O (left), TFE (center) and DPC (right). Non assigned resonances are indicated with *. The low solubility of Ct-TACC-α4 wt in H_2O, precluded the measurement of the ^{13}C NMR data at natural abundance (C, left panel).

Once analyzed the secondary structure through the chemical shift deviations, the calculation of the preferred conformations based on NOE signals of the peptides in TFE and DPC was performed, matching with the conditions where the helix populations are higher (**Table 2.3**). In the presence of TFE and DPC micelles, many NOEs typical of helical structures were observed (sequential HN_i-HN_{i+1}, and non-sequential $H\alpha_i$-HN_{i+3} and $H\alpha_i$-$H\beta_{i+3}$), as shown in **Figure 2.9**.

The calculated structures can be seen in **Figure 2.10** and **Figure 2.11**. In TFE, the four peptides form well-defined helices, with low backbone RMSD values (mean value of 1.4, 1.1, 1.0 and 1.1 Å for TACC-α2, Ct-TACC-α4 wt, Ct-TACC-α4 mut and XMAP-pCt, respectively. See **Table 2.4**). The helices of the long TACC peptides show a certain degree of freedom at the N- and C-termini. Interestingly, the structural comparison of these isolated helices in TFE with previously reported theoretical models of the full domain [77], reveals that they are similar, with backbone RMSD values of 2.4 and 1.3 Å for TACC-α2 and Ct-TACC-α4, respectively. In the case of Ct-TACC-α4 mut, the Asp mutations did not severely affect the conformation, as the RMSD value between the Ct-TACC-α4 mut and Ct-TACC-α4 wt structure is 1.6 Å. Whereas these structures represent the preferred conformers in these conditions, their total population is less than 100 %, as indicated by the CD spectra and NMR conformational chemical shifts. Therefore, it is important to stress that other conformers, which are conformationally diverse, are also present.

In DPC, and due to the broadening of the NMR lines, only the structures of TACC-α2 and XMAP-pCt could be calculated. In these cases, the preferred structures are also well defined with low backbone RMSD values (mean value of 1.4 and 0.2 Å for TACC-α2 and XMAP-pCt, respectively) (See **Table 2.4**). In these last two peptides, it was possible to compare the structures determined in TFE and DPC. Small backbone RMSDs (excluding the N- and C-termini residues) are observed: 1.8 and 1.5 Å for TACC-α2 (in TFE vs. DPC) and XMAP-pCt (in TFE vs. DPC), respectively. These low RMSD values indicate that the backbone conformations are similar in the two solvent systems.

In TFE and DPC cosolvents, the preformed helical conformations get stabilized, as demonstrated by CD and NMR. These structural observations are in accordance with the SAXS model previously described for the whole protein [77]. Therefore, we can conclude that the designed peptides are good models to mimic the entire protein system and are well adapted in order to define experimentally the geometry of the TACC protein's binding site.

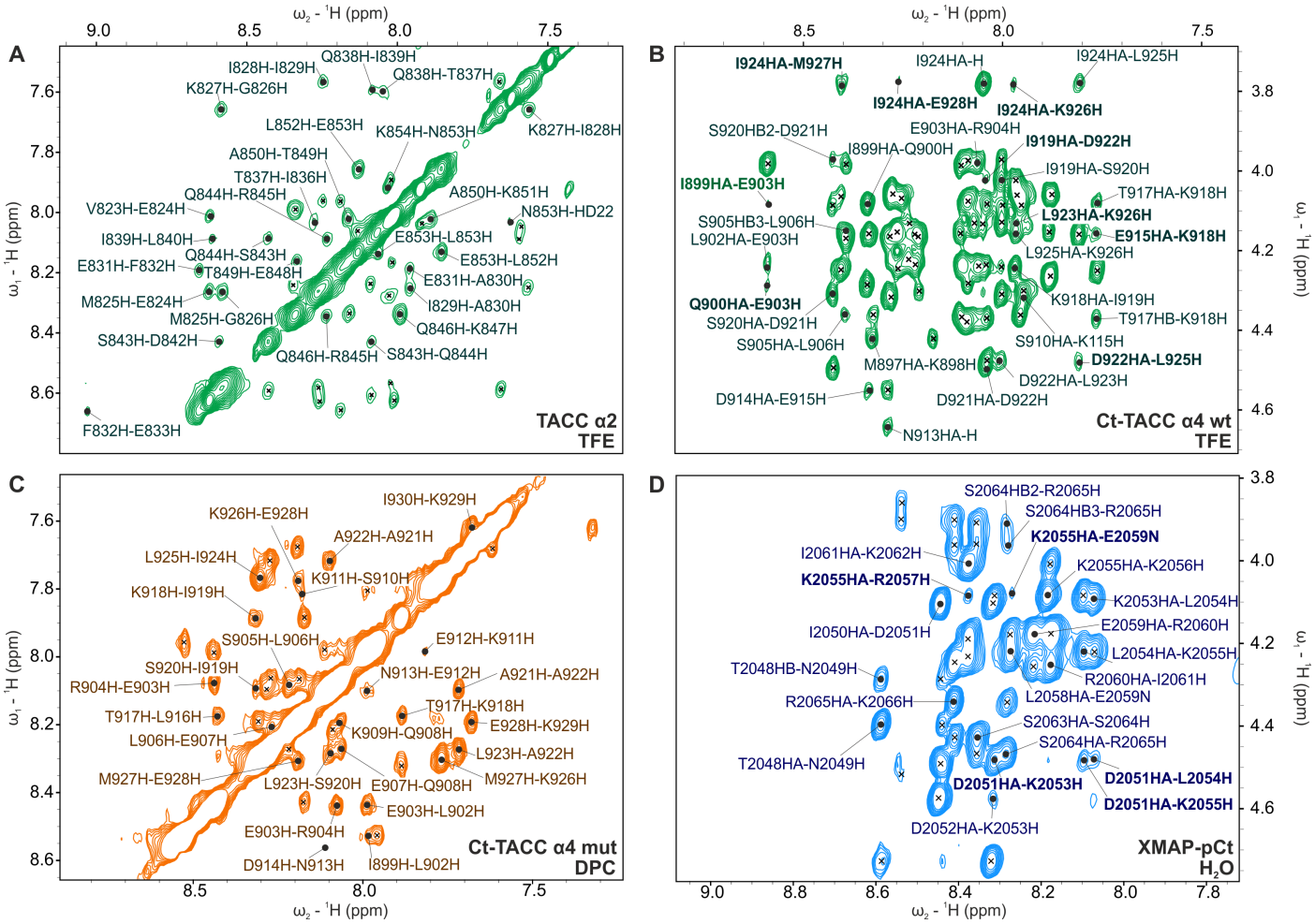


Figure 2.9. Regions of the ^{1}H-^{1}H NOESY spectra. HN-HN region of TACC-α2 in 30 % TFE, pH 5.5 (A), Hα-HN region of Ct-TACC-α4 in TFE, pH 5.5 (B), HN-HN region of Ct-TACC-α4 mut in 20 mM DPC, pH 5.5 (C), Hα-HN region of XMAP-pCt in H_2O and pH 5.5 (D). NOE signals are identified by labeling the pair of protons accounting for the NOE effect. Non-sequential NOEs are in bold.

Table 2.4. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for TACC-α2, Ct-TACC-α4 wt, Ct TACC-α4 mut and XMAP-pCt in 30 % TFE and 20 mM DPC

	30 % TFE				20 mM DPC	
	TACC-α2	**Ct-TACC-α4 wt**	**Ct-TACC-α4 mut**	**XMAP-pCt**	**TACC-α2**	**XMAP-pCt**
Upper Limit distance restraints (from NOEs)						
Total	154	161	89	128	158	63
Short-range |i-j| ≤ 1	30	109	26	43	43	39
Medium-range 1 < |i-j| < 5	124	52	63	85	115	14
Long-range |i-j| ≥ 5	0	0	0	0	0	0
φ/ ψ Dihedral angle constraints (from chemical shifts)	64	64	64	35	60	35
Average CYANA target function value	0.05 ± 0.1	0.34 ± 0.1	0.03 ± 0.1	0.36 ± 0.1	0.58 ± 0.1	0.06 ± 0.1
Averaged maximum violation per structure						
Distance (Å)	0.01 ± 0.1	0.30 ± 0.1	0.01 ± 0.1	0.03 ± 0.1	0.03 ± 0.1	0.01 ± 0.1
Dihedral angle (º)	0.40 ± 0.1	0.50 ± 0.1	0.30 ± 0.1	1.40 ± 0.4	2.00 ± 0.1	0.50 ± 0.1
Pairwise RMSD (Å)						
Backbone atoms	2.4 ± 1.3	1.5 ± 0.8	1.4 ± 0.7	0.9 ± 0.4	1.4 ± 0.7	0.5 ± 0.2
All heavy atoms	3.2 ± 1.3	2.3 ± 0.7	2.1 ± 0.6	1.9 ± 0.4	2.1 ± 0.6	1.6 ± 0.2
Ramachandran plot (%)						
Favorable	95.8	98.9	99.7	88.1	96.9	99.7
Allowed	3.3	1.1	0.3	5.8	3.1	0.3
Outlier	0.9	0	0	6.1	0	0

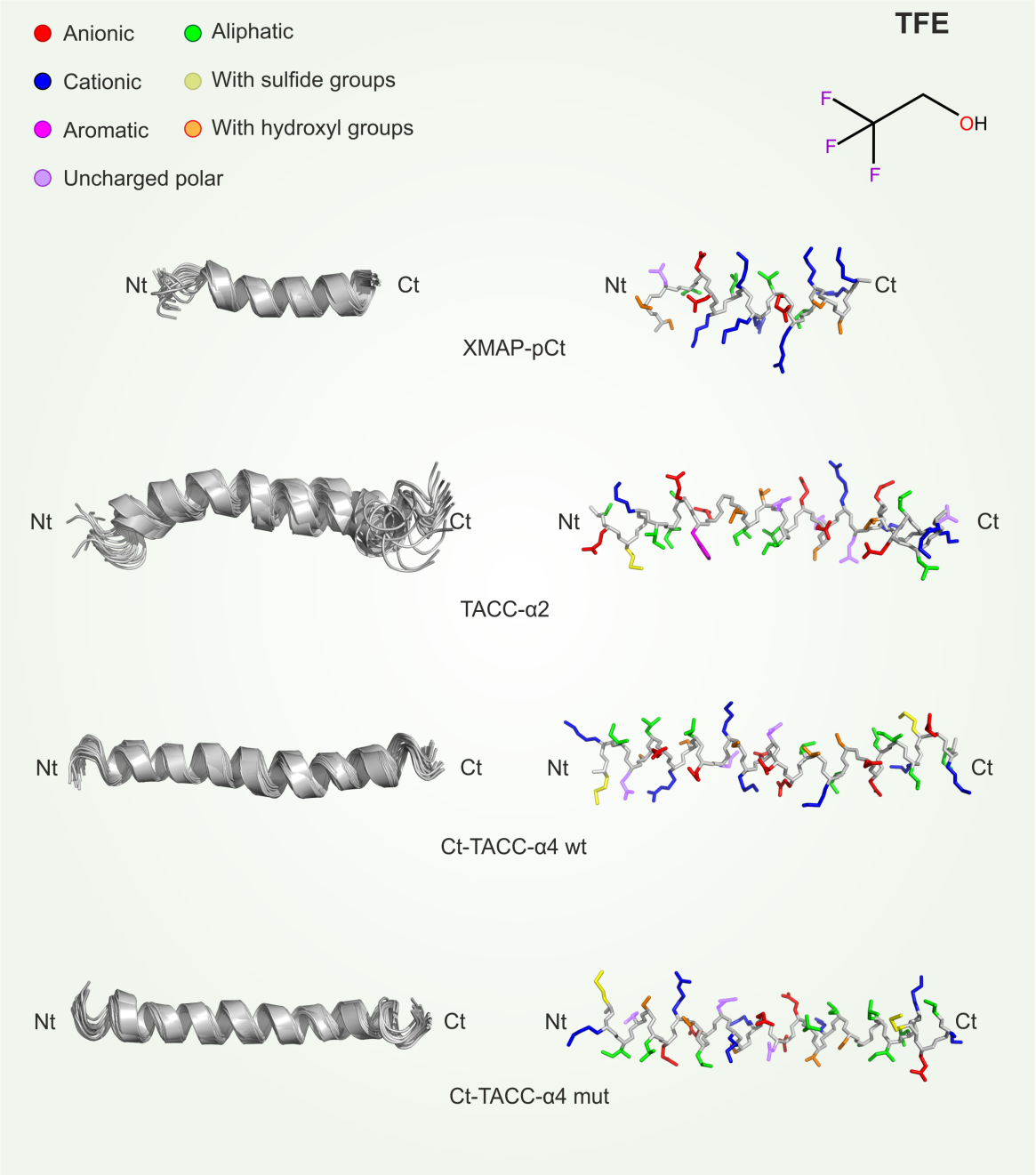


Figure 2.10. Ensemble of the 20 structures with the lowest target function values of TACC-α2, Ct-TACC-α4 wt, Ct-TACC-α4 mut and XMAP-pCt in 30 % TFE, 25 ºC, pH 5.5. The superposition of the backbone of the 20 lowest target function structures in each family is represented as ribbon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Anionic: red, cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

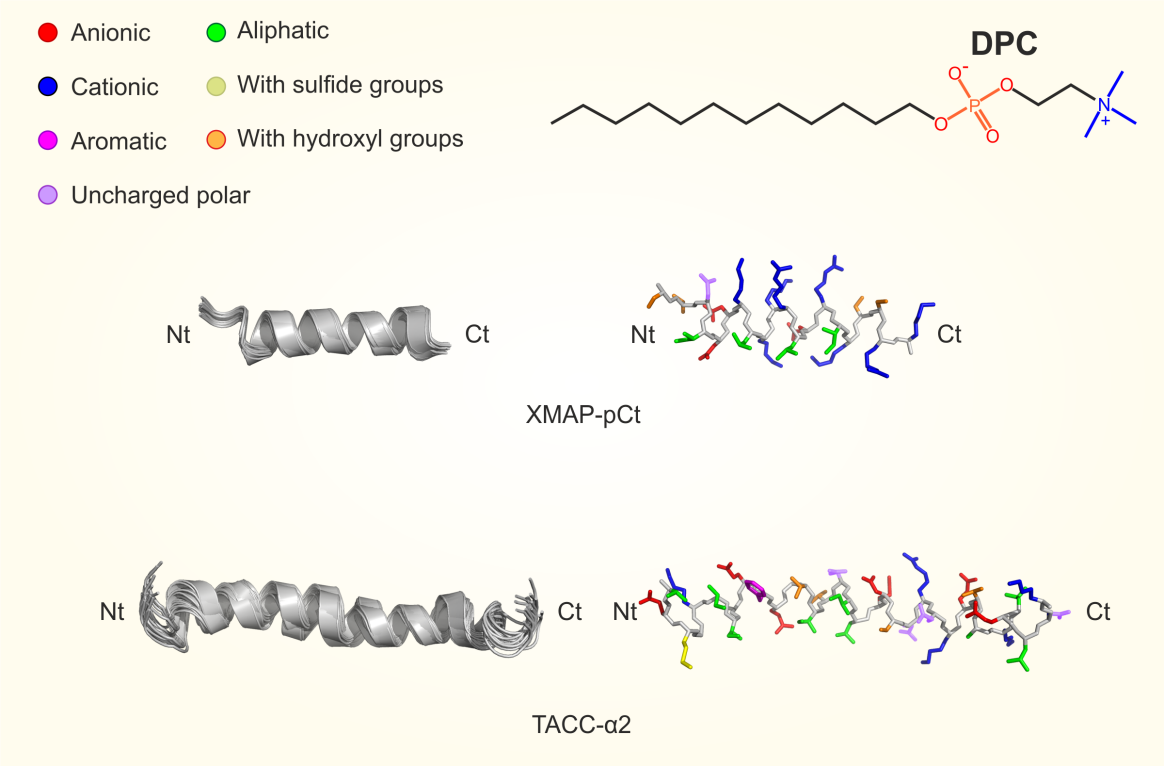


Figure 2.11. Ensemble of the 20 structures with the lowest target function values of TACC-α2 and XMAP-pCt in 20 mM DPC, 25 ºC, pH 5.5. The superposition of the backbone of the 20 lowest target function structures in each family is represented as ribbon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid. Anionic: red, cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

2.3.4 CD and NMR characterization of peptide mixtures

Following the previous methodology and to find out if certain regions of the TACC centrosomal protein are prone to interact, CD spectra for all possible combinations, particularly five pairs and two trios, were recorded in the same conditions as the isolated peptides. The interaction was tested comparing the experimental and theoretical (sum of individual contributions) CD data of the different mixtures. No significant differences were found for the peptide pairs (TACC-α2 + Ct-TACC-α4 wt; Ct-TACC-α2 + TACC-α4 mut; TACC-α2 + XMAP-pCt, Ct-TACC-α4 wt + XMAP-pCt and Ct-TACC-α4 mut + XMAP-pCt) in H_2O, TFE or DPC (**Figure 2.12** and **Figure 2.13**, see further pages). This indicates that the interactions, if present, are not detectable by CD. Only in the trio mixtures at 1:1:1 ratio (TACC-α2 + Ct-TACC-α4 wt + XMAP-pCt and TACC-α2 + Ct-TACC-α4 mut + XMAP-pCt), the experimental CD spectra showed differences with respect to the theoretical sum

of the isolated peptides, except in pure water solutions (**Figure 2.14 A**, see further pages). Although the observed differences are small, they are solid indications of interaction among the components. In both trios in TFE, and in the one containing Ct-TACC-α4 mut in DPC at pH 10.0, the experimental CD spectra revealed an increased α-helix population in the mixture when compared with the free components; while in the trio TACC-α2 + Ct-TACC-α4 wt + XMAP-pCt in DPC, pH 3.0 and 5.5, the experimental CD spectra show a diminished α-helix population in DPC.

Finally, the trio containing the Ct-TACC-α4 mut presents a different response with pH, and no variations were observed at acidic pH. This is interesting as it shows a different behavior depending on the presence of the Asp residues at positions 921 and 922, at the reported physiological pH of the centrosome [107].

As mentioned above, the SAXS model of XTACC3-XMAP215 is a 1:2 complex. For this reason, we tested by CD the trio samples at 1:1:2 instead of 1:1:1 peptide ratio. At pH 5.5 in H_2O, the CD data in the 1:1:2 samples did not show changes (**Figure 2.15**, see further pages) indicating, like in the 1:1:1 preparations, that no interaction is present. In TFE, the difference between the experimental measurements and the theoretical sum of the individual peptide signals has the opposite tendency than that observed in the 1:1:1 ratio. This indicates that the interaction is not favored in these conditions, in which one peptide (XMAP-pCt) is in molecular excess. Finally, as commented before, the selected sequences have shown low solubility and other peptide ratios in the mixture different to 1:1:1 tend to precipitate with time. For all these reasons, with the concentration requirements for NMR experiments, forced us to use only competent 1:1:1 mixtures for the NMR study. Finally, as commented before, the selected sequences have shown low solubility and other peptide ratios in the mixture different to 1:1:1 tend to precipitate with time. These liminations together with the concentration requirements for NMR experiments, forced us to use only 1:1:1 mixtures for the NMR study.

Thus, to strengthen the conclusions from CD results, NMR experiments were performed for the soluble trio TACC-α2 + Ct-TACC-α4 wt + XMAP-pCt at a 1:1:1 molar ratio in two different solvents: in H_2O at 5 ºC and 30 % TFE at 25 ºC (**Figure 2.16**, see further pages). The other trio was not soluble at the NMR concentrations in both H_2O and TFE; while in DPC, the lines were excessively broad and therefore not possible to assign the spectrum.

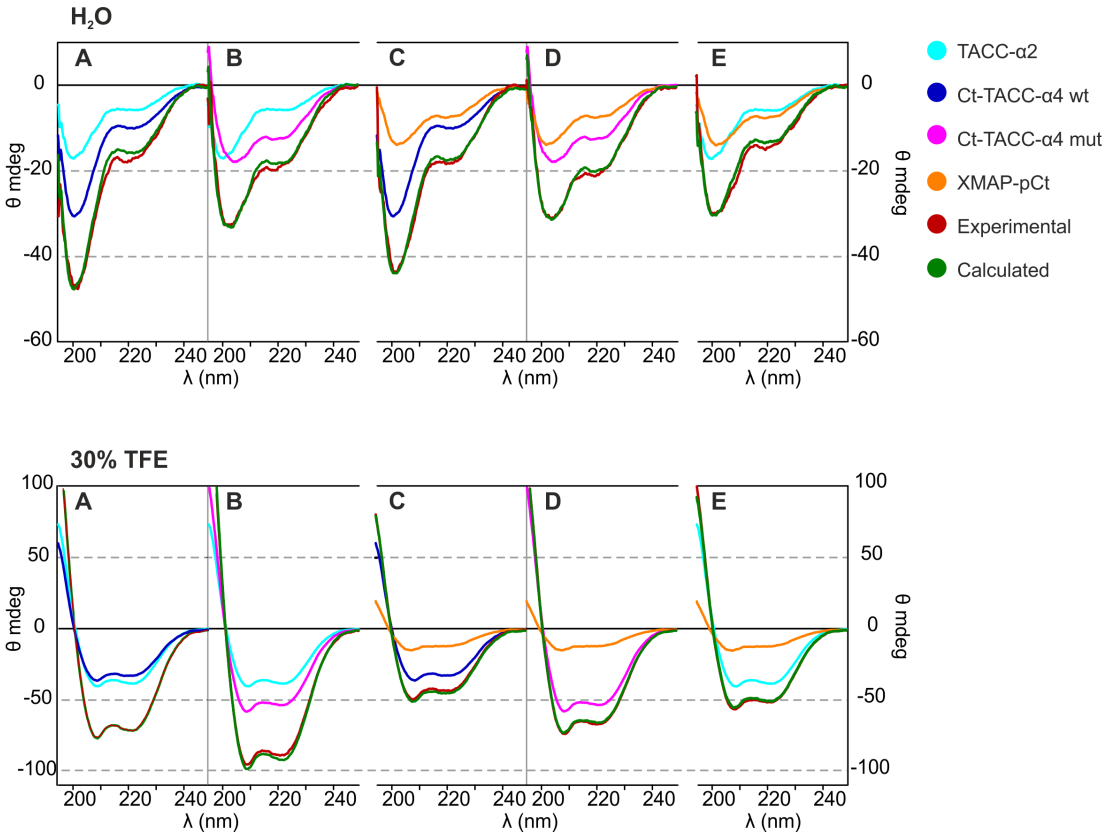


Figure 2.12. Far UV-CD (θ, mdeg) data of pair peptide mixtures in H_2O and 30 % TFE, 5 ºC, pH 5.5. TACC-α2 + Ct-TACC-α4 wt (A), TACC-α2 + Ct-TACC-α4 mut (B), Ct-TACC-α4 wt + XMAP-pCt (C), Ct-TACC-α4 mut + XMAP-pCt (D), TACC-α2 + XMAP-pCt (E). CD data of individual and isolated peptides are represented as: TACC-α2 (cyan), Ct-TACC-α4 wt (dark blue), Ct-TACC-α4 mut (magenta) and XMAP-pCt (orange). Theoretical calculated CD values (as the sum of the individual contributions) of the different mixtures are represented in green and the experimental CD values of the pair mixtures in red.

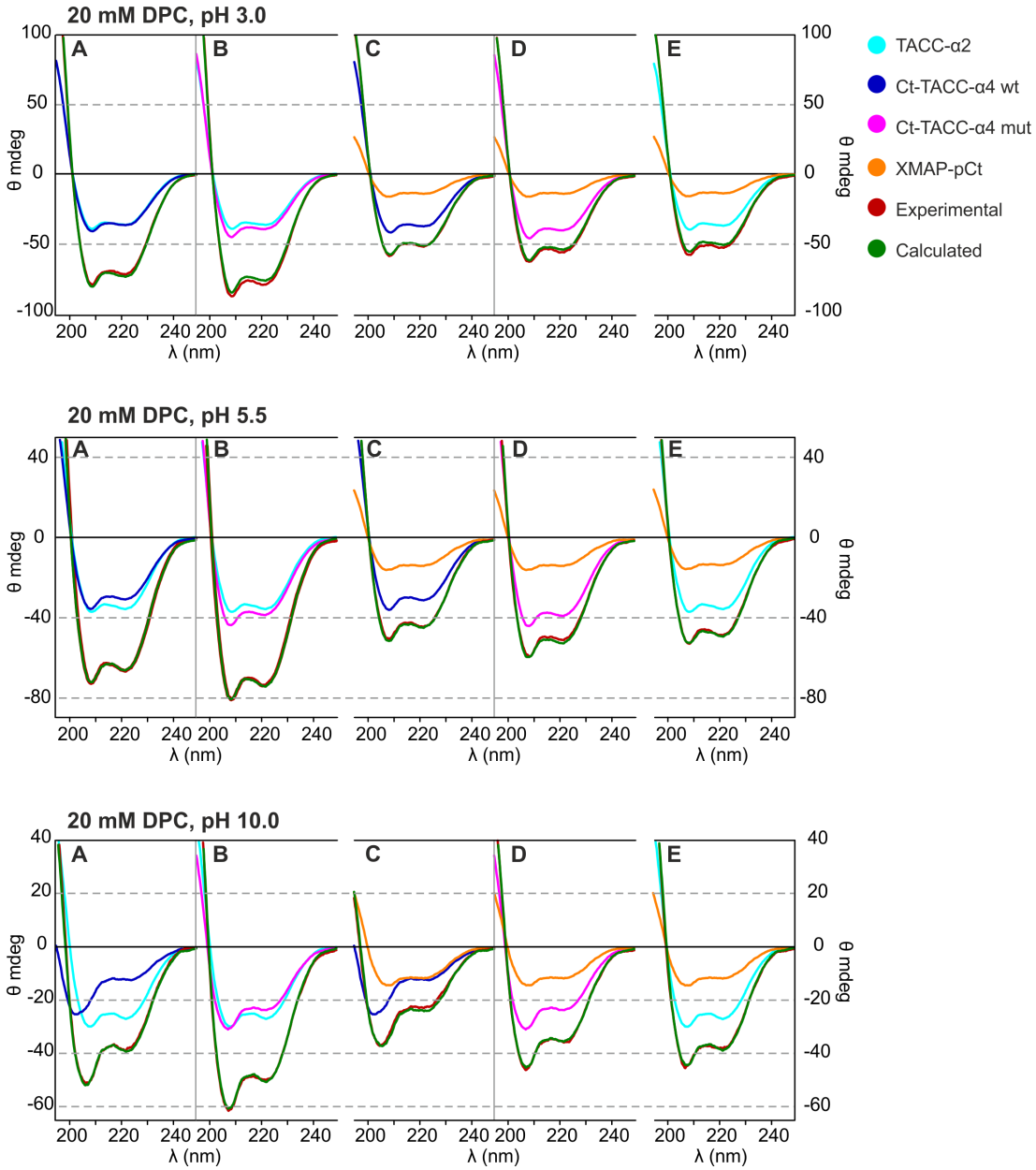


Figure 2.13. Far UV-CD (θ, mdeg) data of pair peptide mixtures in 20 mM DPC, 5 ºC, pH 5.5, 3.0 and 10.0. TACC-α2 + Ct-TACC-α4 wt (A), TACC-α2 + Ct-TACC-α4 mut (B), Ct-TACC-α4 wt + XMAP-pCt (C), Ct-TACC-α4 mut + XMAP-pCt (D), TACC-α2 + XMAP-pCt (E). CD data of individual and isolated peptides are represented as: TACC-α2 (cyan), Ct-TACC-α4 wt (dark blue), Ct-TACC-α4 mut (magenta) and XMAP-pCt (orange). Theoretical calculated CD values (as the sum of the individual contributions) of the different mixtures are represented in green and the experimental CD values of the pair mixtures in red.

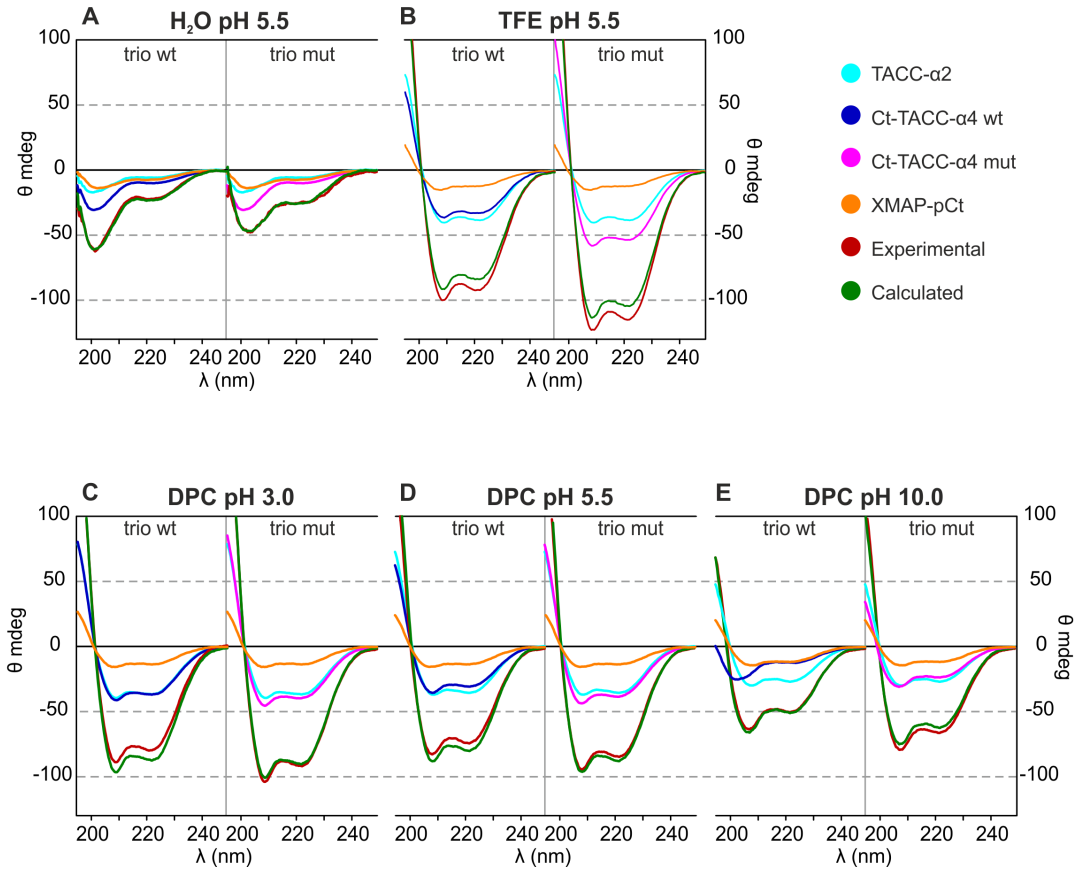


Figure 2.14. Far UV-CD of trio peptide mixtures in different conditions. H_2O, 5 ºC, pH 5.5 (A), TFE, 25 ºC, pH 5.5 (B), DPC, 25 ºC, pH 3.0 (C), DPC, 25 ºC, pH 5.5 (D), DPC, 25 ºC, pH 10.0 (E). Left panels correspond to TACC-α2 + Ct-TACC-α4 wt + XMAP-pCt and right panels to TACC-α2 + Ct-TACC-α4 mut + XMAP-pCt mixtures. CD data of individual and isolated peptides are represented as: TACC-α2 (cyan), Ct-TACC-α4 wt (dark blue), Ct-TACC-α4 mut (magenta) and XMAP-pCt (orange). Theoretical calculated CD values (as the sum of the individual contributions) of the different mixtures are represented in green and the experimental CD values of the trio mixtures in red.

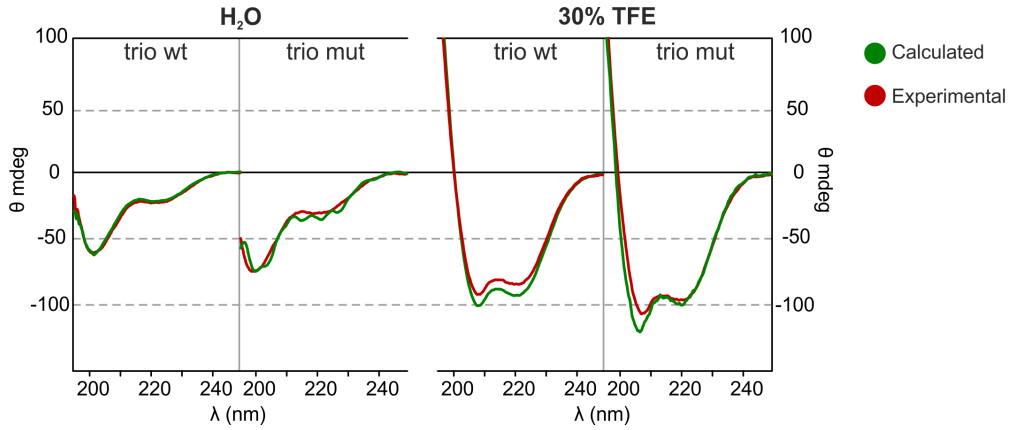


Figure 2.15. Far UV-CD of trio peptide mixtures in different conditions and 1:1:2 ratio. H_2O, 5 ºC and TFE, 25 ºC; pH 5.5. Left panels correspond to TACC-α2 + Ct-TACC-α4 wt + XMAP-pCt and right panels to TACC-α2 + Ct-TACC-α4 mut + XMAP-pCt mixtures all in ratio 1:1:2 (TACC-α2: TACC-α4 (wt or mut): XMAP-pCt). Theoretical calculated CD values (as the sum of the individual contributions) of the different mixtures are represented in green and the experimental CD values of the trio mixtures in red.

In agreement with the CD data, the spectra obtained for the mixture in H_2O did not show significant differences compared with the sum of the three NMR spectra of the individual peptides (**Figure 2.16**). In fact, all Hα differences are very small, less than 0.01 ppm. These measured differences (maximum difference of 0.009 ppm for Hα of Asn855 in TACC-α2) are in the range of the estimated experimental error (about ± 0.01 ppm) and are randomly distributed throughout the peptide sequence. However, the situation was different in TFE. All backbone signals could be assigned unequivocally in the mixture (**Table A.13**), and meaningful differences that are not uniformly distributed spanning the peptide sequences were found. This can be seen in **Figure 2.17**, which shows the weighted NMR chemical shift differences for the three peptides. For TACC-α2 and Ct-TACC-α4 wt, differences larger than the mean values are found in specific locations. In TACC-α2 most of them are concentrated at the N-terminus, Gly826 and Ile829-Thr835. On the contrary, in Ct-TACC-α4 wt, the largest changes are seen for Leu906 and Glu907 and at the C-terminus (Glu915-Met927). For XMAP-pCt, only the N-terminus (Ser2047-Asn2049) region was significantly affected. This is compatible with an antiparallel arrangement of TACC-α2 and Ct-TACC-α4 wt, and the antiparallel position of XMAP-pCt with respect to TACC-α4. The changes observed in the NMR spectra of the trios in comparison with the fingerprint of the individual peptides allows to identify exactly which residues are displaced and consequently, determine the ones highly and slightly involved in the interaction. Also, from the experimental point of view, it is important to highlight that NMR spectroscopy is capable and suitable for detecting very weak interactions in short peptides with atomic resolution.

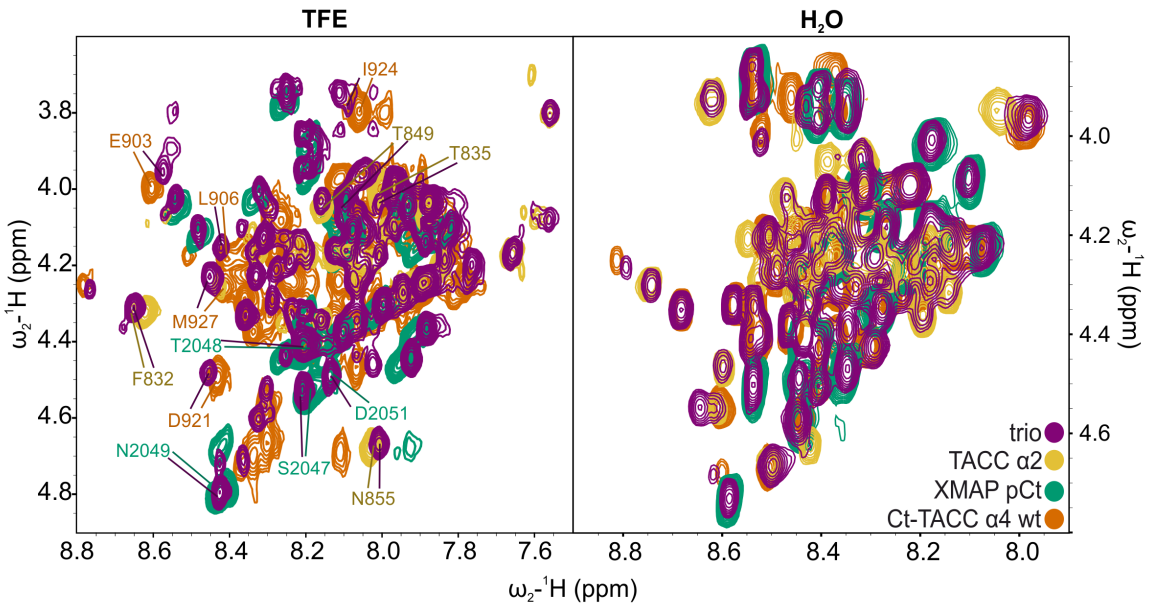


Figure 2.16. Superposition of the finger print region of the TOCSY NMR spectra of the individual peptides and in mixture in 30 % TFE and H_2O. (TACC-α2: dark yellow; Ct-TACC-α4 wt: orange; XMAP-pCt: green; trio mixture: purple). Some shifted signals (HN-Hα) are labeled with the corresponding residue number and one letter code.

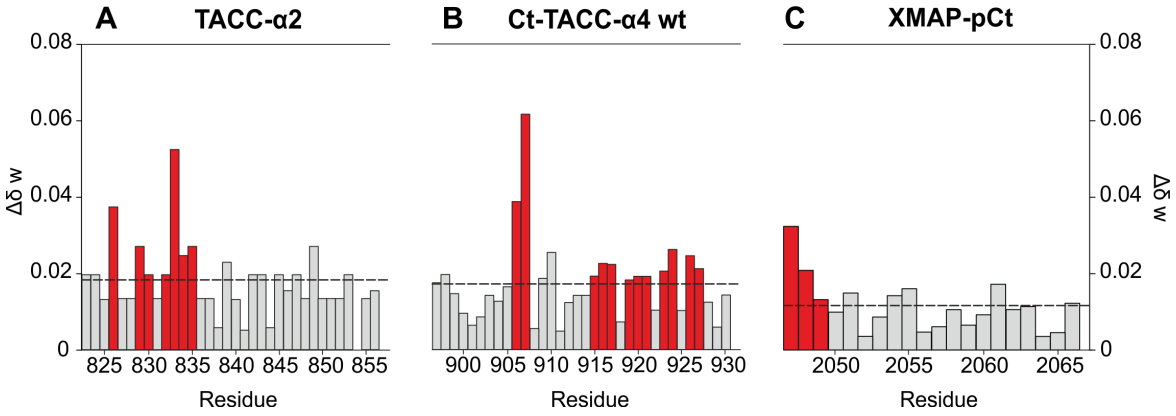


Figure 2.17. Weighted NMR chemical shift differences probe the interaction in the mixture, TACC-α2 (A), Ct-TACC-α4 wt (B), XMAP-pCt (C). Weighted values ($\Delta\delta$ w) are calculated as shown in (**Equation 2.3**), where $\Delta\delta H\alpha_{interaction}$ and $\Delta\delta HN_{interaction}$ are the chemical shift differences of the corresponding peptide protons alone and in the mixture conditions in 30 % TFE solution, 25 ºC. Mean values are represented by the dashed line. The regions of each peptide most affected by the interaction are depicted in red.

2.3.5 Molecular modeling of the TACC-α2 + Ct-TACC-α4 + XMAP-pCt assembly

In order to get a model of the interactions between TACC-α2 and Ct-TACC-α4, based on the experimental data, a computationally study was performed in collaboration with Dr. M. Mompeán (IRICA, University of Castilla-La Mancha). As it is explained in the experimental section, simulated annealing calculations in the torsion angle space were performed.

Either a hexa-Gly or a dodeca-Gly linker was used to join both TACC helices, thus mimicking the full-length context while increasing the local concentration to afford sampling of the helix-helix packing. Although molecular dynamics in the Cartesian space is usually the method of choice for this type of studies, its application was not suitable in the present system, as no experimental interactions were detected between TACC-α2 and Ct-TACC-α4 in binary mixtures. In this case, an exhaustive computational study of the two helix interactions would require them to be initially placed far away from each other in a non-interacting configuration, and long computationally demanding time-scales may be needed to capture the folding events.

Enhanced sampling methods are, on the other hand, a good approach to accelerate conformational changes [108]. However, the reliable application of this method would require the atomic resolution structure of the TACC coiled-coil folded state. Therefore, the simulated annealing in torsion angle space was chosen as a tool to speed up the calculations and explore the energy landscape of the TACC-α2-$(Gly)_n$-Ct-TACC-α4 (n = 6, 12) system at a reasonable computational cost. Many of the numerous different conformations generated following this approach presented the electrostatic interaction between Asp921 (Ct-TACC-α4) and Lys827 (TACC-α2) deduced from the SAXS model. Using this intermolecular contact as an unambiguous restraint, the most conservative NMR-driven protein-protein molecular docking was implemented, which included the third component XMAP-pCt. We calculated the complex with the 1.1.1 stoichiometry as the NMR chemical shift perturbation data used as input to drive the docking correspond to that specific situation. It was observed that in the absence of the Asp921-Lys827 contact, the docking solutions did not converge towards a defined ensemble that fulfills the experimental chemical shift perturbation data. Nevertheless, when the Asp921-Lys827 salt bridge is included in the calculations, the docking solutions can accommodate both the reported and the currently obtained experimental observations.

Similar docking solutions obtained for both TACC-α2-$(Gly)_6$-Ct-TACC-α4 + XMAP-pCt and TACC-α2-$(Gly)_{-12}$-Ct-TACC-α4 + XMAP-pCt suggest that the length of the $(Gly)_n$ linker does not affect the overall fold (**Figure 2.18**) when the above mentioned salt bridge

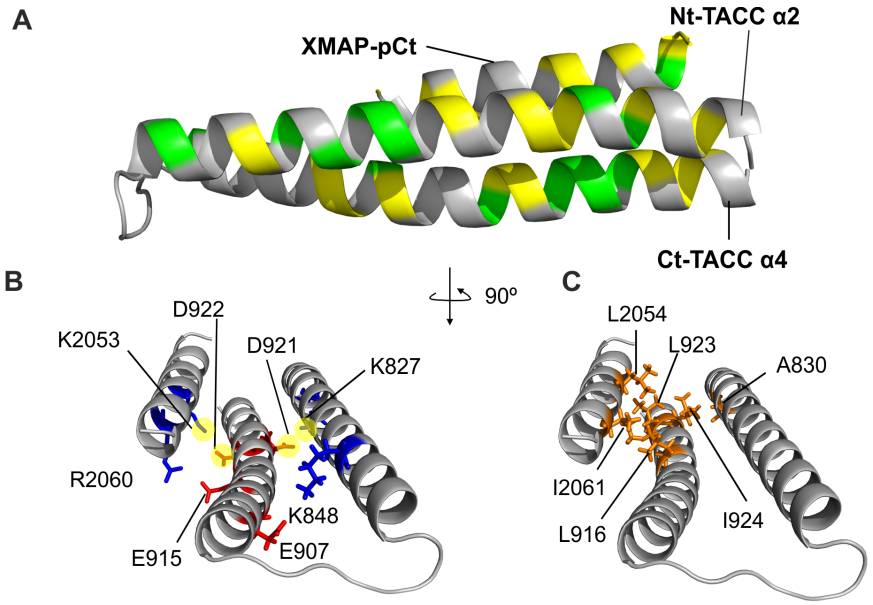


Figure 2.18. Structural 3D computed model of TACC-α2, Ct-TACC-α4 wt and XMAP-pCt, showing residues affected by the interaction. Residues strongly affected are in yellow and other residues with above average weighted values are displayed in green (A). Electrostatic (B) and hydrophobic (C) interactions between helices are represented and labeled. Arg and Lys side chains are in blue, Asp and Glu in red, and Leu, Ile and Ala in orange. Asp and Lys salt bridges are highlighted with yellow circles.
* Model generated by Dr. M. Mompeán.

is present. These results are in accordance with the NMR data, and unveil a number of interesting structural features that are in agreement with the resulting three-helix bundle. Therefore, TACC-α2 and Ct-TACC-α4 might pack with the fold reported from SAXS data through Asp921-Lys827 (Ct-TACC-α4-TACC-α2) interaction, in addition to a second favorable electrostatic interaction between Glu907 (Ct-TACC-α4) and Lys848 (TACC-α2), both of which are compatible with the hydrophobic packing of Ile924 (Ct-TACC-α4) and Ala830 (TACC-α2). As for the XMAP-pCt helix, the present results show that it likely establishes two sets of non-polar and ionic interactions with TACC-α4. In addition, key residues, namely Glu915-Leu916, Asp922-Leu923 (Ct-TACC-α4) and Lys2053-Leu2054, Arg2060-Ile2061 (XMAP-pCt) are oriented so that hydrophobic interactions between Leu916-Ile2061 and Leu923-Leu2054 may occur concomitantly with the formation of the two salt bridges between Glu915-Arg2060 and Asp922-Lys2053.

2.4 Discussion

The objective of this study is to describe the inherent structural properties of two regions of TACC (TACC-α2 and Ct-TACC-α4) and to highlight the interacting regions with XMAP with atomic information. Moreover, we would like to understand if the mutations Asp921 and Asp922 to A921 and A922 in TACC-α4 have any influence on the protein's structure. As well, it would be interesting to understand why these Asp residues are key for molecular recognition and test the proposed mechanism for the binding of the C-terminus domain of XMAP with the coiled-coil domain of TACC.

2.4.1 Short linear peptides are good models to mimic native protein structure

Here, our NMR and CD results show that the isolated peptides TACC-α2 and Ct-TACC-α4 tend to adopt helical conformations and their helical population is affected by the hydrophobicity and the pH of the media. In TFE and DPC cosolvents, they adopt preferred helical structures which points to the stabilization of the same preformed conformation in conditions that are analogous to the intracellular or membrane-like environments. This tendency is compatible with the secondary structure adopted by these sequences in the whole protein. In fact, the structural similarity of the model (**Figure 2.18**) with the conformations of the isolated helices (**Figure 2.10** and **Figure 2.11**) determined experimentally in this study reveals that the peptides are good models for the geometry of the TACC binding site. For the peptide with the Asp mutations at positions 921-922 (which are crucial residues for the interaction of the full protein with XMAP), [77] the structure was not affected and its helical tendency was even higher than the observed by the native sequence in most of the tested conditions. This is interesting, as it suggests that the lack of interaction of TACC carrying these mutations with XMAP is not related with the loss of the secondary structure, but with the specific interactions in which those Asp's side chains participate in the quaternary structure.

It is also well known that the centrosome as an organelle, carries a net negative charge with an isoelectric point around 3.1 [107]. In this regard, our CD data show that all peptides become more structured at lower pH values (at least in the presence of DPC micelles that mimics the hydrophobic and crowded cellular location), being these results compatible with the biological conditions. In the same way, similar results were described for other centrosomal proteins [78, 106].

2.4.2 Short linear peptides are good models to study protein interactions

Once the structure of the peptide components was known, our interest was focused on the pairwise molecular interactions. As commonly done [40, 82], here we have interpreted the changes in the population of the secondary structures as indicative of an interaction between the different peptides. Neither the CD nor NMR data could detect interactions between peptide pairs. These results suggest that either the expected interactions are transient and too weak to be seen in our conditions, or the interactions occur without changes in the secondary structure population of the peptides or, more likely, no interaction is present under our experimental conditions. In this last case, it is possible that the interaction requires the three components (TACC-α2, Ct-TACC-α4 and XMAP-pCt) simultaneously, or involves other regions of the full-length proteins not included in these peptides.

In the case of the peptide trios, we found three different situations employing CD data. First, for the trios in water (**Figure 2.14**) no differences between the sum of the individual spectra and the spectrum of the corresponding trio are observed. Following the interpretation of the pair-wise mixtures, this can be attributed to weak or flickering interactions that do not affect the secondary structure, or that no interaction is present. The second pattern we observed is the gain of α-helix structure as a consequence of the molecular interaction, as seen in both TFE samples (**Figure 2.14 B**) and in DPC at pH 10 for the Ct-TACC-α4 mut trio (**Figure 2.14 E**). This points to a reinforcement of the preformed structure as a consequence of new stabilizing interactions between partners or with the solvent during an induced fit recognition mechanism. Finally, we also observed a slight loss of structure due to interaction in DPC media at acidic pH for the Ct-TACC-α4 wt trio (**Figure 2.14 C** and **D**).

On the bases of our data, it seems that the simultaneous presence of the three preformed helices is a prerequisite for the interaction, since in the cases where one or more peptides has low helical content no interaction is detected. This has been observed for all the trios in aqueous solution (**Figure 2.14 A**) and in DPC at pH 10 for the Ct-TACC-α4 wt trio (**Figure 2.14 E**).

The data corresponding with the acidic media in DPC deserves further attention for two reasons: i) the importance of the Asp residues 921 and 922 for binding, whose side chains will be neutral at low pH, and ii) the fact that the interaction happens in the centrosome, whose reported physiological pH is also acidic [107]. *A priori*, we expected an increment of the structure under these conditions, but none of our results support this proposal; in the case of the Ct-TACC-α4 wt trio, there is a decrease in structure and in the Ct-TACC-α4 mut trio,

the helicity was only slightly modified. One possible explanation is that the formation of the native TACC-XMAP complex requires a structural rearrangement, which results in a loss of the global helical content at acidic pHs. This hypothesis does not rule out the role of the Asp residues as being essential for the function and the interaction. In the case of the TFE media, there is an increment of the helical content due to the interaction, even in the presence of the two Asp residues. This can be attributed to the stabilizing effect of the TFE being able to counterbalance the effect of the Asp residues mentioned above.

The NMR experiments allowed the analysis and quantification of the interaction of the most relevant mixture TACC-α2 + Ct-TACC-α4 wt + XMAP-pCt. In general, affected regions in TACC-α2 and Ct-TACC-α4 are in agreement with the mutagenesis and functional studies [77], and corroborate the contribution of both charged and hydrophobic side chains towards the formation of the complex. Surprisingly, these effects are not seen when we studied the TACC-α2 and TACC-α4 pair, suggesting that in the full-length protein, other stabilizing interactions contribute to maintain these two helices together. In this regard, XMAP-pCt seems to favor a close interaction between TACC-α2 and Ct-TACC-α4 to make the binding possible.

In the simplified model presented here, the simultaneous presence of the preformed helices of all components (TACC-α2, Ct-TACC-α4 wt, and XMAP-pCt) is a prerequisite to enable the observation of the interaction. As we have shown, in the absence of intracellular-like cosolvents, the peptides have low helical content and then no interaction is detected. We think that the data presented here improve the current understanding of the specific molecular recognition and biological function of these proteins.

Finally, we should mention that all our data correspond to the 1:1:1 complex and not with, in principle, the biological 1:1:2 complex. In spite of this, we think that the results obtained with this simple model are illustrative, allow to detect interactions that are structural and environmentally dependent, and can describe the initial steps of the recognition event and not necessarily the last step of the interaction. This approach can probably be useful for other systems in which no other information can be available by traditional methods.

2.4.3 Structure-function relationships

It should be mentioned that the initial objective of the present study was to see if we could detect any sign of a possible biological 1:1:2 complex as, for instance, the identification of two binding faces for XMAP. However, our computational results did not reveal a preference for XMAP binding on the other side of the antiparallel arrangement of the XTACC helices

under the experimental conditions used in this study. This is understandable, since these models were generated by an NMR-driven approach, and they are able to explain the pattern of NMR chemical shift perturbations of the 1:1:1 complex. However, we cannot exclude other binding modes for XMAP.

The resulting assembly consists of the antiparallel packing of TACC-α2 and Ct-TACC-α4, and the antiparallel interaction of the latter with XMAP-pCt. It is remarkable that the NMR chemical shift pattern illustrated in **Figure 2.17** exhibits a high degree of correspondence with the calculated atomistic model. For instance, the antiparallel arrangement of XMAP-pCt and Ct-TACC-α4 is consistent with the chemical shift variations in the N-terminus residues of XMAP-pCt, as well as changes around the Lys2053-Leu2054 and Arg2060-Ile2061 regions. Similarly, spectral changes in the complementary region spanning residues Glu915-Leu923 of Ct-TACC-α4 are explained within the context of this interaction, as well as in residues Ile924-Met927 that are contacting the segment Ile829-Thr835 in TACC-α2 (**Figure 2.18**).

Furthermore, an explanation for the lack of TACC-XMAP-pCt interaction when both Asp 921 and 922 are mutated to Ala residues can be proposed on the basis of our model. As can be seen in **Figure 2.18**, mutations at those positions result in a crucial breakage of the salt bridges established by Ct-TACC-α4 with both TACC-α2 (through Asp921) and XMAP (via Asp922).

Interestingly, these results could be extended to the homologous human domains. The sequence alignment of XTACC3 and XMAP215 with the corresponding proteins of different species including TACC human isoforms (**Figure 2.19**), indicates a high degree of conserva-

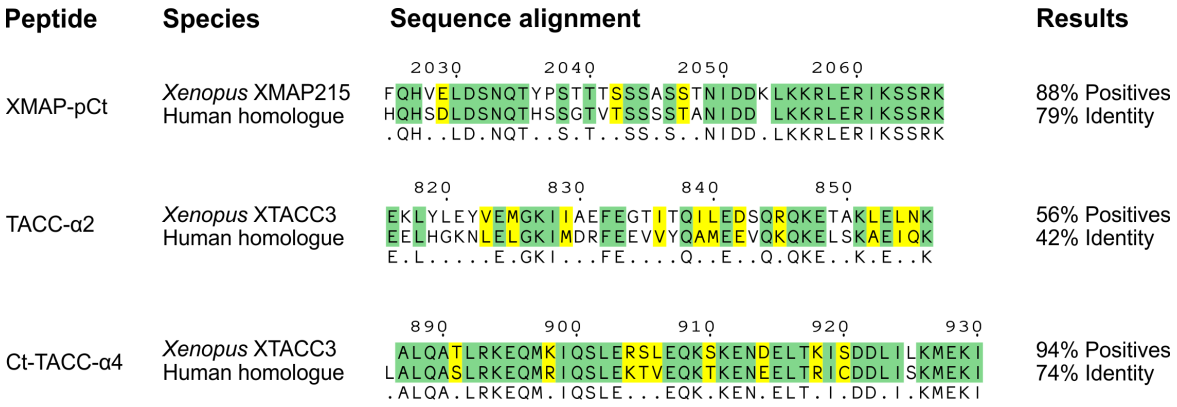


Figure 2.19. Alignment of sequences of the *Xenopus laevis* XMAP215 and XTACC3 with their human homologs. The positives percentage indicates the number and fraction of residues for which the alignment scores have positive values (yellow), indicating similar residues, the identity reflects the percentage of residues that are identical (green). Identical residues are listed on the bottom rows. Analysis generated with ESPrint [109].

tion, specially in those regions corresponding to Ct-TACC-α4 and XMAP-pCt. This suggests that the herein described molecular basis of the interaction is conserved. Bearing in mind that the role of the human homologs in tumorigenic processes has been reported [110], it is tempting to propose that new beneficial processes against cancer can be envisaged by focusing on the residues shown to be important for the interaction in this study.

2.5 Conclusions

According to the proposed objectives and the results described and discussed above, the main conclusions of this chapter are:

1. It has been proven how complex systems can be simplified to successfully surpass the current technological limitations. In this case, the XTACC3-XMAP215 interaction has been described by CD and NMR based on isolated and mixed peptides, and with the use of cosolvents.

2. All single peptides have helical tendencies in intracellular-like surroundings that match with the secondary structure adopted by those sequences in the full-length proteins. The population of these structures depends on the media and pH. However, they are not forming the predicted coiled-coil arrangements. One plausible reason explaining the absence of noticeable coiled-coil structures might be that longer sequences are needed in order to form the coiled-coil. Likewise, it is possible that other weak or transitory interactions in the biological conditions might take part in the stabilization of the coiled-coil.

3. For the XTACC3 peptide with the Asp mutations at positions 921 and 922, which are crucial residues for the interaction with XMAP, the preferred structure was not affected and the helical tendency was even higher than that observed for the native sequence in most of the tested conditions. This finding suggests that the lack of interaction of TACC carrying these mutations with XMAP is not related to the loss of the 3D structure, but with the specific interactions in which the Asp residues side chains participate.

4. We have probed by CD (changes in band intensities) and NMR (changes in chemical shifts) that the designed peptides interact in solution at a 1:1:1 ratio in the presence of intracellular-like cosolvens that stabilize the helical structures, which do not occur in pure water. This evinces that the simultaneous presence of all preformed helices should be a prerequisite for binding.

5. NMR data of the peptide mixtures afford the identification of the specific residues that are affected by the interaction as follows: i) in TACC-α2 most of them are concentrated at the N-terminus, Gly826 and Ile829-Thr835; ii) in Ct-TACC-α4, the largest changes are seen for Leu906 and Glu907 and at the C-terminus, Glu915-Met927; iii) for XMAP-pCt, only the N-terminus (Ser2047-Asn2049) region was significantly affected. This is

compatible with an antiparallel arrangement of TACC-α2 and Ct-TACC-α4, and the antiparallel position of XMAP-pCt with respect to TACC-α4.

6. The information obtained by NMR spectroscopy was used to drive a simple computer-based structure of the current interaction. In this model, TACC-α2 and Ct-TACC-α4 might pack with the fold reported from SAXS data through the Asp921-Lys827 (Ct-TACC-α4-TACC-α2) interaction. In addition, a second favorable electrostatic interaction between Glu907 (Ct-TACC-α4) and Lys848 (TACC-α2), and the hydrophobic packing of Ile924 (Ct-TACC-α4) with Ala830 (TACC-α2) were found to stabilize the helix bundle. In the ternary complex, XMAP-pCt helix likely establishes two sets of non-polar and ionic contributions with TACC-α4 as follows: Leu916-Ile2061, Leu923-Leu2054; and Glu915-Arg2060, Asp922-Lys2053, respectively.

7. The fact that we have probed the interaction at the 1:1:1 molar ratio can be relevant, as the model can describe the initial steps of the recognition event and not necessarily the last step of the full interaction represented by the functional XTACC3/XMAP215 1:2 complex.

8. Finally, as the sequences studied here are highly conserved throughout evolution and given that the human homologs are involved in carcinogenic processes, the present findings on the specific residues related with the direct interaction might be used in the future as targets to design new therapeutics.

3

Structural analysis of protein-membrane interactions employing peptides derived from the HIV-1 ectodomain gp41 in cosolvents

3.1 Introduction and objectives

The human immunodeficiency virus (HIV) is a retrovirus classified into the Genus *Lentivirus*, within the subfamily Orthoretrovirinae [111]. After being identified as the cause of the acquired immunodeficiency syndrome (AIDS) in 1983 [112], it has been considered a global health problem with 36.7 million people infected [112–114]. There are two types of HIV viruses (1 and 2) that present the same basic genetic arrangement, intracellular replication pathways and clinical symptoms, but the transmission and progression to AIDS due to HIV-2 is much slower than in HIV-1 [115]. HIV-1 is probably the most important type from the sanitary point of view, as it is the cause of the global pandemia, whereas HIV-2 is located almost exclusively in Western Africa.

The symptoms of the HIV infection in form of AIDS are caused by the system's inability to offer an immune response against infections through the production of T-cells [112, 116]. Once the infection is established, it triggers a persistent and progressive process that finally leads to the complete destruction of the CD4+ subset of T-cells [117, 116], leaving the infected body totally defenseless.

3.1.1 Main structural components of the HIV viral particle

HIV is characterized for its high genetic variability and mutation rates, and also for its fast replication cycle that makes it difficult to investigate towards drug design [118–120]. As seen in **Figure 3.1**, the virus is spherical with a diameter of approximately 120 nm of diameter [121]. The outer membrane of the viral envelope is composed by the lipid bilayer derived from the membrane of a previously infected cell. It contains some of the original proteins from the infected host cell and other viral proteins known as envelope proteins (Env).

Between the viral capsid and the lipid bilayer, a matrix composed by the viral protein p17 guarantees the integrity of the virion particle **Figure 3.1**. p17 is a structural protein that participates in the early stages of the virus replication and acts as a viral cytokine [122]. Recently, its misfolding has been related with the formation of toxic amyloidogenic assemblies that induce neurocognitive disorders [123]. The conical capsid is composed by roughly 2000 units of associated viral p24 proteins [124] and surrounds the nucleocapsid, which is composed by p6 and p7 proteins [125, 126]. The RNA is bound to the nucleocapsid proteins and consists of two (+) RNA genomes; once the virus binds to the cell, the enzymes essential for the virion formation such as the reverse transcriptase, proteases, ribonucleases and integrases

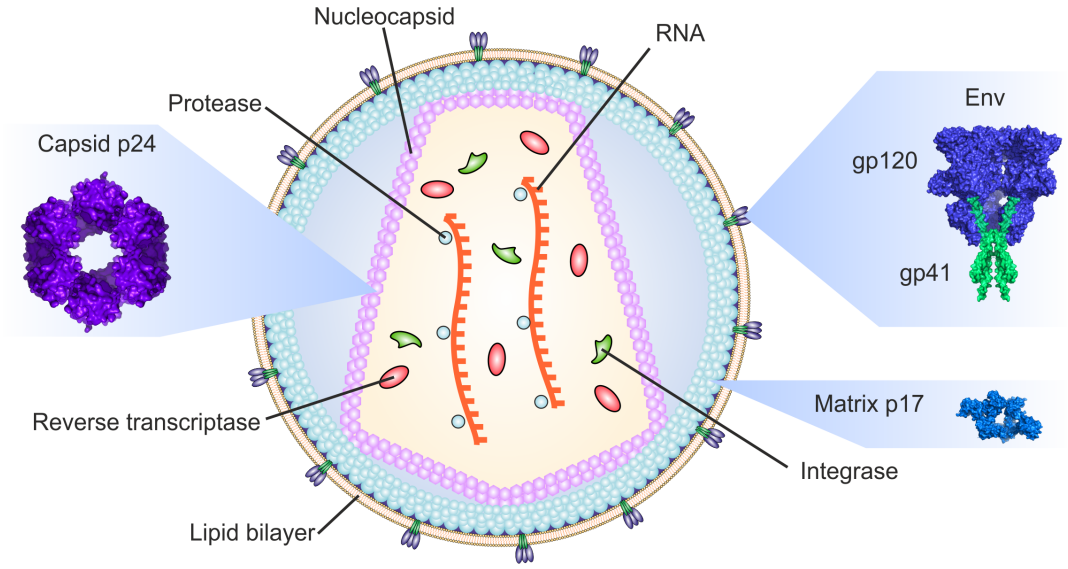


Figure 3.1. Schematic image of the viral structure of HIV. The overall structure comprises around 120 nm of diameter and is spherical. A lipid bilayer and several different proteins constitute the outer membrane; the Env-glycoprotein complex is composed by the gp120 and gp40 glycoproteins, which are responsible for the host cell recognition and binding. The matrix surrounds the capsid and is assembled by the association of p17 proteins. The capsid wraps the essential virion contents: RNA, the reverse transcriptase, proteases and integrases.

are released and the reverse transcriptase transcribes the RNA into DNA inside the infected cell [127, 126].

3.1.2 The Env complex is essential for T-cell recognition and infection

HIV is an enveloped virus that introduces into the macrophages and CD4+ T-cells. The envelope protein complex (Env) is initially produced as the precursor gp160, and is essential for the viral entry as it allows the virus to recognize and bind into the host cells, to adhere and fuse with the cellular targets and initiate the infectious cycle [128–131]. The Env complex is composed by two glycoproteins after the maturation of the gp160 precursor: gp120 and gp41 (**Figure 3.1**). Three gp120 glycoproteins are associated as a cap, and three gp41 units act as a stem that anchors the cap into the membrane [128, 132].

The entry into the host cell is mediated through the interaction of the viral trimeric gp120 with the CD4 (cluster of differentiation) binding domains, which is a glycoprotein found in the cell surface (**Figure 3.2 B**). This elicits a conformational change of the Env complex and exposes the chemokine receptor binding domains of gp120 that allows it to bind to the CCR5 (C-C chemokine receptor type 5) receptor in the cell [129, 126, 131] (**Figure 3.2 C**); the

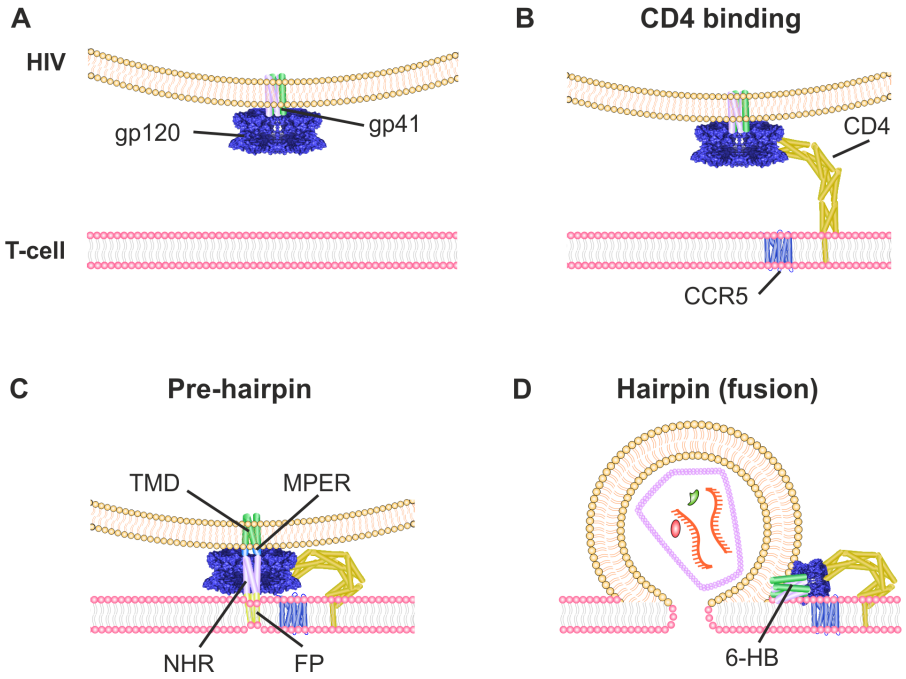


Figure 3.2. Schematic representation of the cell recognition and fusion during infection. A: In a free virion state, gp41 proteins are embedded into the Env-complex. B: The T-cell CD4 protein recognizes and binds to gp120, C: gp120 binds to the T-cell co-receptor CCR5 and the FP domain of gp41 anchors to the membrane of the T-cell. Several conformational changes take place, D: the filamentous structure of gp41 collapses and folds, a six helical bundle (6-HB) is formed and the energy liberated during the process brings the membranes together and fusion occurs.

N-terminus fusion peptide (FP) of gp41 penetrates into the host cell membrane, while the transmembrane domain (TMD) portion of gp41 is still anchored into the viral membrane, and subsequently, the pre-hairpin gets formed. This filamentous structure collapses and forms a six-helix bundle (6-HB) or hairpin trimer of the CHR (C-terminus helix region) regions. Then, the NHR (N-terminus helix region) and CHR regions interact and the two membranes merge (**Figure 3.2 D**).

Each stage in the complex process of the HIV entry can be a suitable target for the development of new antiviral molecules [133, 128, 129]. One of the most promising strategies is to target the gp41 regions of the Env complex, given that the antibody binding to gp41 would block the membrane fusion step [133, 131].

3.1.3 Structural basis of the gp41 glycoprotein

The fact that the sequence of gp41, unlike gp120, is highly conserved makes it an appropriate target for antibody recognition. Gp41 is composed by 345 amino acids and has a molecular

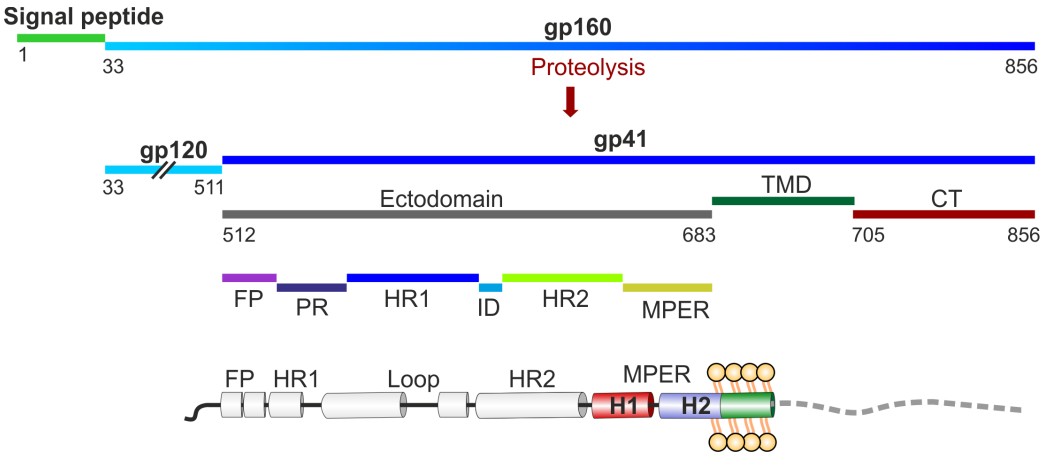


Figure 3.3. Domains and sequence of gp41. Glycoprotein gp160 is the precursor of the glycosylated proteins gp120 and gp41. The envelope protein gp41 is made by 345 amino acids in which three different domains are distinguished: the ectodomain with 172 residues, the transmembrane domain (TMD) with 21 and the cytoplasmic tail (CT) with 152 amino acids. Different constituents compose the ectodomain: the fusion peptide (FP), the polar region (PR), the amino- and carboxy-terminus heptad-repeat regions (HR1 and HR2, respectively), the immunodominant region (ID) and the membrane-proximal external region (MPER). The MPER region has two identified helical segments: H1 and H2.

mass of 41 kDa. It contains three main defined domains (**Figure 3.3**) denoted as the ectodomain or extracellular region (with 172 residues), the transmembrane domain (TMD, 22 residues) and the cytoplasmic tail (CT, 151 residues) [134, 130].

The ectodomain is also subdivided into various regions (**Figure 3.3**). The fusion peptide (FP) is the region closest to the N-terminus and is followed in sequence by the polar region (PR) [135]; afterwards, two hydrophobic regions referred to the heptad-repeat regions HR1 and HR2 (also known as N-helix NHR and C-helix CHR) form α-helical coiled-coil structures [128, 136, 137, 130, 138]. The immunodominant region (ID) lies between the HR1 and HR2 regions, and finally, the membrane-proximal external region (MPER) is a Trp-rich zone close to the transmembrane domain, which is followed by the cytoplasmic tail domain [135, 130].

The fusion peptide (FP) is a hydrophobic region composed primarily by Gly, Leu and Phe amino acids, and the order of the residues belonging to its conserved hydrophobic center is essential for its function, as their randomization, which keeps its net hydrophobicity, alters its function [130]. It is also the region closest to the N-terminus and, therefore, might be embedded into the gp120-gp41 complex. The binding of gp120 with CD4 leads to a conformational change, which exposes the FP region and lets it to penetrate into the target cell, causing the cell membrane destabilization and, therefore, leading to a pore formation that allows the virus to internalize into the cell [139, 135].

The heptad-repeat regions HR1 and HR2 are linked through a disulfide bond, where the immunodominant region acts as a hydrophilic loop between them [135]. The HR motifs form a stable six-helix bundle (6-HB) that draws the viral and cell membranes close together through the formation of a bundle of three HR1 helices (**Figure 3.2 D**) that folds over a hydrophobic groove antiparallel with respect to the other three HR2 regions [128, 138]. These heptad regions have been proposed previously as a target towards drug design. Indeed, there are currently various clinically approved peptides (enfuvirtide, T-20, fuzeon) derived from HR1 and HR2 that can block the formation of the bundle and control the HIV infection [140–144].

The membrane-proximal external region (MPER) comprises the last 24 residues of the ectodomain and it is well known that it is essential for the fusogenic process. As well, its highly conserved sequence is recognized by several neutralizing antibodies [145, 146].

The transmembrane domain (TMD) anchors the Env complex to the lipid bilayer (**Figure 3.2**) and is almost entirely composed by hydrophobic residues that ensure its docking into the lipids. It is also a highly conserved region, and mutational studies suggest that it exerts important roles during the membrane fusion stage [147–149]. However, its transmembrane nature makes it difficult to express and purify, and two different proposals based on its predicted structure have been considered: one suggests that it is an α-helix that crosses the membrane once and maintains the cytoplasmic tail inside the virus particle [150, 151], while the other one proposes that the TMD crosses the membrane three times [152, 153].

Finally, the cytoplasmic tail (CT) has been found to be essential for the spreading of the infection. Its truncation leads to a defect of the Env complex incorporation into the virions and therefore, the CT region might also be a suitable target in HIV treatments [135].

3.1.4 Broadly neutralizing antibodies against MPER of gp41 as therapeutic molecules

Some individuals develop broadly neutralizing monoclonal antibodies upon the infection of HIV against the Env complex and consequently, there is an increased interest on their clinical use [154]. Within the Env complex, the MPER region is well known as being targeted by several antibodies such as 2F5, 4E10, Z13 and 10E8 [133, 145, 155, 156, 132, 146] (**Figure 3.4**).

The exceptionally high degree of conservation of the MPER region makes it a great choice for use in immunotherapy. However, there is a lack of information on how these antibodies act [133, 84] and different approaches employing peptide epitopes belonging to their targets have been used [133, 130].

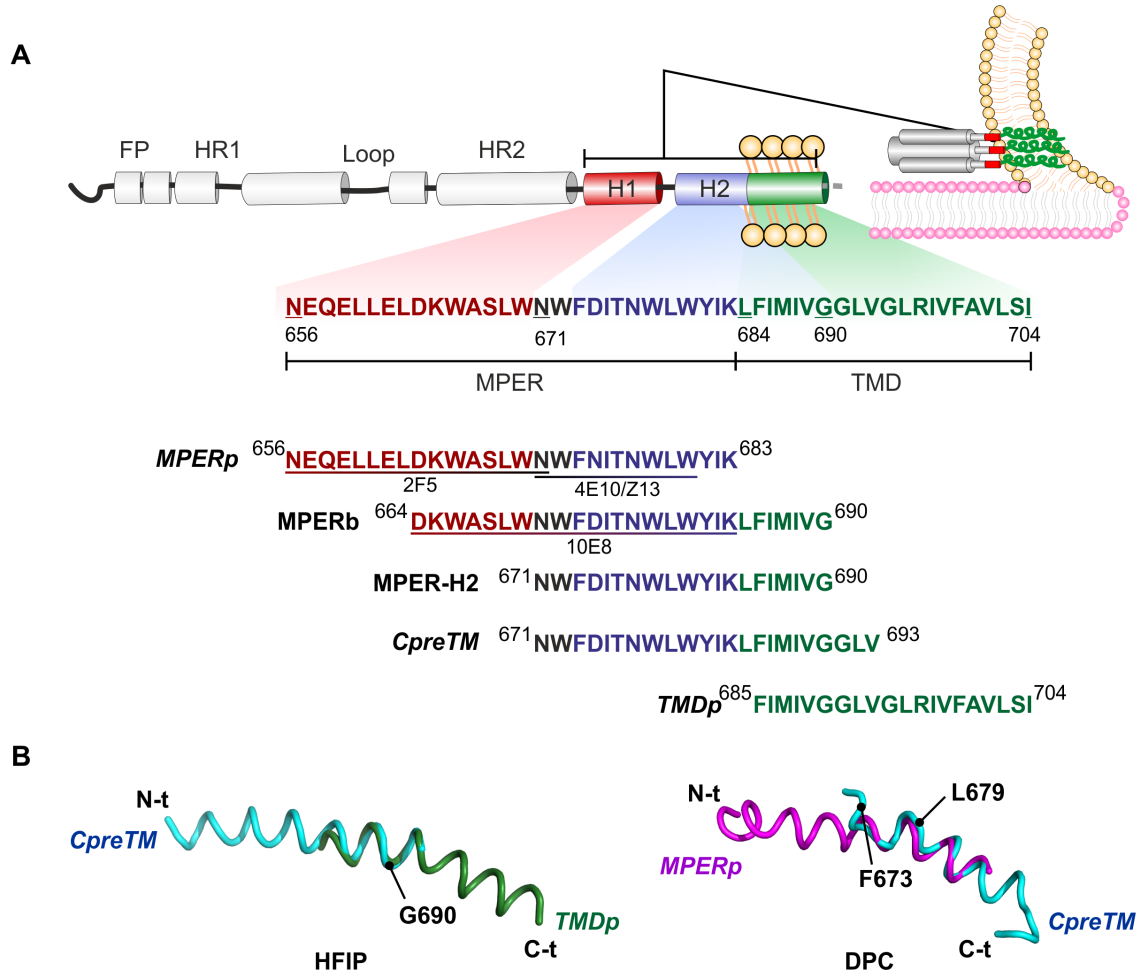


Figure 3.4. Comparison of the structural elements of the peptides studied previously (in italics) and in this study (MPERb and MPER-H2). The peptides span a region previously identified to be involved in the anchoring and fusion of the membranes. A: Regions corresponding with the three distinct segments H1, H2 and TMD are shown in red, blue and green, respectively. Theoretical models predict that the H2 segment spans until Lys683, while experimental results include it until Gly690. Regions corresponding to the epitopes of the 2F5, 4E10, Z13 and 10E8 antibodies are underlined. B: Superposition of the previously obtained structures in ribbon representation: CpreTM (cyan), TMDp (green) and MPERp (magenta).

In this regard, the team led by Dr. M. A Jiménez at the IQFR in collaboration with Dr. J. L. Nieva at the Biophysics Unit of the University of the Basque Country (UPV) is pioneer in studying the structure of immunogenic gp41-derived peptides in membrane-like media, as well as their binding with HIV antibodies. Thus, their results on peptides encompassing different MPER segments are the foundation of the current study.

3.1.5 Simplification of the Env complex system

Since our group has special interest in understanding a particular complex macromolecular association consisting on the protein-membrane recognition of the Env complex at atomic resolution, there are several biophysical techniques that are complementary in the study of the structure of complex systems, of which the most widely used are NMR spectroscopy, X-ray crystallography and cryo-electron microscopy. Solution NMR is a unique technique to obtain structural information regarding these dynamic complexes in different environments, but its inherently low sensitivity and the dependence on dynamic time scales sometimes limits its applicability. Thus, the correct design of reduced and simplified molecular constructions is needed to obtain structural information for these specific cases, among them, the system examined in this study.

As mentioned above, our group together with Nieva's laboratory aim to gather knowledge about the Env complex through the characterization of peptides derived from the gp41 MPER region, which contain the epitopes of monoclonal antibodies. Prior to this study, three peptides containing the 2F5 and 4E10 epitopes (**Figure 3.4**) were characterized: MPERp (Asn656-Lys683), CpreTM (Asn671-Val693) and TMDp (Phe685-Ile704) [133, 132, 130]. As explained below and as a general hallmark, these peptides are helical and exhibit kinks at determined positions.

Prior to the experimental studies on MPER peptides on the basis of molecular dynamics simulations, the region between residues Lys681 to Arg707 the structure was predicted to be in an α-helix conformation and with kinks around the residues Lys683 and Arg696 [157]. However, NMR studies performed in our group employing the peptides MPERp, CpreTM and TMDp, which have overlapping sequences covering the region between Asn656 and Ile704, revealed that all of them adopt helical conformations with kinks at positions that depend on the media [133]. Thus, MPERp in HFIP shows a kink at positions Lys665-Trp666, while in DPC micelles, the kink is located between the residues Phe673 and Ile675. Concerning CpreTM and TMDp in HFIP media, they consist of a continuous helix spanning the sequence Asn671-Ile704, which breaks at Gly690, where a kink is found (**Figure 3.4 B**). The structure of CpreTM in DPC micelles shows two kinks: one between residues Asn671 and Phe673 and the other at Gly690. Nevertheless, the low solubility of the peptide TMDp impeded measurements within DPC micelles. It is interesting to note that the NMR results concerning the peptide CpreTM (Asn671-Val693) reveal that the kink predicted by the computer models located at Lys683 is not present and indicate the existence of a continuous α-helix, including the C-terminus segment of the MPER domain and the N-terminus of the TMD domain [133, 132]. This result highlights the importance of acquiring experimental data,

since predictions are sometimes wrong. Based on these results, a model for the insertion of the TMD region into the membranes has been proposed [133].

Taking into account these previous results, we propose to study peptides whose sequences contain the continuous helix found in CpreTM, and TMDp, denoted as H2, and to obtain insights into the mechanism in which the HIV-1 MPER epitope is recognized by the broadly neutralizing antibody 10E8, which remains unclear.

3.1.6 Objectives

Previous studies suggest that HIV infections can be successfully neutralized during their first stages. In this context, we aim to obtain new information about the initial steps of the proposed broad neutralization mechanism. As mentioned before, gp41 glycoprotein is a highly conserved sequence in the virus, and for this reason it is more suitable as a target than other more variable entities, like the gp120 glycoprotein. In addition, the fact that the epitopes recognized by the neutralizing monoclonal antibodies are linear in the case of gp41 and discontinuous in gp120, facilitates all types of structural studies and interpretation of the results in the former. Considering the good results obtained with the studies of the MPER and TMD peptides containing the 2F5 and 4E10 epitopes [133, 84], in this study, we focus on peptides that include the key residues of the 10E8 epitope, which involves part of the H1 and H2 helices. These peptides encompass the full continuous helix H2, as experimentally found in the peptide CpreTM [133]. Our goal is to understand the role of this MPER-TMD region of gp41 in the membrane hairpin formation during the first stages of the infection and its neutralization by the 10E8 antibody.

To do this, two overlapping peptides were designed and characterized: i) one spanning the terminus part of the H1 and the complete H2 helices (Asp664-Gly690) of the MPER-TMD regions, which contains the full 10E8 epitope, named MPERb; and ii) a shorter version named MPER-H2, covering only the H2 helix region (Asn671-Gly690) of MPER-TMD. In this regard, it has been demonstrated that the helical scaffold comprising the MPERb sequence within gp41 is biologically relevant towards the production of new neutralizing antibodies.

The natural environment of the sequences studied here is the membrane proximal external region of the cell. This fact together with the highly hydrophobic nature of the peptides, determines the use of solvents of low polarity to mimic the natural media in the proposed structural studies. Also, given that the main components of the cell membrane are phosphatidylcholine lipids, we also employ DPC (dodecylphosphocholine) micelles to mimic the lipid bilayer and to evaluate the possible interactions between the peptides and the membrane lipids.

Therefore, the specific objectives proposed are:

- To design new overlapping peptides spanning the sequences corresponding to the epitope of the 10E8 broadly neutralizing antibody comprising part of the MPER and TMD regions.

- To fully characterize at atomic resolution the structures of these peptides in different cosolvents in order to obtain biologically relevant structural information to determine the implication of these regions during the infective process.
- To compare the obtained results with previous models [133], in which the existence of a kink at the MPER region that might help on its insertion into the membrane is hypothesized.
- To determine more specifically the zones that might be crucial during the membrane fusion process and propose a new model on how this process occurs using membrane-like cosolvents.
- To clarify the mechanism by which the epitope is recognized by the 10E8 antibody based on the structure and binding properties of the new designed peptides derived from our NMR data and additional data obtained by our collaborators at the Biophysics Unit of the University of the Basque Country.

3.2 Materials and methods

3.2.1 Chemicals

The deuterated compounds [D_{38}] DPC (dodecylphosphocholine) (98 %) and D_2O (99.9 %) were obtained from Cambridge Isotope Laboratories (USA); [D_2] HFIP (1,1,1,3,3,3-hexafluoro-2-propanol) (99 %) was obtained from Eurisotop (France). The percentages of deuteration are indicated in parenthesis. Non-deuterated DPC was purchased from Avanti Polar Lipids (USA).

3.2.2 Peptide synthesis

MPERb and MPER-H2 peptides were made by Fmoc-solid phase synthesis and provided by the group of Dr. José L. Nieva (Biophysics Unit, Department of Biochemistry and Molecular Biology, University of the Basque Country). Peptide sequences are listed in **Table 3.1**; the numbering of the peptide residues matches with the one in the wt sequence of gp41 prior to maturation and post-translational modifications.

Table 3.1. Sequences of the peptides studied. Residues belonging to helix H1 and H2 are in red and green, respectively and those of the solubility tags are in italics.

Peptide (protein sequence #)	Sequence	Theoretical pI
MPERb (664-690)	*KKKK*DKWASLWNWFDITNWLWYIKLFIMIVG*KKKKK*	10.30
MPER-H2 (671-690)	NWFDITNWLWYIKLFIMIVG*KKKKK*	10.13

3.2.3 NMR sample preparation

NMR samples were prepared by dissolving the lyophilized peptide (1-2 mg) in 0.5 mL of 2 mM HEPES buffer, pH 7.0 with either 20 mM [D_{38}] DPC in H_2O/D_2O (9:1 v/v), 20 mM [D_{38}] DPC in D_2O, 20 mM [D_{38}] DPC/non-deuterated DPC (1:1 v/v) in H_2O/D_2O (9:1 v/v), [D_{38}] DPC/non-deuterated DPC (1:1 v/v) in D_2O, [D_{38}], 25 % [D_2] HFIP in H_2O/D_2O (9:1 v/v) and 25 % [D_2] HFIP in D_2O. Peptide concentrations were between 0.9 and 1.8 mM. The pH was adjusted to 7.0 when necessary by adding minimal amounts of NaOD or DCl, measured with a Hamilton glass micro-electrode.

All samples were placed in 5 mm NMR tubes, and contained DSS as the internal reference for ^1H chemical shifts.

3.2.4 NMR spectra acquisition, assignment and structure calculation

NMR spectra were recorded at 25 and 35 ºC with a Bruker Avance spectrometer, operating at 600 MHz (^{1}H), equipped with a z-field gradient cryoprobe.

Phase-sensitive two-dimensional total correlated spectroscopy TOCSY spectra were obtained by using 20 and 60 ms mixing times. The nuclear Overhauser enhancement spectroscopy NOESY mixing time was 150 ms. ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra were recorded at ^{13}C natural abundance. ^{13}C δ-values were indirectly referenced by using the IUPAC-IUB recommended ^{1}H/^{13}C chemical shift ratio (0.25144953) [101]. Water signal was suppressed by either presaturation or by using a 3-9-19 pulse sequence. The 2D data matrices were multiplied by a square-sine-bell window function with the corresponding shift optimized for every spectrum and zero-filled prior to Fourier transformation. Baseline correction was applied in both dimensions. Data were processed with the standard TOPSPIN program (Bruker Biospin, Karlsruhe, Germany). NMR and sample conditions are summarized in **Table 3.2**.

Table 3.2. NMR spectra conditions of the peptides studied.

Peptide	Solvent	pH	Temperature (ºC)
MPERb	25 % HFIP/H_2O	7.0	25, 35
	25 % HFIP/D_2O	7.0	25, 35
	20 mM DPC/H_2O	7.0	25, 35
	20 mM DPC/D_2O	7.0	25, 35
	20 mM 50 % deuterated DPC/H_2O	7.0	35
MPER-H2	25 % HFIP/H_2O	7.0	25, 35
	25 % HFIP/D_2O	7.0	25, 35
	20 mM DPC/H_2O	7.0	25, 35
	20 mM DPC/D_2O	7.0	25, 35
	20 mM 50 % deuterated DPC/H_2O	7.0	35

3.2.5 NMR spectra assignment and interaction mapping

The NMR spectral assignment of each peptide in the different conditions was performed by following the well-established sequential-specific methodology based on homonuclear spectra [94] with the help of the SPARKY software [95]. The ^{13}C resonances were identified on the basis of the correlations between the hydrogen and the bound carbon atom present in the ^{1}H -^{13}C-HSQC spectra. All assigned chemical shifts are listed in **Appendix B**.

The NMR spectral assignment of the DPC resonances in the partially deuterated DPC mixtures was based on its previously reported chemical shifts [158]. Micelle-peptide interactions were identified based on the ^{1}H assignment of both components. Intermolecular interactions were evaluated from the intermolecular NOEs found in the partially deuterated micelle media, and comparing the new signals appeared with the NOEs not found with the fully deuterated micelles.

3.2.6 NMR estimation of helix populations

Helix populations were quantified from the ^{1}Hα or ^{13}Cα secondary chemical shift values, applying the formula for either the carbon or proton as described previously in **Equation 2.2**.

3.2.7 Structure calculation

Structure calculations were done with the program CYANA 2.1 by using distance and dihedral angle constraints derived from NMR parameters and using the standard iterative procedure for automatic NOE assignment of the program [96]. Distance constraints were obtained from the cross-peaks present in 2D [^{1}H-^{1}H]-NOESY spectra (150 ms mixing time), which were integrated using the standard SPARKY integration sub-routine [95]. Dihedral angle restraints, for Φ and Ψ angles, were derived from ^{1}Hα, ^{13}Cα and ^{1}Cβ chemical shifts using the TALOS-N webserver [97]. The 20 lowest target function structures in the final cycle of the CYANA protocol were selected as the final structure ensembles. The quality of these structures was assessed using PROCHECK/NMR [159] as implemented at the Protein Structure Validation Suite server (PSVS: http://psvs-1_4-dev.nesg.org). The structural ensembles calculated for MPERb have been deposited at the PDB data bank with accession codes 2NCT (in HFIP) and 2NCS (in DPC). The structural statistics data for these structures are provided in **Table 3.4**. The structures for MPER-H2 have not been deposited because, excluding the Lys-tag, it has less than 24 residues, which is the minimal length allowed by the PDB databank. The structural ensembles were visualized and examined using MOLMOL [98] and PyMOL [99].

3.3 Results

3.3.1 Design of optimized peptide epitopes for 10E8

As mentioned in the introduction section, 10E8 is a potent neutralizing antibody that interacts with the gp41 ectodomain. Although some data concerning this interaction is already known, the exact structural details of the MPER portion able to interact with 10E8 remains unclear. Following other studies performed in our group [133, 132, 130], two peptides (**Table 3.1** and **Figure 3.4**), named MPERb and MPER-H2, were designed to be characterized by solution NMR within non-polar (HFIP) and membrane-like environments (DPC micelles). These media have been established as good mimetics of the conditions found before the fusion of MPER with the membrane [133, 160, 132, 130]. MPERb covers the C-terminus part of helix H1 plus the entire length of H2 (**Figure 3.4**). This construction spans all the residues (664-690) reported to establish contacts with the 10E8 antibody [155]. MPER-H2 spans exclusively the residues of helix H2 (671-690) and was designed to understand the structural influence, if any, of H1 on H2 or *vice versa*.

An analysis of the sequences (residues 664 to 670 and 671 to 690) based on previous theoretical and experimental studies [133] indicates that they might have a high propensity to fold as α-helical structures and to present a homogeneous distribution of aliphatic residues, rather than a specific location on one of the faces of the helix (**Figure 3.5**). From the experimental point of view, this high hydrophobicity is an important problem, as it is likely to promote oligomerization and result in low water-soluble samples. To overcome these problems and achieve the sample concentrations required for NMR studies, a Lys tag was added in both ends of the MPERb peptide and at the C-terminus of MPER-H2. This strategy

Figure 3.5. α-helical wheel representations for the structures formed by the peptides MPERb (left) and MPER-H2 (right). The homogeneous distribution of the hydrophobic residues (in green) can be appreciated.

has been applied in many cases and has been reported to not affect the interpretation of the results in relation with the biological processes [133, 161, 162].

3.3.2 Secondary structure of the MPERb and MPER-H2 peptides

The good quality of the NMR spectra in 25 % HFIP allowed a full assignment of the signals corresponding to the MPER peptides. In DPC micelles, the quality of the spectra was not so good because of the signal broadening that made the assignment task more complex. This effect was more accentuated in the MPER-H2 peptide than in MPERb. Despite of this, the assignment was finally completed in DPC for both peptides. The only incompletely assigned residues were the Lys tags as a consequence of signal overlapping.

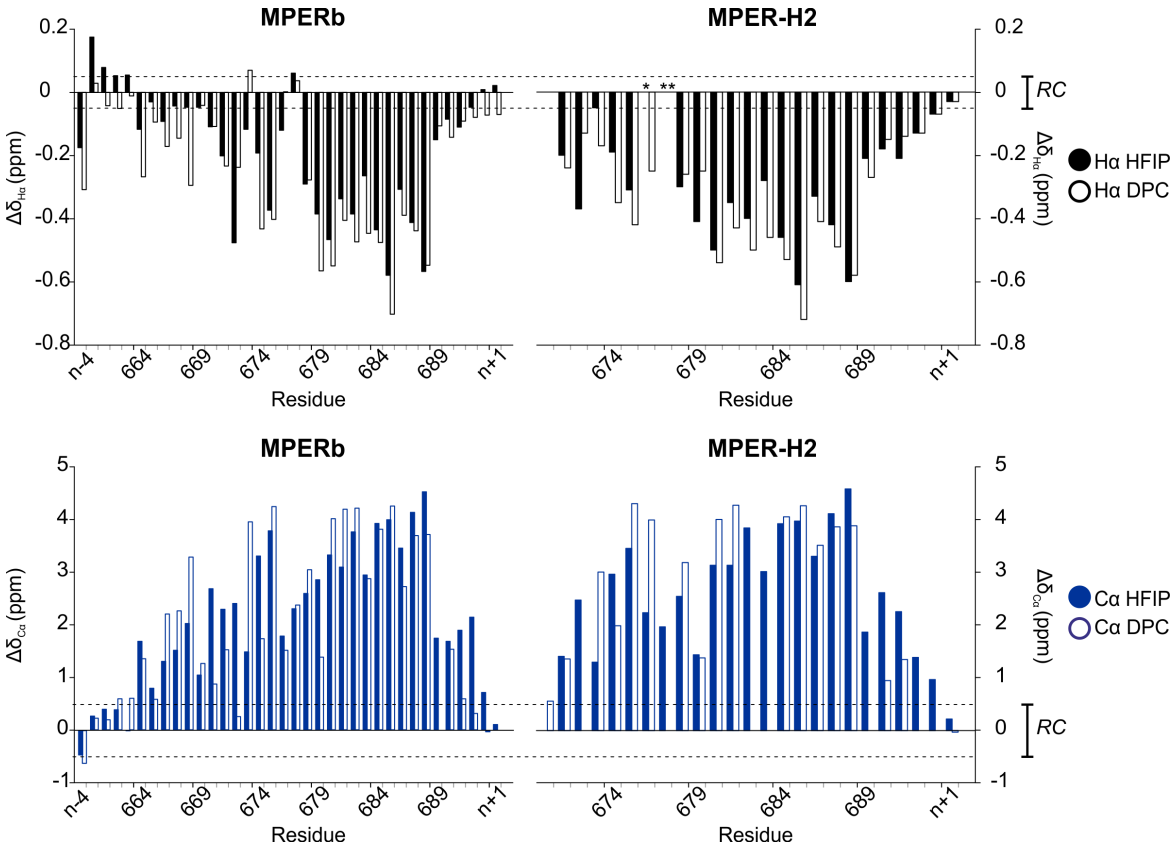


Figure 3.6. Hα (black) and Cα (blue) secondary chemical shifts observed for the peptides MPERb and MPER-H2, pH 7.0. Negative Δδ Hα values of large magnitude (> 0.05 ppm) and large positive Δδ Cα values (> 0.5 ppm) are indicative of helical conformations. Dashed lines indicate the random coil (RC) ranges.
* Unassigned nuclei.

After assignment, a qualitative analysis of the NMR parameters was carried out. First, the Hα and Cα chemical shifts of both peptides show large differences with respect to the random coil values ($|\Delta\delta| > 0.05$ ppm for the Hα and > 0.5 ppm for the Cα, see **Figure 3.6**). The large magnitude of Δδ values and the observed signs (negative for the Hα and positive for the Cα) indicate that the MPERb and MPER-H2 peptides adopt helical structures in both HFIP and DPC micelle media. As seen in **Figure 3.6**, the magnitudes of the chemical shifts deviations at the Lys tags (N- and C-terminus residues in MPERb and C-terminus in MPER-H2; labeled as n - i or n + i) are quite small. This indicates that they might adopt random conformations. It is also noticeable the presence of certain regions (Thr676, Asn677) with lower Δδ Hα values than the mean. This might be explained by the chemical shift anisotropy effects produced by aromatic residues located nearby or at i ± 3 or i ± 4 positions (Phe673, Trp678, Trp680, Tyr681).

Given that only pairs of spatially closed protons produce NOE signals (distances usually lower than 5 Å), the strongest evidence on the formation of α-helical structures is provided by the NOE data. In the case of the peptides MPERb and MPER-H2, the presence of non-sequential NOEs $d_{\alpha N}$ (i, i + 2), $d_{\alpha N}$ (i, i + 3), $d_{\alpha N}$ (i, i + 4), $d_{\alpha\beta}$ (i, i + 3) and d_{NN} (i, i + 2) (**Figure 3.7** and **Figure 3.8**), which are characteristic of helices, evidences that the peptides form α-helical structures in HFIP and in DPC micelles.

The helical populations estimated from the ^{1}Hα and ^{13}Cα chemical shifts (see Method's **Section 3.2.6**), excluding the Lys solubilization tags (**Table 3.3**), are high for both peptides: 61 % and 73 % for MPERb in HFIP and DPC, respectively, and 100 % for MPER-H2 in both cosolvents.

Table 3.3. Comparison of the percentages of global helical populations estimated for MPERp, MPERb, MPER-H2, CpreTM and TMDp peptides from Δδ Hα and Cα NMR in 25 % HFIP and 20 mM DPC, 35 ºC, pH 7.0.
n.d., not determined.
* Peptides studied previously by our group [133, 132, 130].

	Helix population (%)			
	HFIP		**DPC**	
Peptide	**Hα**	**Cα**	**Hα**	**Cα**
MPERp*	59.7	n.d.	59.7	n.d.
MPERb	61.0	98.2	73.1	97.4
MPER-H2	100	100	100	100
CpreTM*	70.0	100	86.5	n.d.
TMDp*	71.3	100	n.d.	n.d.

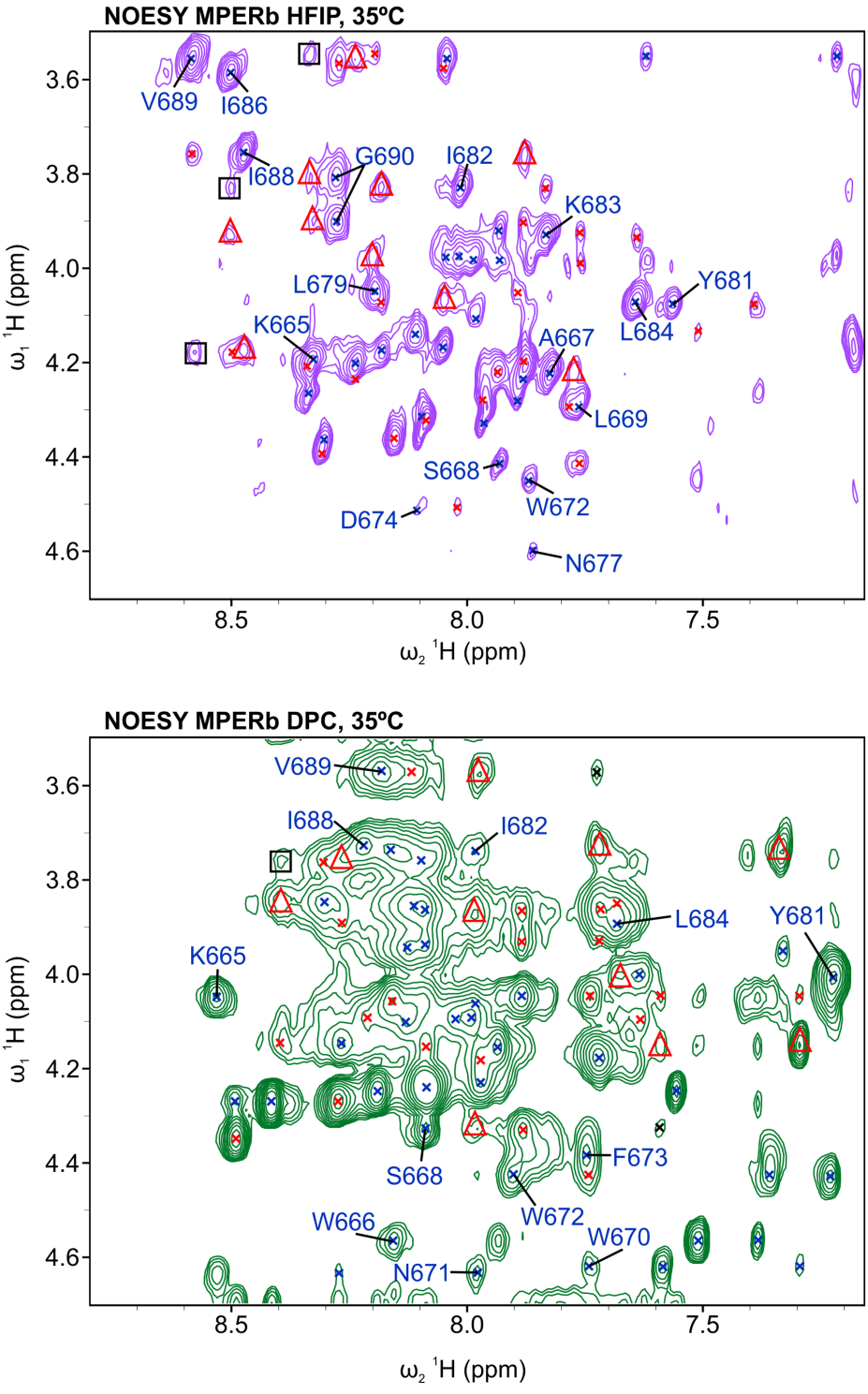


Figure 3.7. Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of MPERb in 35 % HFIP and 20 mM DPC containing 2 mM HEPES buffer, 35 ºC, pH 7.0. Intraresidual NOEs are indicated with blue crosses, intraresidual HN-Hα are labeled, non-sequential $d_{\alpha N}$ (i, i + 1) signals are indicated with red crosses, $d_{\alpha N}$ (i, i + 3) in red triangles and $d_{\alpha N}$ (i, i + 4) with black squares.

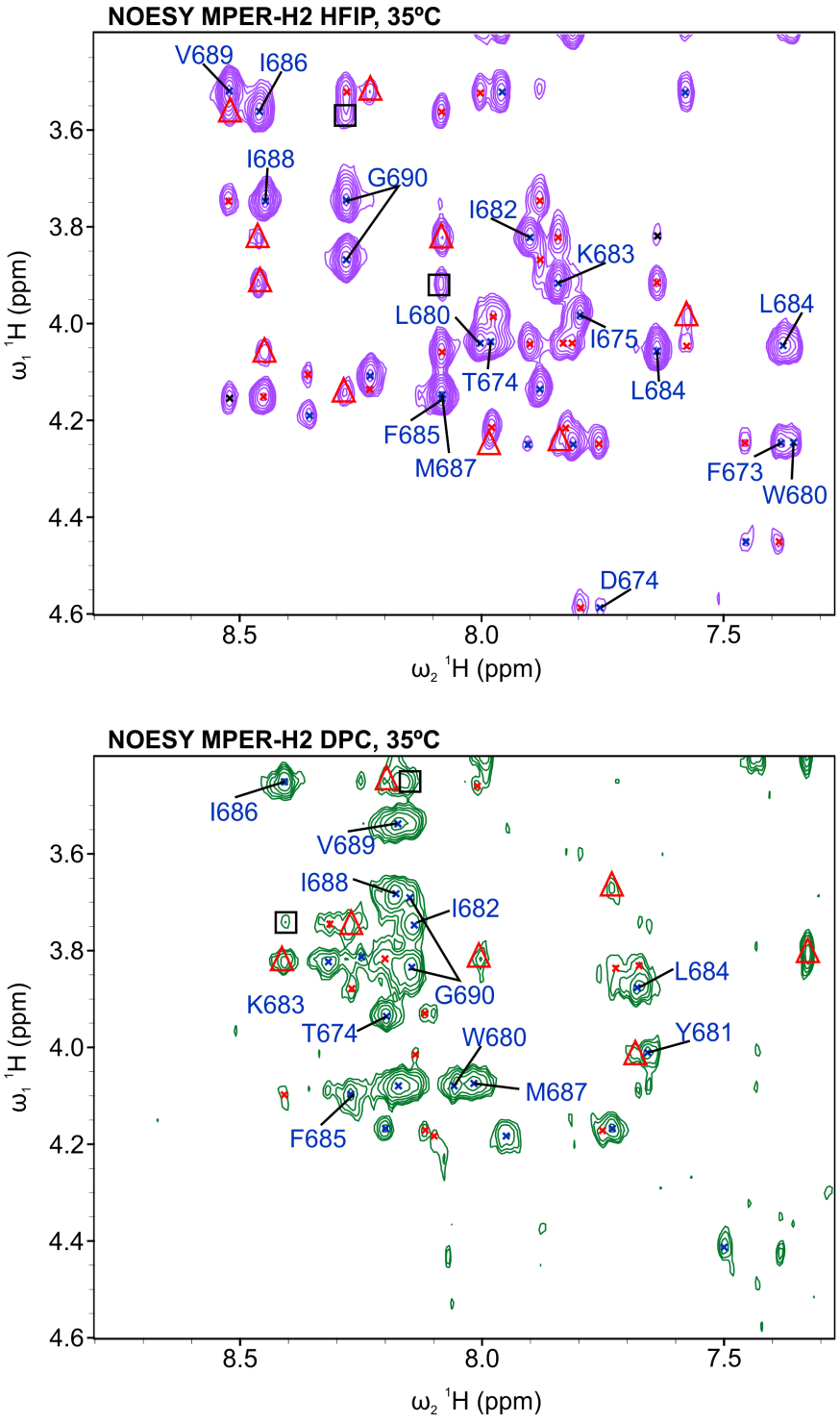


Figure 3.8. Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of MPER-H2 in 35 % HFIP and 20 mM DPC containing 2 mM HEPES buffer, 35 ºC, pH 7.0. Intraresidual NOEs are indicated with blue crosses, intraresidual HN-Hα are labeled, non-sequential $d_{\alpha N}$ (i, i + 1) signals are indicated with red crosses, $d_{\alpha N}$ (i, i + 3) in red triangles and $d_{\alpha N}$ (i, i + 4) with black squares.

3.3.3 3D NMR structure of the MPER peptides

Once analyzed the qualitative information about the secondary structure, the calculation of the preferred conformations of the peptides in HFIP and DPC was performed on the basis of the experimental restraints. It is important to take into account that the results obtained by solution NMR register the peptide conformational average and the evaluation of the data correspond to the most populated conformers in solution. However, as the α-helices formed by the peptides MPERb and MPER-H2 are highly populated (**Table 3.4**), the structure calculations are meaningful. The calculated structures can be seen in **Figure 3.9** and **Figure 3.10** (see further pages). The N- and C-termini comprising the Lys tags show a certain degree of freedom which increases the RMSD values of the overall MPERb and MPER-H2 peptides. Excluding these Lys tags, the resulting structures in both media were well defined, as demonstrated by the small RMSD values for the backbone atoms of residues spanning 664 and 690 of MPERb and from 671 to 690 of the MPER-H2 peptide (**Table 3.4**). The calculated conformers of MPERb were deposited in the Protein Data Bank with the PDB codes 2NCT for the peptide in HFIP and 2NCS in DPC micelles.

The presence of a kink (**Figure 3.11**, see further pages) in the MPERb peptide between Asn671 and Trp672 residues is also remarkable. This observation is in accordance with the structural motif present in the previously reported NMR structure of the peptide CpreTM (**Figure 3.4**) [133], in which the two defined regions H1 and H2 can be inferred.

Table 3.4. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for MPERb and MPER-H2 in 25 % HFIP and 20 mM DPC, 35 ºC, pH 7.0.
* Excluding Lys tags.

	MPERb		MPER-H2	
	HFIP	**DPC**	**HFIP**	**DPC**
Upper Limit distance restraints (from NOEs)				
Total	361	445	356	375
Short-range |i-j| $\leq$ 1	279	311	259	250
Medium-range 1 < |i-j| < 5	82	134	97	125
Long-range |i-j| $\geq$ 5	0	0	0	0
Averaged total number per residue	10.0	12.4	14.2	15.0
φ/ ψ Dihedral angle constraints (from chemical shifts)	29/29	26/26	23/22	17/15
Average CYANA target function value	0.01 ± 0.1	0.11 ± 0.1	0.01 ± 0.1	0.05 ± 0.1
Averaged maximum violation per structure				
Distance (Å)	0.02 ± 0.1	0.03 ± 0.1	0.01 ± 0.1	0.02 ± 0.1
Dihedral angle (º)	0.05 ± 0.1	0.40 ± 0.2	0.01 ± 0.1	0.09 ± 0.1
Pairwise RMSD (Å)				
Backbone atoms	2.7 ± 1.0	2.7 ± 0.8	0.7 ± 0.2	1.9 ± 0.6
All heavy atoms	3.8 ± 0.9	3.7 ± 0.7	1.7 ± 0.3	2.9 ± 0.7
Backbone atoms*	0.7 ± 0.2	0.4 ± 0.1	0.4 ± 0.1	1.0 ± 0.6
All heavy atoms*	1.6 ± 0.3	0.9 ± 0.1	1.4 ± 0.3	1.7 ± 0.7
Ramachandran plot (%)				
Favorable	91.8	88.8	93.6	80
Allowed	8.4	11	6.4	20
Outlier	0.5	0.2	0	0
Favorable*	100	100	96.7	95.5
Allowed*	0	0	3.3	4.5
Outlier*	0	0	0	0

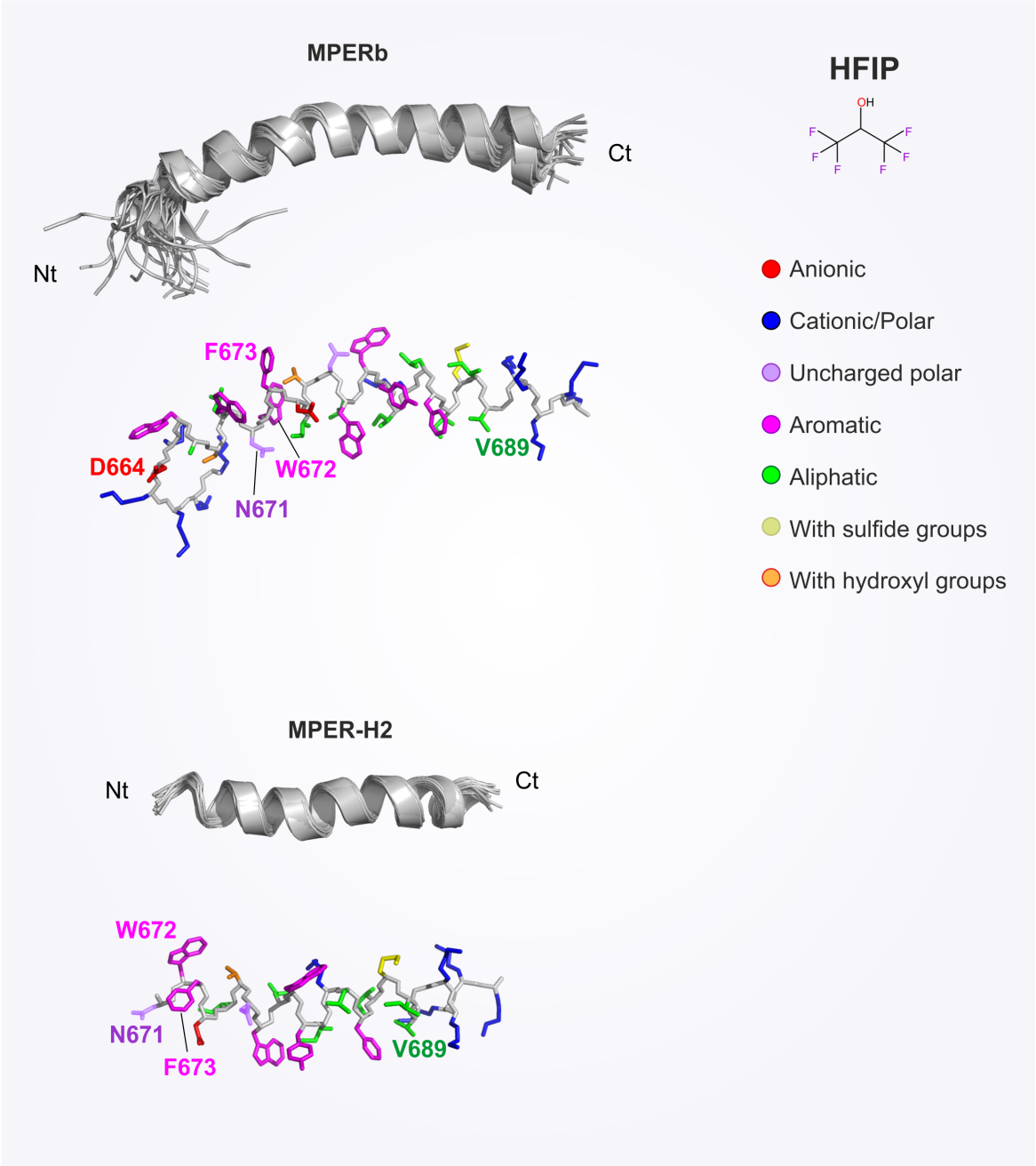


Figure 3.9. Ensemble of the 20 structures with the lowest target function values of MPERb and MPER-H2 in 25 % HFIP, 35 ºC, pH 7.0. The superposition of the backbone of the 20 lowest target function structures in each family is represented as ribbon. Side chains of the lowest target function structure in solution are displayed in different colors depending on the type of amino acid.
Anionic: red, cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

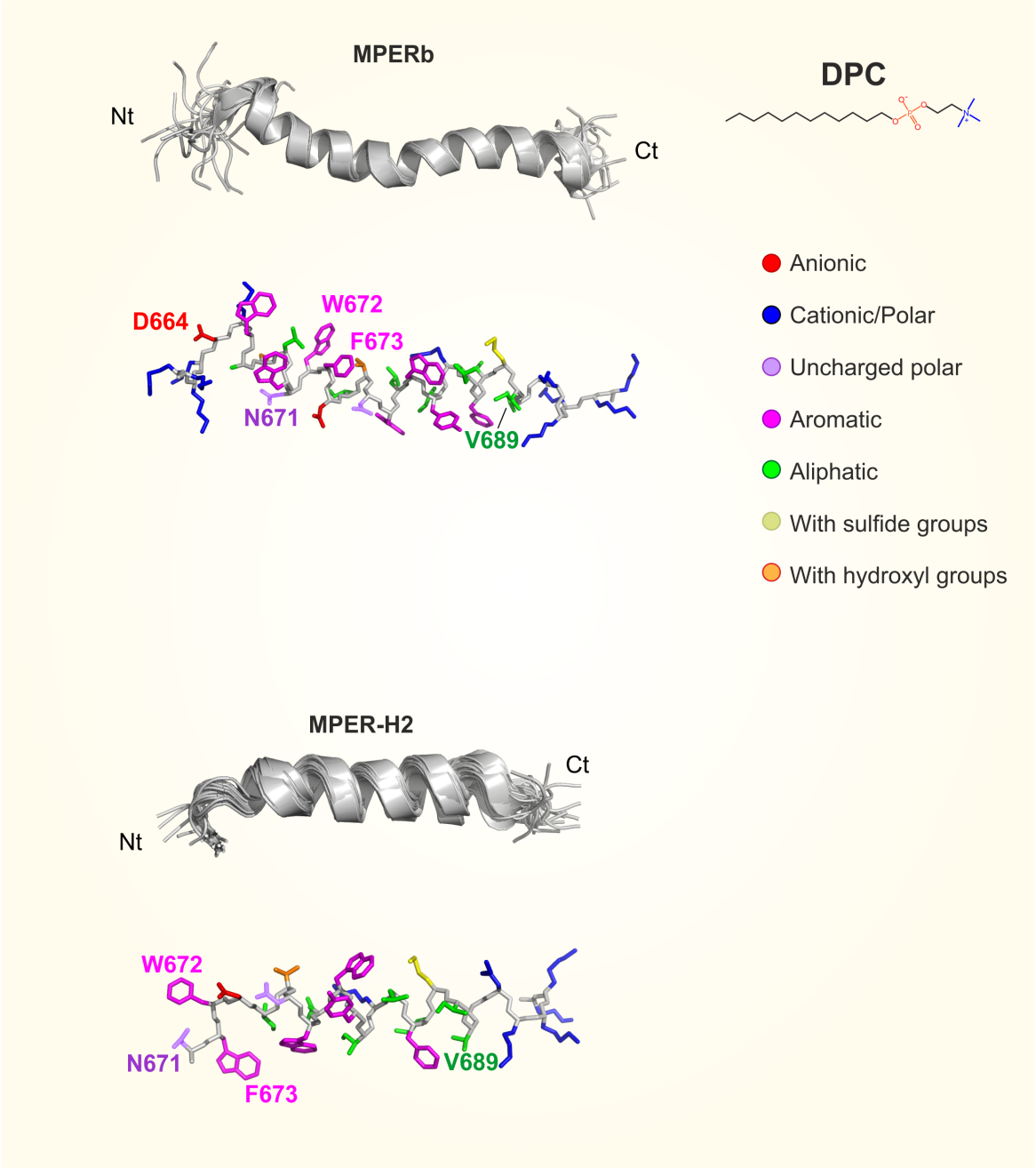


Figure 3.10. Ensemble of the 20 structures with the lowest target function values of MPERb and MPER-H2 in 20 mM DPC, 35 ºC, pH 7.0. The superposition of the backbone of the 20 lowest target function structures in each family is represented as ribbon. Side chains of the lowest target function structure in solution are displayed in different colors depending on the type of amino acid. Anionic: red, cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

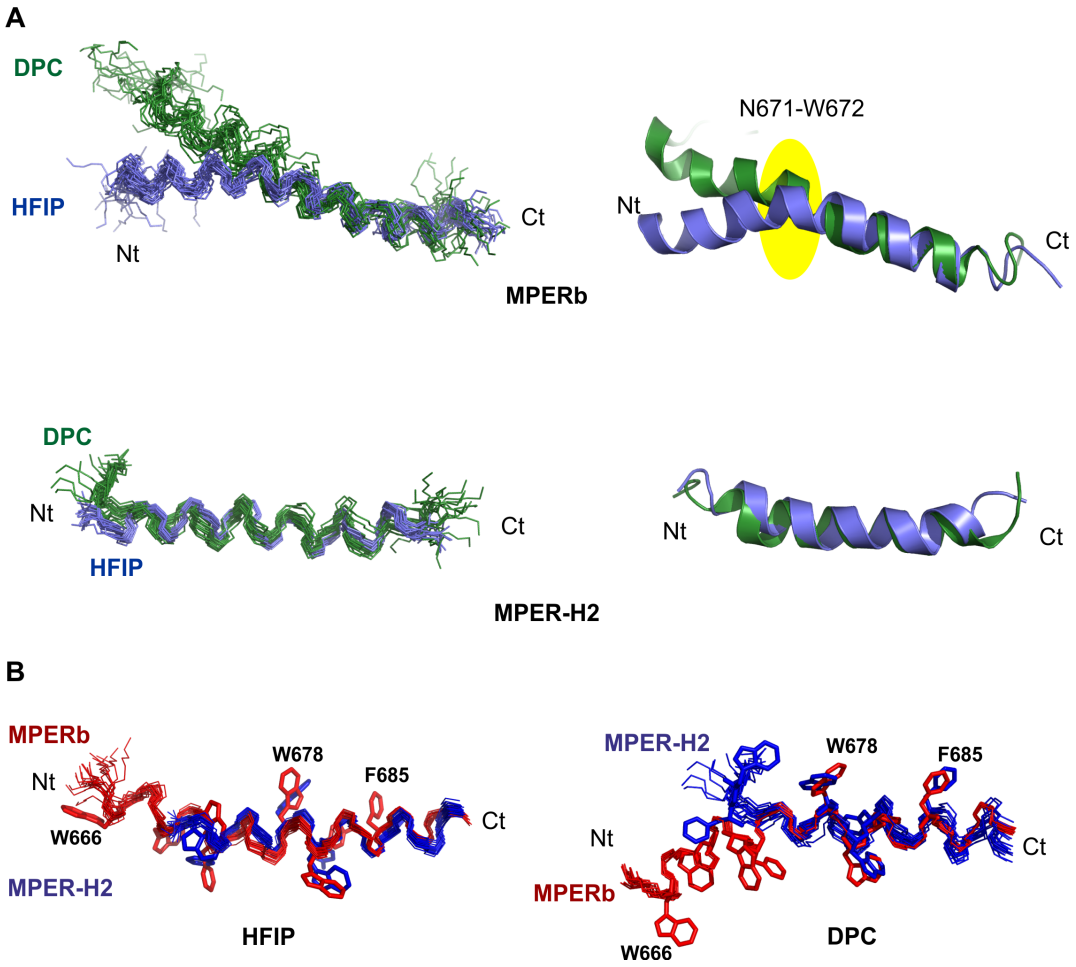


Figure 3.11. Structures adopted by MPERb and MPER-H2 in 25 % (v/v) HFIP and 20 mM DPC. A: Superposition of the 20 lowest target function conformers overlaid onto the backbone atoms of the peptide in HFIP and DPC (left), and superposition of representative NMR structures displayed in ribbon representation (right). A kink (highlighted in yellow) can be inferred for the MPERb peptide in DPC and HFIP between residues 671 and 672. B: Superposition of the 20 lowest target function structures of MPERb (red) and MPER-H2 (blue) peptides in HIFP (left) and DPC (right), Lys tags were removed and aromatic residues of the lowest target function structure are shown.

3.3.4 MPER peptides interact with membrane mimetics

To test if possible intermolecular interactions with DPC micelles take place, 2D ^{1}H-^{1}H-NOESY spectra of both peptides in a partially protonated (50 % deuterated) media of 20 mM DPC were recorded and compared with the spectra with fully deuterated DPC. Signals corresponding with the intermolecular peptide-DPC NOEs appear only in the partially deuterated DPC media (**Figure 3.12**). These intermolecular NOEs proved the presence of direct interactions between the peptides and micelles. As seen in **Figure 3.12**, Trp, Tyr and

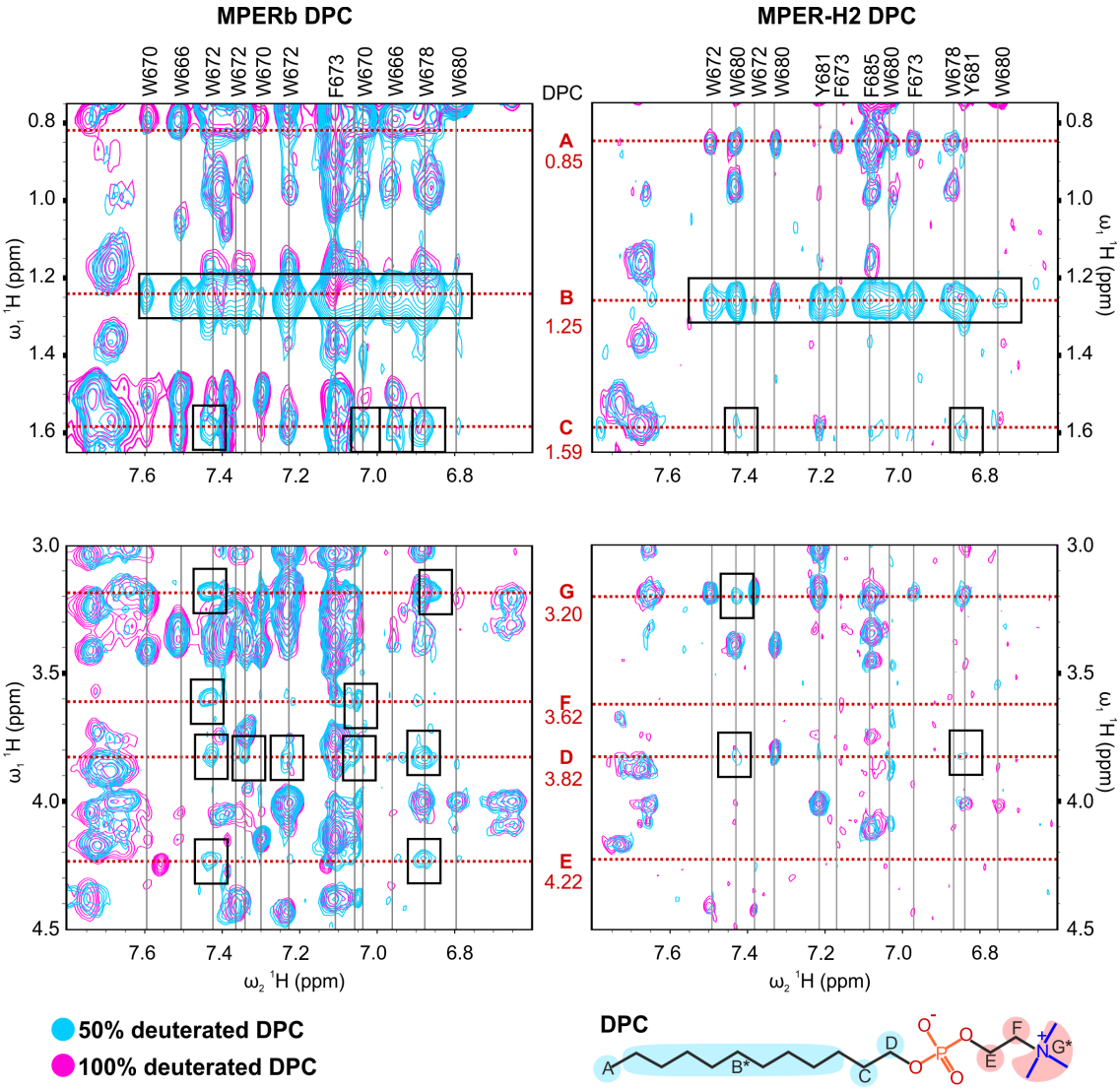


Figure 3.12. Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of MPERb and MPER-H2 in H_2O/D_2O (9:1 ratio per volume) containing 2 mM HEPES buffer, 20 mM DPC/DPC-d38 1:1, 35 ºC, pH 7.0. DPC chemical shifts are indicated with red dashed lines. The new set of blue signals (boxed) correspond with intermolecular NOEs between the DPC micelles and the peptides. * Protons with the same chemical shift.

Phe aromatic rings are crucial for the interaction to take place and the highest intensities are found between the hydrophobic tail of the micelles and the aromatic rings of Trp666, Trp670, Trp672, Trp678 and Phe673. In general, there are less intermolecular NOE signals within the polar heads than with the hydrophobic tails of the micelles in both peptides. In the MPERb peptide (**Figure 3.12**), a representative number of them belong to the aromatic ring of Trp666. The MPER-H2 peptide, which has a shorter sequence than MPERb, shows less NOE signals with the polar head of the micelle.

3.4 Discussion

As mentioned before, the highly conserved residues of the gp41 glycoprotein are essential during the fusion of the virus with the host cell membrane [163, 164]. In this study, we provide new structural insights that might help understand the relationship between structure and function of the linear helical epitope of the 10E8 antibody (MPERb peptide) from a biological point of view in membrane-like environments. The presence and distribution of certain kind of residues (aliphatic) in the MPER region, and specially, the aromatic rings of Trp amino acids, allows it to form multiple non-covalent interactions with lipids [165–168].

3.4.1 H1 and H2 in MPER are structurally independent modules

The NMR comparative and qualitative analysis of the secondary structure (**Figure 3.13**) adopted by MPERb and MPER-H2 peptides in DPC and HFIP media, reveal that the profiles are similar, this means that both peptides adopt comparable overall helical structures and indicates the absence of stabilizing interactions between H1 and H2, at least in the free state. This is also supported by the absence of long-range NOEs ($|i - j| \geq 5$). This finding validates the relevance of the information obtained individually for each peptide. The structure of helix H2, which is shared by the two peptides, is nearly identical in both peptides, as shown by the

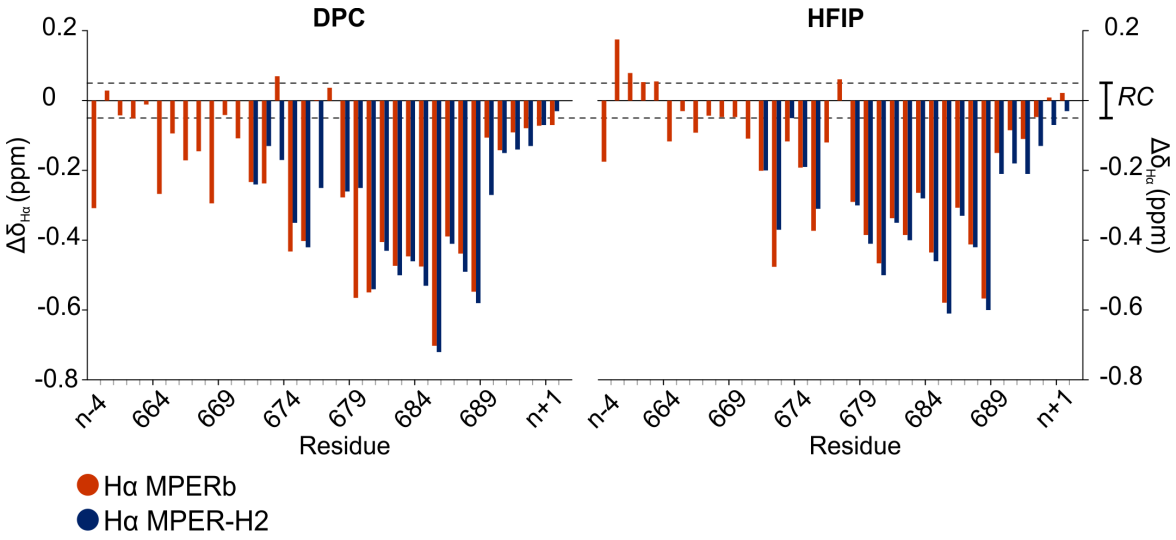


Figure 3.13. Comparison of the Hα secondary chemical shifts observed for the peptides MPERb and MPER-H2 in DPC and HFIP, pH 7.0. The profiles are similar in both peptides and are characteristic of an α-helix structure formation. Dashed lines indicate the random coil (RC) ranges.

almost perfect superposition of the backbone atoms of MPERb and MPER-H2 structures in both HFIP and DPC micelles (**Figure 3.11**). In the case of the MPERb peptide in HFIP, there is a kink in the helix that comprises the residues Asn671 and Trp672 (the transition region between H1 and H2). This might implicate that the helix is certainly flexible in this zone, making it possible to have different relative orientations between H1 and H2, and so it could play an important role during the cell membrane recognition.

3.4.2 Bending in the helices can be found in different regions depending on the solvent

The results obtained in this Thesis regarding to the MPERb and MPER-H2 complement previous studies [133, 132], in which three different overlapping peptides (**Figure 3.4**) spanning the MPER and TMD regions were characterized. The qualitative comparative analysis of the secondary structure adopted by all these peptides (MPERp, MPERb, MPER-H2, CpreTM and TMD; (**Figure 3.4** and **Figure 3.14**) revealed that all of them adopt well defined helical structures. In HFIP and DPC micelles, the profiles of the chemical shift deviations ($\Delta\delta$) for all peptides are very similar, showing only slight differences in magnitude (**Figure 3.14**), suggesting small differences in the population of ordered conformations. This is corroborated by the estimated helix populations (see **Table 3.3**). For the peptides MPERb and CpreTM, the positive $\Delta\delta$ Hα value observed for Trp678 may be explained by either chemical shift anisotropy caused by surrounding aromatic residues or it might be due to a discontinuity in the helical structure. The structures calculated for all peptides exhibited well defined backbones, and showed little variability in the orientations of the side chains (**Figure 3.11**, see further pages).

Based on the structures obtained for all peptides (those in the past and the ones solved in this study), a congruent model of the overlay can be assembled (**Figure 3.15**). It should be mentioned that the structures in HFIP are good models for the pre-hairpin state, whereas those in DPC are usually accepted as the structure adopted in the later hairpin step of the membrane fusion mechanism [133, 132, 130]. As seen in the figure, the superposition of all structures in HFIP is consistent, and even if the side chains of the aromatic residues present certain degree of flexibility, their positions are still close in the different peptides. The superposition of the structures in DPC does not include the TMD region, given that the structure of the peptide TMDp could not be determined in DPC. Nevertheless, the superposition of the rest of the peptide structures in DPC is also consistent, except for the residues spanning Asn671 and Ile675, which are partially extended in CpreTM and MPER-H2; and helical, though in a

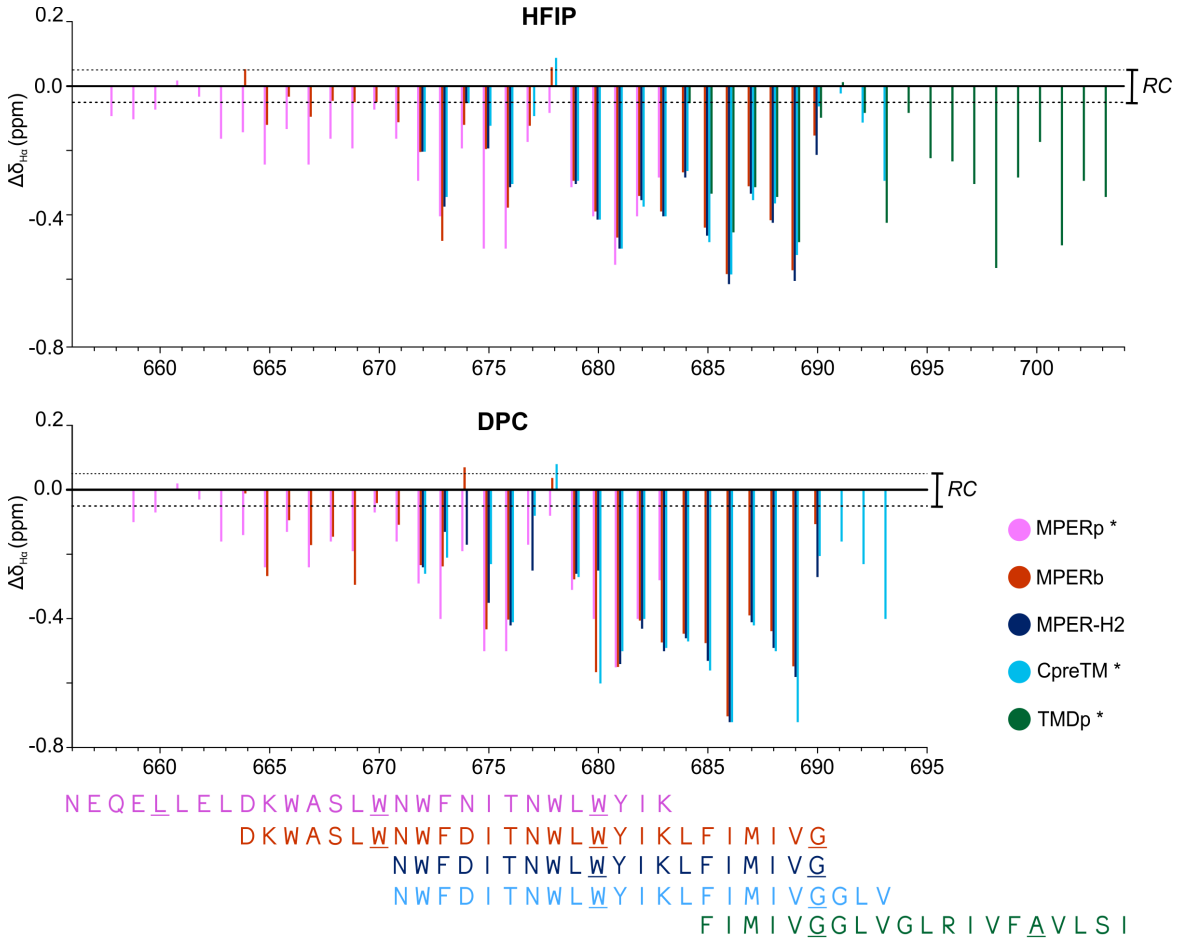


Figure 3.14. Comparison of the Hα secondary chemical shifts observed for all the peptides studied previously* and in this study in 25 % HFIP and 20 mM DPC micelles, pH 7.0. The profiles are similar for all peptides and the differences found are in magnitude, characteristic of an α-helix structure formation. Dashed lines indicate the random coil (RC) ranges. For clarity, every tenth position of the sequence is underlined.
* Peptides previously studied by our group [133, 132, 130].

kink, in the longer MPERp and MPERb peptides. These differences are likely explained by the existence of stabilizing interactions provided by precedent residues present in the longer peptides that are absent in the shorter peptides, in which these residues are located at their N-terminus segment. Concerning MPERp and MPERb peptides, the α-helix found in DPC suffers a kink with an angle of 28 º between residues Asp674 and Ile675, at the beginning of the H2 region of MPER, which is not observed in HFIP. However, the structure in HFIP shows a kink with an angle of 21 º between Leu692 and Val693, that corresponds to the TMD region in the TMDp peptide. As a whole, the MPER-TMD region contains several points where the helical structures can bend or not depending on the conditions, which very likely are important for the different stages of the membrane fusion process leading to infection.

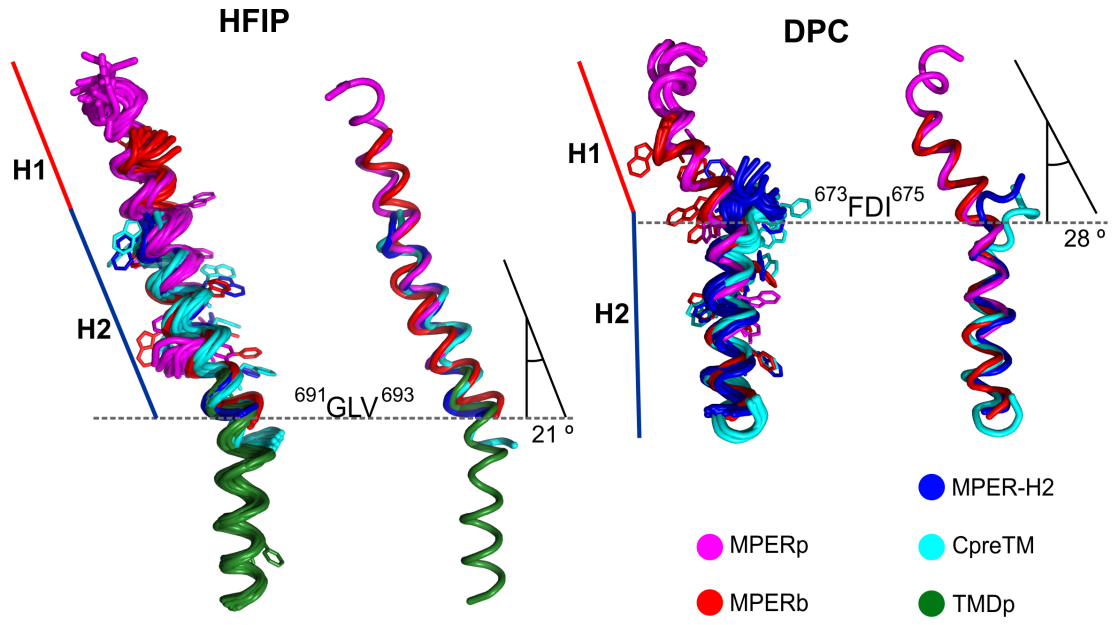


Figure 3.15. Models for gp41 MPER-TMD organization in HFIP and DPC micelles as inferred from the calculated NMR structures. Left: Overall structural model of the MPER-TMD regions deduced from the combination of the 20 lowest target function NMR structures of MPERp (magenta, PDB: 2M8M), MPERb (red, PDB: 2NCT), MPER-H2 (dark blue), CpreTM (cyan, PDB: 2MG2) and TMDp (green, PDB: 2MG1) in 25 % HFIP. The kink at the ^{691}GLV^{693} location can be inferred (dashed lines). Right: Overall structural model of the MPER-TMD regions deduced from the combination of the 20 lowest target function NMR structures of MPERp (magenta, PDB: 2M8O), MPERb (red, PDB: 2NCS), MPER-H2 (dark blue) and CpreTM (cyan, PDB: 2MG3) in 20 mM DPC. The kink at the ^{673}FDI^{675} location (dashed lines) can be inferred, the NMR structure of TMDp in DPC could not be obtained. H1 and H2 segments are also indicated.

3.4.3 Peptide-membrane interaction and insertion

Combining the analysis of the spectra with the partially deuterated mix of DPC micelles and the previous results, the model of the interaction of the helices in the membrane can be refined adding new information derived from the NOEs between DPC and the peptides MPERb and MPER-H2 (**Figure 3.12**). For MPERb, NOE data is in accordance with the proposal that the position of the residues belonging to the N-terminus are closer to the polar heads, and the residues closer to the C-terminus get embedded into the hydrophobic tails of the micelle, and then, anchored into the membrane interacting with the hydrophobic tails of the lipids [169, 133]. In the case of the MPER-H2 peptide, the first seven residues contiguous to the N-terminus of MPERb are not present. Therefore, its sequence lacks certain important amino acids for membrane interaction, like Trp666 and Trp670. This can be the explanation for the few intermolecular NOE signals observed with the polar heads of the micelle.

In this regard, it should be mentioned that the interaction of peptides with micelles is a dynamic event and probably different complexes are in equilibrium in the media [170]. This is important from the biological point of view, as membrane fluidity plays a key role in the insertion mechanism of several processes, including infection [171]. On the other side, and because of this dynamic scenario, the information derived from the observed NOEs cannot discriminate if the insertion of these short peptides is the same as the longest ones, like TMDp, which contains the transmembrane domain. Unfortunately, the NMR spectra of TMDp could not be assigned in DPC micelles, so no information about intermolecular NOEs could be obtained for this peptide. Therefore, the model of the insertion of the helices within membranes is susceptible to be refined at atomic resolution if new data is available in the future.

3.4.4 The structural flexibility found in the peptides might contribute to the membrane fusion mechanism

In previous studies, the combination of both NMR experiments [84] and X-ray crystallography analysis [133] suggests that the kink between H1 and H2 regions at positions Asn671 and Trp672 is the structural trigger of the membrane fusion mechanism by gp41. Though, in DPC media, the kink found comprised the ^{673}FDI^{675} region, suggesting that during the hairpin formation, this flexible region might thermodynamically favor the changes towards the formation of the six helix-bundle.

Considering the structural features of all MPER peptides examined up to now, the kinks observed in the free peptides sometimes are maintained in the bound states, and sometimes not. Different kinks are also characterized depending on the peptide or the media. These variable results can be just the description of the dynamic nature of these sequences, making them flexible enough to contribute to connect the 6-helix bundle with the viral membrane during the fusion process (**Figure 3.2**) and once the recognition is established, the helices might get straightened.

3.4.5 Broadly neutralizing antibodies bind differently to distinct MPER epitopes

According to previous studies, the preservation of a continuous MPER-N-TMD helix insertion should be key for the MPER epitopes to be recognized by the antibodies [133]. Therefore, the structural information of the MPERb epitope peptides bound to antibodies is essential for

the correct understanding of the neutralizing process. In this regard, in collaboration with other groups, the complexes of peptides covering this region bound to the 10E8 and 4E10 Fab antibodies were crystalized and their structures were solved by X-ray crystallography. Thus, on the basis of the X-ray structures of the complexes and the NMR structures of the free peptides commented before, a model of the recognition of this epitope by the 10E8 and 4E10 antibodies have been proposed [84] (**Figure 3.16 A** and **B**).

In the crystal structure obtained with the 10E8 antibody and the MPERb peptide, the interaction region is well defined, with specific residues of the MPER epitope (Trp678, Trp680, Tyr681 and Phe685) located nearby certain antibody residues (Tyr105, Phe107, Trp108 and Tyr111). Interestingly, residues between Asp664 and Ser668 (H1 region) of MPERb adopted a dynamic state and were not observed in the electron density map. Consequently, as the recognition does not fix the position of the H1 helix, we can propose that this region might not be essential during the recognition of the Env complex by the antibody. This fact supports the finding that the kink found by NMR (at Phe673-Ile675 position) might give certain flexibility that allows H1 to adopt random positions, at least in the presence of the Fab molecule. The H2 region adopts a helical conformation in which the Gly690 was also unobserved in the electron density, thus, in accordance with a dynamic nature. Interestingly, residues between Leu669 and Asn671 did not show the characteristic kink found for the free peptide. One possibility may be that the inherent flexibility found by these residues in the free state might allow the plasticity needed by the antibody to bind to the epitope and, once bound, the recognition might tend to produce a slight conformational change and straighten that portion of the helix.

In the case of the 4E10 antibody, the arrangement of the complex with MPER changes, but the peptide maintains the same structure as in the free state. An interaction map can be traced as well, defining a different subset of residues in MPER (Trp666, Trp670 and Phe673) involved in the interaction with the antibody (Tyr32 and Phe100). The orientation of the peptide is different with respect to the previous complex (estimated difference of 30º), and no random structures are present (**Figure 3.16**). It is also interesting to note that the binding affinity of the MPER-H2 peptide with the 10E8 antibody is essentially equal compared to the peptide MPERb, which contains the H1 helix that is dynamic in the complex with the antibody (K_d values measured by our collaborators by ITC are 9.6 ± 1.0 nM for MPERb and 10.0 ± 1.6 nM for MPER-H2 [84]).

The 2F5 antibody was also crystalized with the MPERp peptide. The interaction interface can be drawn with the identified active residues of MPER (Trp666 and Trp670) and those in the antibody (Pro98 and Arg100) (**Figure 3.16 C**). The crystal structure of the complex

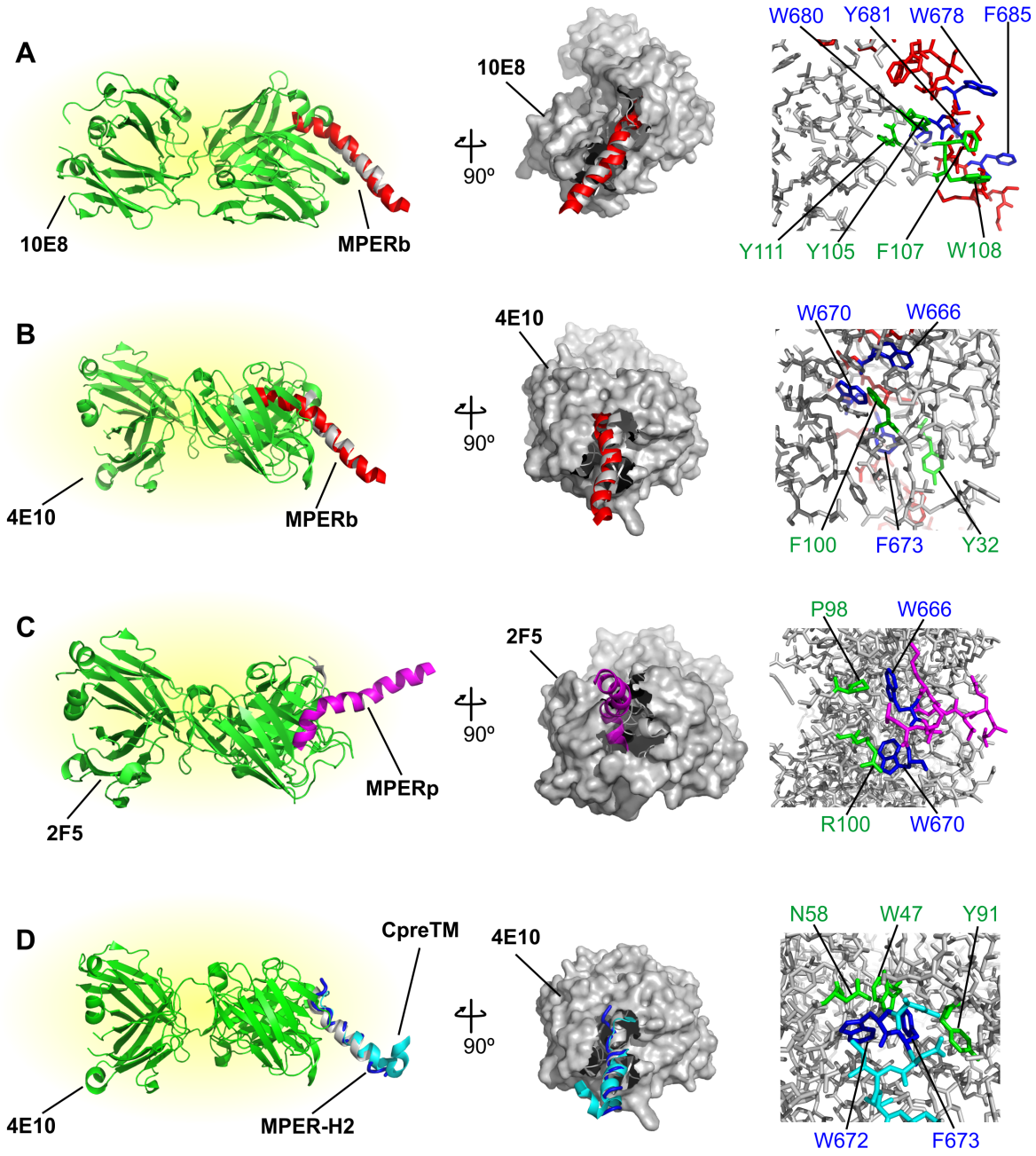


Figure 3.16. 10E8, 4E10 and 2F5 epitope organization and putative mechanisms of antibody recognition. The lowest target function NMR structure obtained in DPC micelles is fitted into its corresponding Fab-bound crystal structure (depicted in green and gray). A: MPERb peptide (red) and the Fab-bound structure of the 10E8 antibody (PDB code 4G6F). B: MPERb peptide (red) and the Fab-bound structure of the 4E10 antibody (PDB code 4WY7). C: MPERp peptide (magenta) and the Fab-bound structure of the 2F5 antibody (PDB code 3D0L). D: MPER-H2 (dark blue) and CpreTM (cyan) peptides and the Fab-bound structure of the 4E10 antibody (PDB code 4WY7). Residues involved in the antibody and peptide interaction are depicted in green and blue, respectively.

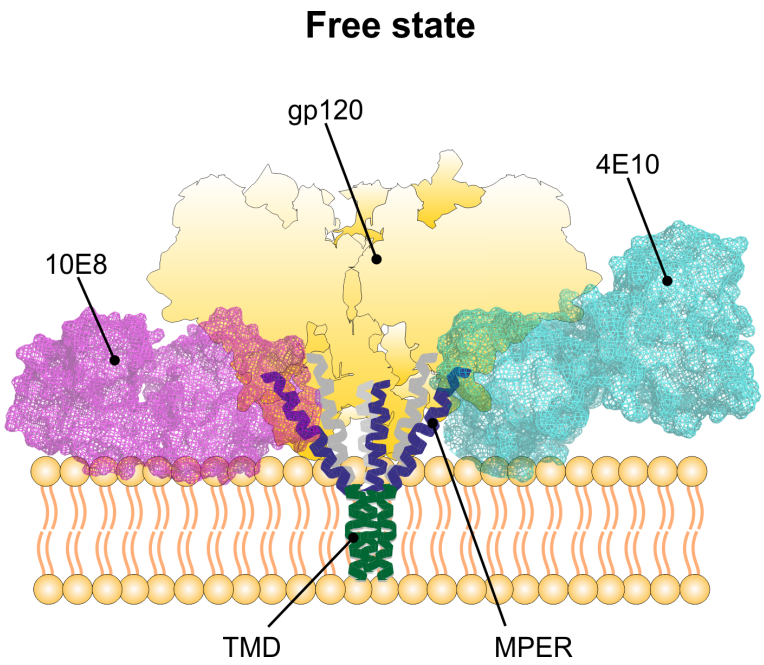


Figure 3.17. Schematic representation of the Env complex and its antibody neutralization. Based on the combined structures calculated in HFIP and DPC from the NMR data, the overall structure during the pre-hairpin state might be present to bind the six-helix bundle (6-HB) with the membrane and the kink located at the ^{691}GLV^{693} region might bring the flexibility required to bind the membranes. The 10E8 (magenta) and 4E10 (cyan) epitope binding at the viral interface blocks the membrane fusion process.

shows that, like in the isolated state, the peptide adopts a helical structure interrupted by a bend between the residues Asp664 and Trp666 (**Figure 3.16 C**) [132]. Finally, the 4E10 antibody bound to the CpreTM peptide was also crystalized (**Figure 3.16 D**) and the structure reveals that the segment comprising the residues Asn671 and Phe673 also has to reorient itself in order to bind with the antibody. In this case, crucial residues of the recognition site are Trp672 and Phe673 of MPER, and Trp47, Asn58 and Tyr91of the antibody.

In summary, different antibodies bind to the MPER region in various ways, they approach to the helices with different orientations, and the peptides sometimes suffer conformational changes. Nevertheless, in all cases, the role of the aromatic rings in the stabilization of the complexes is important. We can use all the available information to construct a picture of the involvement of gp41 during the viral infection and its antibody neutralization. In this model, regarding the different orientation of the peptide helix in the complexes, we can propose that the 10E8 antibody might approach parallel to the membrane plane, while 4E10 with an angle of difference of 30º with respect to 10E8 (**Figure 3.17**). This means that 10E8 might also contact with the surface of the viral membrane when bound to its MPER epitope. Further fluorescence assays [84] confirmed that the 10E8 antibody binds specifically in the MPER H2

region, in which Trp680 can be localized in the interface between the antibodies (10E8 or 4E10) and the viral membrane. It can also be proposed that the flexibility located in the H1 region might help to reduce the entropic cost of this binding process. Linking all these results, it can stand out that the general flexibility and the aromatic residues are key to determine how the antibodies contact their epitopes at the membrane interface and to define which epitopes can be reachable for the antibody recognition and block the infective process (**Figure 3.17**).

3.5 Conclusions

1. Given the limitations of solution NMR to obtain full structural information of large complexes, the use of simplified systems is required. In this study, the correct design of peptides belonging to the MPER regions of gp41 and their NMR structural study in two different media (a non-polar cosolvent and DPC micelles) provide insights into the system on its biological context.

2. In both HFIP and DPC micelles, the peptides adopt helical structures with kinks located in different regions. This variability in the location of the kinks might be the source for the flexibility that lets the virus to bind with the T-cell membrane and allow hairpin formation.

3. The presence of aromatic residues spanning the helix was known to be important for the affinity of gp41 with membranes. The most important elements identified here that directly interact with membranes belong to the aromatic rings of Trp, Tyr and Phe amino acids.

4. In this study, the structure of the epitope for the broad neutralization antibody 10E8 was successfully obtained and analyzed as a simplified version, combining these results with previous experiments and other biophysical techniques, a model of the full system could be proposed.

5. As the gp41 glycoprotein is a highly conserved sequence in HIV and plays an important role during the T-cell membrane fusion, the employment of antibodies targeting gp41 epitopes is an interesting therapeutic approach. The structural information obtained regarding to crucial regions for the antibody recognition may elucidate how the neutralization of the virus by these antibodies might occur.

6. We propose that new broadly neutralizing antibodies targeting gp41 can be designed based on the characteristics of the MPER-TMD regions analyzed in this study. A proposition can be that new antibodies should bind specifically to the MPER region that is transiently exposed to the solvent at the membrane surface, taking into account the epitope's orientation, depth and angle of penetration into the membrane.

4

Structure and molecular association in sequences derived from the antimicrobial peptide protegrin-1 in the presence of membrane mimetics

4.1 Introduction and objectives

The antibiotic era started at 1930's and 1940's when humans discovered the wide potential of certain molecules (antibiotics) to kill microorganisms and cure infections. Since then, antibiotics have been widely used without sufficient control (inappropriate treatment and overuse), and microorganisms had started to adapt to this new evolutionary pressure. Thus, the best adapted microbes able to resist the action of the antibiotics, those with physiologically or genetically enhanced capacity to survive high doses of antibiotics, are selected to gain or maintain their infective potential [172]. As a consequence, antibodies may result in preferential growth of resistant microorganisms, while the drug inhibits the growth of the susceptible ones.

Nowadays, the problem concerning the rising of antibiotic resistance is critical from a health point of view. As the pharmaceutical industry fails to act, it will mean that the treatment options for many patients will be severely limited. It is becoming urgent to return to the situation when the antibiotic era started, when no or few resistance was present, and in which the only 20 different classes of known antibiotics worked efficiently during the following 60 years [173]. In this regard, antimicrobial peptides (AMPs) have been demonstrated to be a potential solution to successful antibiotic discovery and production [174].

4.1.1 The antibiotic crisis

As mentioned before, the negligent abuse of antibiotics, the unawareness of the warnings against their overuse and poor control practices created a high selection over pathogens. The resistant microorganisms developed immunity against drugs, and consequently, the current infection treatments became more challenging [175–177]. Even when some microorganisms were naturally capable to overcome these molecules before the antibiotic era, its abuse provoked the survival of new resistant pathogens within the previously susceptible ones [172]. The importance of a correct control program has been demonstrated previously in England, where the application of good nursing practices through prudently prescriptions of a combined old and new antibiotics, prophylactic measures of constant hand washing and good infection controls lead to a dramatic decrease of Meticillin-Resistant *Staphylococcus aureus* (MRSA) and *Clostridium difficile* infections in 2010 [178, 179].

Some of the most relevant antibiotic resistant pathogens have been announced by the Centers for Disease Control and Prevention quite recently [180]. The report includes pathogens that can be spread to humans or animals such as MRSA, Glycopeptide-Resistant *S. aureus*,

Toxin Hyperproducing *C. difficile*, Extended-Spectrum β-lactamase- and carbapenemase-producing coliforms [180, 181]. These resistant microorganisms increased their virulence and are a serious sanitary and veterinary problem worldwide. This is an important concern as it causes a raise in mortality, especially within hospital surroundings [175, 182].

4.1.2 Antimicrobial peptides

Diverse peptides with antimicrobial activity have been found naturally acting as a host defense against viruses, fungi, parasites and bacteria [183, 184]. There are currently more than 5500 antimicrobial sequences reported, from which more than 1600 belong to synthetic peptides [185, 174]. In 1939, the first known isolated peptide with antimicrobial activity was obtained from a soil mixture of microorganisms. A culture of an identified Gram-positive aerobic bacillus released a soluble agent capable of lysing other Gram-positive microbial species [186]. In the following year, the first performed *in vivo* experiment showed that mice treated with the antimicrobial agent, called gramicidin, were protected against an intraperitoneal administered dose of a virulent Type I pneumococci G. Also, The *Pneumococcus* culture got sterilized 1 hour after the addition of gramicidin [187]. Later on, diverse AMPs were identified and described. However, once discovered, their potential applications in clinic were studied and their toxicity, such as against human blood cells *in vitro* and on mice through peritoneal administration was also documented [187, 188]. Despite their potential toxicity, the clinical use of AMPs as antimicrobial agents gained interest and it is still under study and consideration [189, 190].

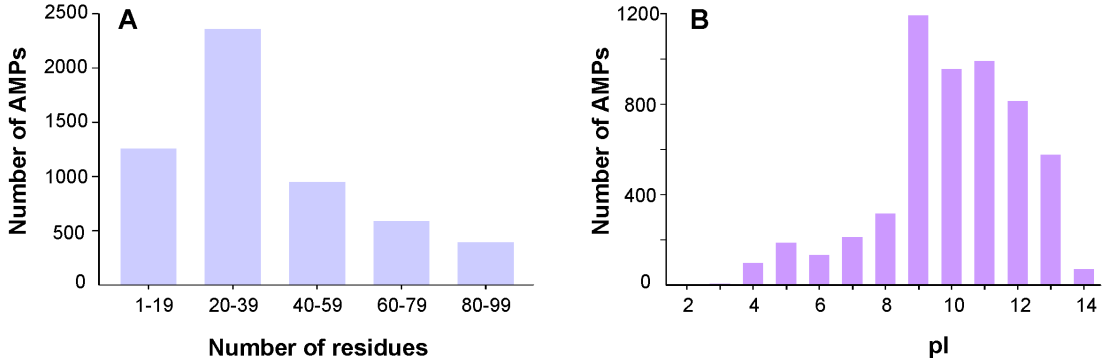


Figure 4.1. Current AMPs statistical distribution based on number of residues and isoelectric point. A) The most frequent peptide size ranges from 20 to 39 residues and is followed by 1 to 19 amino acids. The less frequent distribution corresponds with peptides between 80 to 99 residues. B) 4598 over 5547 AMPs have an isoelectric point of 9 or higher, this means that at pH 7.0 or below, more than 80 % of the peptides in the database posses a positive net charge. [174]

AMPs size can vary from five to over a hundred of amino acids (**Figure 4.1**) [174] and their commonly cationic nature allows them to interact with the cell membrane, disrupting its permeability and causing cell death. Unlikely antibiotics, AMPs typically do not target against a defined cellular activity, but through a more ubiquitous system. Therefore, taking advantage of this general property, they are widely proposed as potential therapeutics facing bacterial resistance [191, 183, 192]. In this regard, it is important to optimize AMP sequences to obtain the best activity, to avoid their potential toxicity to humans, bacterial resistance, and lack of specificity and stability. For these reasons, AMPs are committed to be mutated [193–198].

Structurally, these antimicrobial sequences can adopt any of the classical secondary structures known for proteins and peptides (**Figure 4.2**). They might be found as β-sheets, α-helices, or extended and loop arrangements, but the majority of them are considered to form amphiphilic helices or β-sheets [189, 199–201], being the helical peptides the most widely studied [189], probably because of their physicochemical properties like solubility and stability. The major sources of high-resolution structural information can be obtained by X-ray crystallography or NMR spectroscopy, however, given the size and intrinsic flexibility of these peptides, the use of NMR spectroscopy predominates over the X-ray techniques [202].

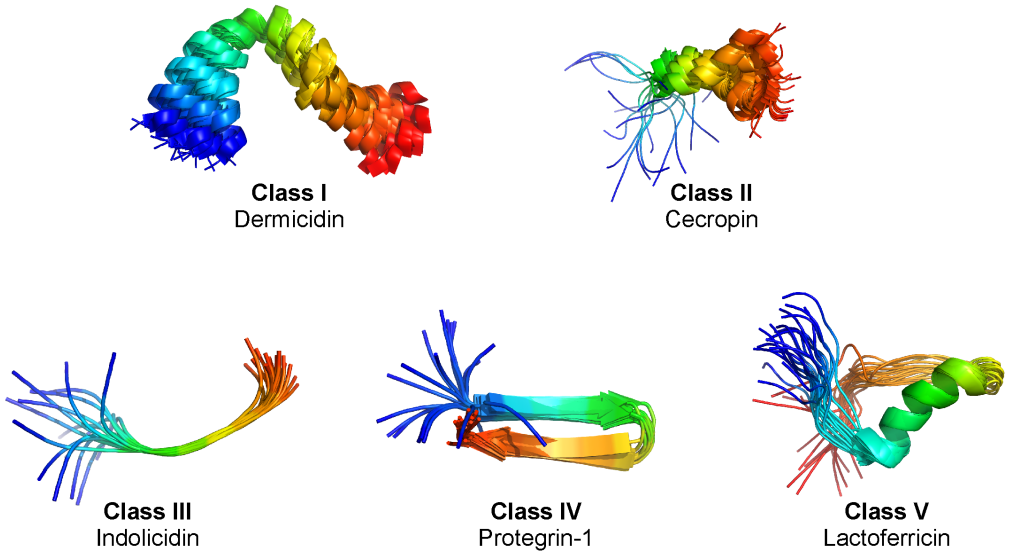


Figure 4.2. Examples of the different classes found in antimicrobial peptides. Class I includes dermicidin (PDB 2NDK), class II cecropins (PDB 2N92), class III indolicidin (PDB 1G8C), class IV protegrin-1 (PDB 1PG1) and class V lactoferricin (PDB 1Z6V) from lactoferrin.
* Structures are not in the same scale.

4.1.3 Classification

AMPs can be classified by different criteria: either structurally or functionally. From a structural point of view, AMPs can be categorized into five main types: α-helices, β-sheet, cysteine-rich, with an unusually high content of a common amino acid and peptides with uncommon and/or chemically-modified amino acids [202]. Some AMPs do not belong to any of these classes and others can possess more than one structural characteristic [203, 204]. Taking into account the amino acid composition and structure, five different classes (Class I to V) can be defined as seen in **Table 4.1** and **Figure 4.2** [205].

AMPs can also be grouped depending on their activity or their targets, such as antiviral, antifungal, antiparasitic or antibacterial [189]; as well, some AMPs display antitumoral properties [185, 174]. Antiviral peptides are capable to neutralize virus infections by integrating

Table 4.1. Classification of antimicrobial peptides based on their amino acid composition and structure adopted.

Class	Description	Examples
I	Anionic peptides, small anionic peptides rich in glutamic and aspartic acids from sheep, cattle and human proteins	Dermicidin, Maximin H5
II	Linear cationic α-helical peptides, short peptides with differentiated hydrophilic or hydrophobic regions; their structure can change from unstructured to helical in micelle media	BP100, Cecropins, Magainin, Pleurocidin, CAP18
III	Cationic peptides enriched with specific amino acids. Pro, His, Trp, Arg or Gly can be over-represented in the sequence	Indolicidin, Abaecin, Pyrrhocoricin
IV	Anionic and cationic peptides that contain cysteine and form disulfide bonds. They may contain several disulfide bonds and form a β-sheet structure.	Protegrin-1, Defensin A, HNP-1
V	Anionic and cationic peptides that are fragments of larger proteins, as they are part of certain proteins, their biological role in immunity is unclear.	Lactoferricin from lactoferrin, Casocidin I from casein

themselves into the host cell membrane or into the viral envelope, causing membrane instability or reducing the virus' binding capability [206–211]. Antifungal peptides can disrupt the fungal cell wall through its permeabilization [212] or by pore formation [213]. The antiparasitic peptides, which belong to the less represented group [189, 185], are also capable to kill their targets via interaction with the cell membrane [214]. In counterpart, antibacterial peptides are the most studied AMPs. Most of them have cationic sequences and usually amphipathic structures, being these properties important for function.

4.1.4 Mechanisms of action of AMPs

Several different mechanisms of action have been proposed for AMPs. Some of them can cause the fragmentation of the lipid bilayer through membrane permeabilization in some organisms [215, 216]. At the same time, in other pathogens they can cause the interruption of essential biological processes, such as protein and DNA synthesis, DNA replication, and other metabolic pathways [205, 217, 218].

In this scenario, the principal mechanism of action studied in AMPs consists of the disruption of the membrane homeostasis, known as membrane-active AMPs. However, this is a complex mechanism that is not completely understood. From the mechanistic point of view, the biological consequences of the membrane destabilization has not a single foundation and probably is the sum of at least two different effects: membrane disruption and enzymatic activity against cell machinery. It is known that the bacterial membrane (**Figure 4.3**) harvests one third of the total proteins of the cell and that they have different functions that are critical for survival [219], including among others: active transport of nutrients, respiration, proton

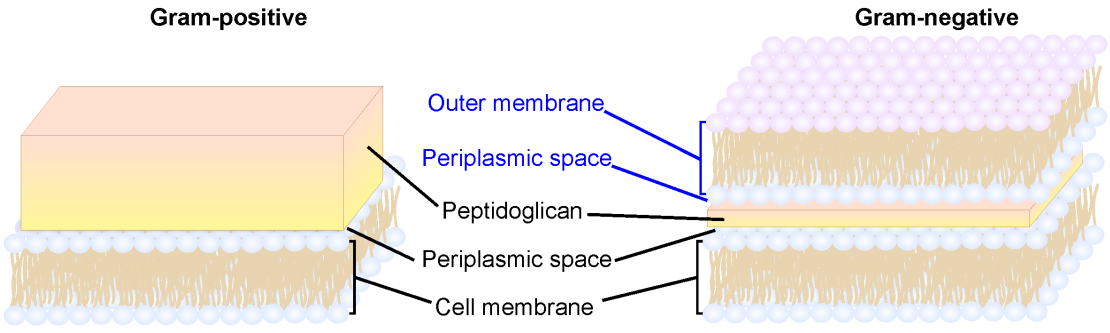


Figure 4.3. Cartoon representation of Gram-positive and Gram-negative bacterial cell membranes. Gram-positive bacteria posses a plasmatic cell membrane, followed by a periplasmic space and covered with a thick peptidoglycan layer. Gram-negative bacteria are composed by a plasmatic cell membrane, followed by a periplasmic space, the peptidoglycan layer is thinner and is followed by the periplasmic space and an outer lipid membrane.

motive forces, ATP generation, and intercellular communication. The function of these proteins can be altered with AMP treatment, implying that AMPs' rapid killing effect is also consequence of this circumstance and not only due to the physical disruption of the membrane [189].

4.1.4.1 Membrane-active AMPs

Given that organisms possess specific membrane compositions, the targets of membrane-active AMPs are usually defined by their interactions with the cell membrane. As we have mentioned before, AMPs can be amphipathic, and so are the majority of membrane-active AMPs. Thus, an initial electrostatic interaction between a positively charged surface of the AMP and a negatively charged cell membrane may occur. Then, the hydrophobic region of the AMP will permit the insertion of the molecule into the cell membrane and afterwards, will cause a membrane destabilization [220].

On the basis of this general process, three main mechanisms of action are proposed [189] as represented in **Figure 4.4**. In the toroidal pore model, AMPs align perpendicularly into the membrane, the polar heads of the lipids interact with the hydrophilic portions of the peptides and their hydrophobic regions get faced, forming a pore. The barrel-stave model is similar to the toroidal pore, but in this case, the staves are proposed to first align parallel to the cell membrane, and subsequently, the barrels are formed and get inserted perpendicularly into the membrane bilayer. The carpet like or detergent like model proposes that the peptides touch the membrane and form a coating over it, then, the molecules penetrate into the lipid bilayer and surround the lipids, forming a micelle and leaving holes. The most commonly accepted mechanisms are the barrel-stave and toroidal pore model, though, the barrel-stave model seems to be unsuitable for the cationic antimicrobial peptides found in eukaryotes [202].

Other less widely accepted models have also been proposed (**Figure 4.5**). In the plane diffusion model, peptide monomers get surrounded by the membrane lipids and induce local membrane perturbations. Consequently, transient pores are formed [202]. In the sinking raft model, the peptides aggregate in the outer face of the membrane bilayer and induce local imbalances, a curvature of the membrane is shaped and peptides tend to redistribute either in the inner or the outer face, producing transient pores while passing [221]. Finally, the electroporation model is similar to the last one, but the cationic peptides produce a potential imbalance, therefore, peptides need to pass through the bilayer to restore the charges [222].

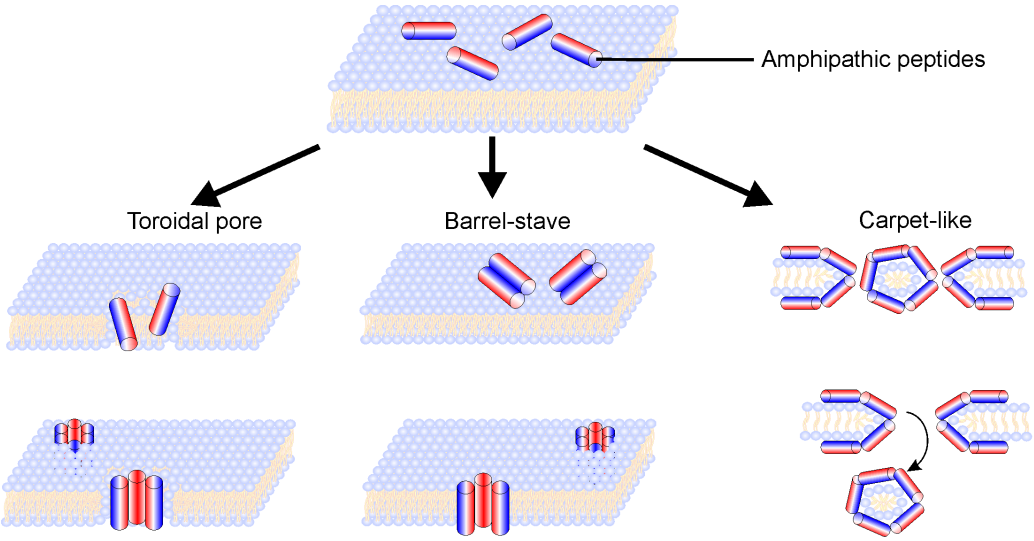


Figure 4.4. Schematic representation of the most popular mechanisms of action of membrane-active AMPs. In the toroidal pore model, the peptides insert into the membrane perpendicularly, and keep in contact with the hydrophilic heads of the lipids. In the barrel-stave model, peptides are disposed like in the toroidal pore model, but the hydrophobic tails of the lipids contact with the hydrophobic face of the peptides. In the carpet-like or detergent-like model, small regions of the membrane are coated by the peptides and create pores by the excision of a micellar section.

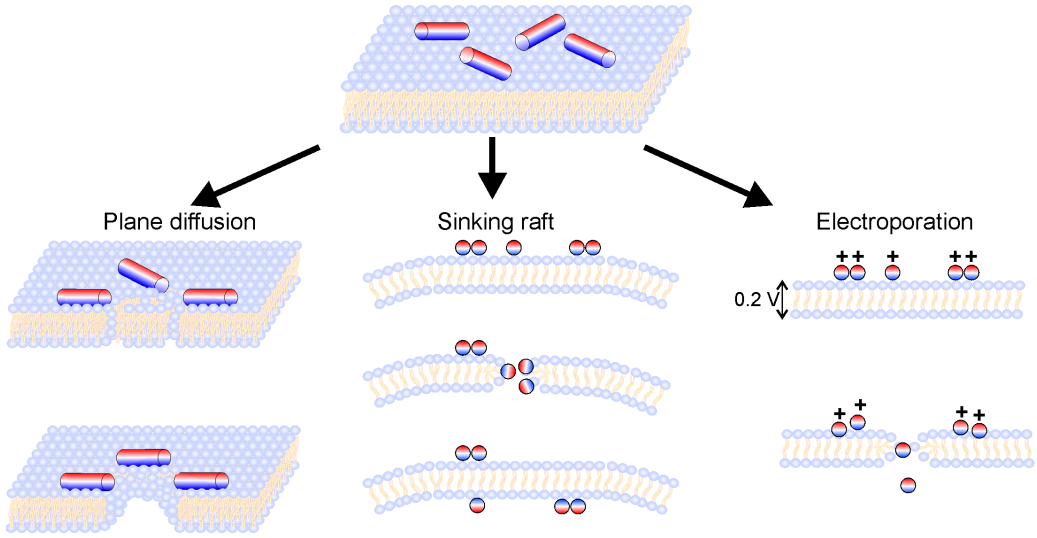


Figure 4.5. Schematic representation of other mechanisms of action of membrane-active AMPs. In the plane diffusion model, the peptides induce local perturbations in the membrane and a pore is formed. In the sinking raft model, the peptides aggregate and induce local curvatures in the membrane; this leads to the redistribution of the AMPs in both faces of the bilayer. The electroporation model proposes that an electric potential is created by the peptides that induce pore formation to restore the potentials.

4.1.4.2 Intracellularly-active AMPs

Although membrane destabilization is the most common known mechanism for AMPs, it was shown that some of these molecules could kill their target cells without causing membrane permeabilization. This suggested that there are other mechanisms of killing cells by AMPs. In fact, certain AMPs do not interact directly with the membrane of their targets; some of them can be internalized through endocytosis, including macropynocytosis and receptor-mediated endocytosis [223, 220]. Once internalized into the cell, these AMPs can take some specific actions, including (as mentioned before) the inhibition of DNA, RNA and protein synthesis [224, 225]. There are also reported peptides that can have multiple targets, such as seminalplasmin, which inhibits RNA polymerase [226] and can activate an autolysin protein too, both cellular processes lead to cell death [227, 228]. In the case of diptericin, this peptide can kill only during the bacterial growth stage, which means that certain developmental metabolic pathways might be interrupted [229].

4.1.5 Important features of AMPs for structure-function relationships

A relationship between structural properties and the mechanism of action have not been found yet. Nevertheless, certain characteristics are important to take into account during the design and testing of new antimicrobial agents. The most relevant physicochemical properties include the peptide's size, charge, hydrophobicity, amphipathicity and solubility [189].

It was commented before that AMPs present secondary structures *per se*, but also upon membrane interaction. In this regard, the peptide length is important for structure formation. Thus, the minimal length required for β-sheet formation is eight amino acids, while for α-helices the estimated number rises to around 22 [40, 230]. The length of the peptide may also affect its potential cytotoxicity. In the case of melittin, a shortened version presents 300 times less toxicity against human erythrocytes compared with the wild type [231].

The net charge of the peptide is also important for their activity as it can determine its potential toxicity against host cells. At the same time, it defines the kind of target through its capacity to interact depending on the lipid composition of the microbial membrane [189, 202].

The hydrophobicity is also a crucial feature; it can determine the range of targets [232] and also, modify its antimicrobial capacity [233]. Almost 50 % of the amino acids present in the sequence of natural AMPs are hydrophobic [234] and an increased number of the hydrophobic residues within the positively charged region of an AMP might increase its activity against membranes [233].

The amphipathicity is one of the key properties that AMPs should have in order to bind properly into the microbial membranes. As membranes are also amphipathic, the peptide should be able to interact with either the polar heads or the hydrophobic tails of lipids [235].

Finally, peptides should be soluble enough in aqueous environments to make them able to interact with cell membranes. If aggregation occurs in the aqueous media prior to the interaction with the membrane, the oligomers might not be able to pass through the lipid bilayer [189]. In summary, it is important to take all these properties in consideration towards the design of new therapeutics based on AMPs.

4.1.6 Protegrins: a promising family of antimicrobial peptides

There is a huge range of possible peptides with antimicrobial properties, starting from naturally isolated, to synthetic ones [174].

Amid the wide range of AMPs described in the literature, protegrins (PG) comprise a family of AMPs that were firstly isolated from porcine leukocytes. Up to now, there are five known protegrins. They are cysteine-rich cationic peptides in which their positive charges are prone to interact and disrupt the membrane permeability probably by forming pores and thus, leading to microbial death [236–239]. Protegrins are composed by 16 to 18 residues with a β-hairpin structure stabilized by two disulfide bonds (**Figure 4.6**) [236, 240]. Due to their cationic nature, protegrins are prone to be absorbed on the surface of anionic membranes. It is known that PG-1 tends to dimerize in the presence of lipids; therefore, it is thought to kill bacteria through the formation of a toroidal or a barrel-stave pore upon oligomerization [241, 242]. Currently, there are two models proposed for PG-1 dimerization depending on the membrane composition, with either a parallel dimer formation in membranes containing cholesterol [242] or an antiparallel association within DPC micelles [243].

Peptide	PDB code	Aligned sequence	Number of residues
PG-1	1PG1	RGGRLCYCRRRFCVCVGR	18
PG-5	2NC7	RGGRLCYCRPRFCVCVGR	18
PG-3	2MZ6	RGGGLCYCRRRFCVCVGR	18
PG-2	2MUH	RGGRLCYCRRRFCICV●●	16
PG-4	-	RGGRLCYCRGWICFCVGR	18

Figure 4.6. Alignment of sequences of protegrins (PG) 1 to 5. All protegrins have four Cys residues in positions 6, 8, 13 and 15 with two disulfide bonds formed between residues 6-15 and 8-13. Different residues compared with PG-1 are highlighted in green. Every fifth position is highlighted in yellow.

4.1.7 Previous structural studies on protegrins

Previous studies have shown that protegrins PG-1, PG-2, PG-3 and PG-5 perform a similar effectiveness as antimicrobial agents; being 12 residues enough to trigger their antimicrobial effects and 4 additional residues (Arg10/Pro10, Arg11, Phe12, Val14/Ile14) for antifungal activity. Therefore, protegrin 2 (RGGRLCYCRRRFCVCV) (**Figure 4.6**) might represent the minimal structure required for both antibacterial and antifungal activity [244, 245].

The 3D structures of some protegrins have been elucidated. All of them adopt a two-stranded antiparallel β-sheets in which the four conserved Cys residues connect both strands through two disulfide bonds between Cys6 and Cys15, and Cys8 and Cys13 [236, 240, 237, 244]. This can be seen in the NMR structure of PG-1, which is the most widely studied [245]. The structure of PG-1 is well defined from residues 5 to 16, being the N- and C-termini more flexible [236] (**Figure 4.7**), the central part of the antiparallel β-sheet is relatively non-polar

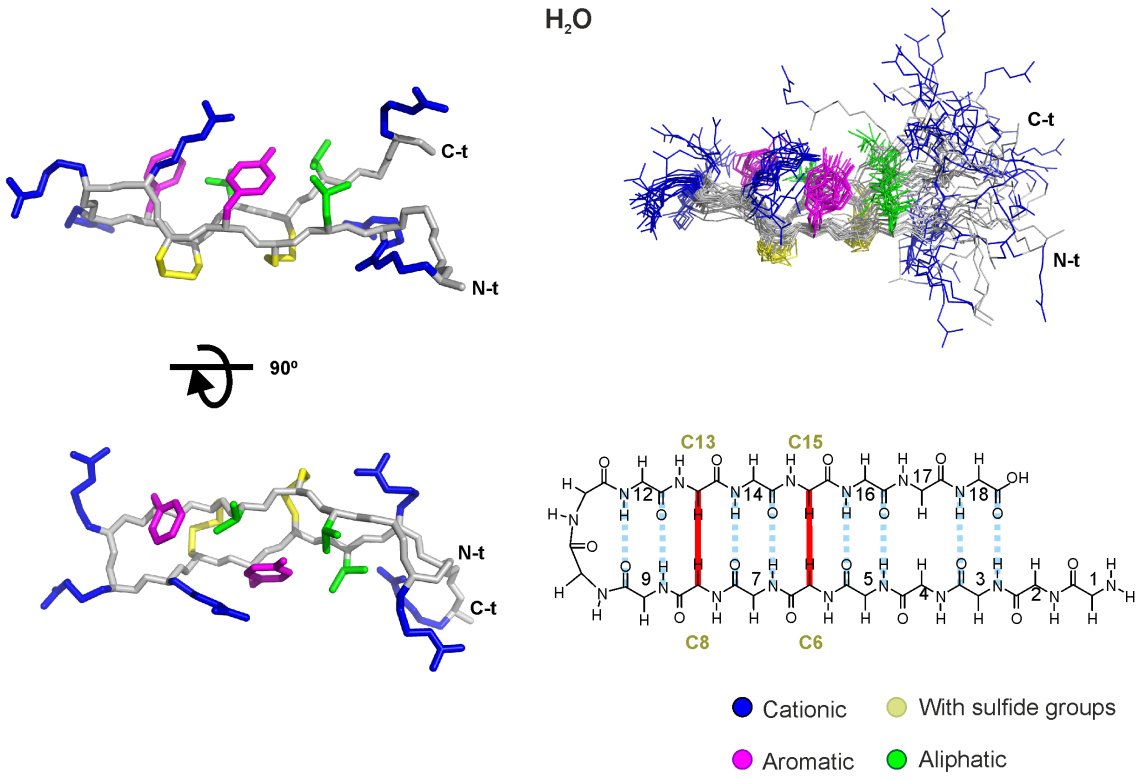


Figure 4.7. 90º rotated views in the horizontal axis of one conformer of PG-1 (PDB: 1PG1). Line representation of the 20 conformers with the lowest target values calculated for PG1 [236] in H_2O and simplified secondary structure representation with the disulfide bonds shown in red and hydrogen bonds in blue dashed lines.

(Cys and Val) at both sides, whereas the positively charged residues are located in the central β-turn region and at the C- and N- termini (**Figure 4.7**). Thanks to this disposition of positively charged residues, protegrins can interact with the surfaces of lipid bilayers through electrostatic interactions.

4.1.8 Objectives

Having in mind the importance of the protegrin family as possible antimicrobial agents, we propose to get new insights about the structure and function relationships of these AMPs by performing a structural study of some PG-1 variants. In particular, we intend to examine the importance of the disulfide bonds for structure stability, membrane interaction and oligomer state. Hence, the analyzed variants are a linear peptide with no Cys residues (PG-T4), and two peptides with a single disulfide bond: one lacking Cys6 and Cys15 (PG-C8C13), and another lacking Cys8 and Cys13 (PG-C6C15; **Table 4.2**). These last two peptides will allow us to analyze the relative importance of the disulfide bonds. We replaced the Cys residues by Thr, as Thr have a high β-sheet propensity [246–248]).

It should be noticed that Cys residues in PG-1 are not only key aminoacids for structure, but also necessary for biological activity. Previous studies with different types of PG-1 variants concluded that the disulfide bonds are essential for pore formation and its broad-spectrum activity [249–251]. However, the process is not yet well understood. In this aspect, we should mention that a full-length PG-1 variant with Cys substituted with Ala adopts a random coil conformation and shows a strongly reduced activity with respect to PG-1 wt [252].

The designed PG-1 variants in this study were characterized by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopies.

Hence, the specific objectives proposed are:

- To design variant peptides derived from PG-1 in which Cys residues are replaced by good β-sheet forming aminoacids (Thr) in different combinations.
- To fully characterize these designed peptides in different media by CD and NMR spectroscopies.
- To determine which disulfide bond is essential for the maintenance of the structural properties and their relation with those adopted by the native PG-1 sequence.
- As it is known that the interaction with membranes is a key step in the functional properties of protegrins, we want to determine if any or the two disulfide bonds are essential for the interaction of PG-1 with membranes.
- To determine if there are structural changes associated with the presence or interaction with membrane mimetics.
- To provide (if possible) a model on how the interaction with membranes occurs.

4.2 Materials and methods

4.2.1 Chemicals

The deuterated compounds [D_{38}] DPC (dodecylphosphocholine) (98 %) and D_2O (99.9 %) were obtained from Cambridge Isotope Laboratories (USA). [D_{25}] SDS (sodium dodecylsulfate) (98 %) was from Sigma. The percentages of deuteration are indicated in parenthesis.

4.2.2 Peptide synthesis

PG-T4 peptide was synthesized by Caslo Aps (Lyngbym, Denmark), peptide sequences are listed in **Table 4.2**. Fmoc-solid phase synthesis protocols were followed and purification was performed by reverse-phase HPLC using a C18 column until reaching 95 % or more purity. PG-C6C15 and PG-C8C13 peptides were synthesized by Lipotec S.A. (Barcelona, Spain). Fmoc-solid phase synthesis protocols were followed and purification was performed by reverse-phase HPLC using a C8 column until reaching 95 % or more purity. The identity of the peptides were confirmed by MALDI-TOF MS.

PG-T4: RP-HPLC: tR = 10.89 min, purity 96.95 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 17-37 % B buffer gradient in 20 min). HRMS: Theoretical molecular mass: 2152.51 (m/z); $[M+H]^+$ found: 2153.24 (m/z).

PG-C6C15: RP-HPLC: tR = 6.93 min, purity 95.90 % (buffer A: 0.1 % TFA in H_2O; buffer B: 0.07 % TFA in CH_3CN; linear B buffer gradient in 20 min). HRMS: Theoretical molecular mass: 2154.15 (m/z); $[M+H]^+$ found: 2154.28 (m/z).

Table 4.2. Sequences of the peptides studied, Cys residues are shown in red, substituted residues are underlined.
* Wild type sequence is displayed for comparison

Peptide	Sequence	Theoretical pI
PG-1 wt*	RGGRLCYCRRRFCVCVGR	12.30
PG-T4	RGGRLTYTRRRFTVTVGR	12.30
PG-C6C15	RGGRLTYCRRRFCVTVGR	12.30
PG-C8C13	RGGRLCYTRRRFTVCVGR	12.30

PG-C8C13: RP-HPLC: tR = 15.76 min, purity 96.30 % (buffer A: 0.1 % TFA in H_2O; buffer B: 0.07 % TFA in CH_3CN; linear B buffer gradient in 20 min). HRMS: Theoretical molecular mass: 2154.15 (m/z); $[M+H]^+$ found: 2154.59 (m/z).

4.2.3 CD spectroscopy

CD spectra were recorded with a Jasco J-815 spectropolarimeter (Tokyo, Japan) equipped with a Peltier PTC-423S system. Aliquots of the NMR solutions with known concentrations confirmed by UV spectrophotometry were taken and diluted into the appropriate volume of buffer to yield the peptide samples at the concentration required for each spectroscopic characterization. Samples were centrifuged 5 min, 13000 rpm prior to CD measuring.

All measurements were carried out by triplicate at 5 ºC and the cuvette path lengths were 0.1 cm for far-UV region (250-195 nm). Each peptide was tested at a concentration of 20, 30, 45 and 60 µM, pH 3.5 in 100 % H_2O or 30 mM DPC micelles.

Isothermal spectra for these samples were acquired at a scan speed of 50 $nm \cdot min^{-1}$ with a response time of 4 s and averaged over four scans for each sample. The same parameters were used to record reference spectra on the buffer/cosolvent solutions. After subtracting appropriate reference spectra from the sample spectra, CD data were processed with the adaptive smoothing method of the Jasco Spectra Analysis software. CD data are given in molar ellipticity units ($[\theta]$, deg $cm^2 \cdot dmol^{-1}$, see **Equation 2.1**).

4.2.4 NMR sample preparation

NMR samples were prepared by dissolving the lyophilized peptide (1-2 mg) in 0.5 mL of solvent; i.e., H_2O/D_2O (9:1 ratio by volume), pure D_2O, 30 mM $[D_{38}]$ DPC in H_2O/D_2O (9:1 v/v), 30 mM $[D_{38}]$ DPC in D_2O and 80 mM $[D_{25}]$ SDS in H_2O/D_2O (9:1 v/v). The peptide concentrations were between 0.9 and 1.8 mM. The pH was adjusted to 3.5 by adding minimal amounts of NaOD or DCl, measured with a Hamilton glass micro-electrode.

All samples were placed in 5 mm NMR tubes, and contained DSS as the internal reference for ^1H chemical shifts. NMR and sample conditions are summarized in **Table 4.3**.

Table 4.3. NMR spectra conditions of the peptides studied.

Peptide	Solvent	pH	Temperature (ºC)
PG-T4	H_2O	3.5	5, 25
	D_2O	3.5	5
	30 mM DPC/H_2O	3.5	25, 35
	30 mM DPC/D_2O	3.5	25, 35
PG-C6C15	H_2O	3.5	5, 25
	D_2O	3.5	5, 25
	30 mM DPC/H_2O	3.5	35
	30 mM DPC/D_2O	3.5	15, 35, 45
PG-C8C13	H_2O	3.5	5, 25
	D_2O	3.5	5
	30 mM DPC/H_2O	3.5	25, 35
	30 mM DPC/D_2O	3.5	5, 15, 25, 35, 45
	80 mM SDS/H_2O	3.5	35, 50

4.2.5 NMR spectra acquisition

NMR spectra were recorded on a range of 5-50 ºC with Bruker Avance spectrometers, operating at 600 and 800 MHz (^{1}H), both equipped with triple resonance, z-field gradient cryoprobes.

Phase-sensitive two-dimensional total correlated spectroscopy TOCSY spectra were obtained by using 20 ms and 60 ms mixing times. The nuclear Overhauser enhancement spectroscopy NOESY mixing time was 150 ms. ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra were recorded at ^{13}C natural abundance. ^{13}C δ-values were indirectly referenced by using the IUPAC-IUB recommended ^{1}H/^{13}C chemical shift ratio (0.25144953) [98, 93]. Water signal was suppressed by either presaturation or by using a 3-9-19 pulse sequence. The 2D data matrices were multiplied by a square-sine-bell window function with the corresponding shift optimized for every spectrum and zero-filled prior to Fourier transformation. Baseline correction was applied in both dimensions. Data were processed with the standard TOPSPIN program (Bruker Biospin, Karlsruhe, Germany).

4.2.6 NMR spectra assignment

The NMR spectral assignment of each peptide in aqueous solution at 5 ºC and in DPC micelles at 35 ºC was performed by following the well-established sequential-specific methodology based on homonuclear spectra [94] with the help of the SPARKY software [95]. The ^{13}C resonances were identified on the basis of the correlations between the hydrogen and the bound carbon atom present in the ^{1}H-^{13}C-HSQC spectra. Spectra acquired at different temperatures were analyzed in order to confirm ambiguous assignments. All assigned chemical shifts are listed in the **Appendix C**.

Chemical shift deviations calculated for the $^{13}C\alpha$ and $^{1}H\alpha$ applying the formula for either the carbon or proton [89]:

$$\delta\Delta\, H\alpha = \delta H\alpha_{observed} - \delta H\alpha_{random\ coil}$$
$$\delta\Delta\, C\alpha = \delta C\alpha_{observed} - \delta C\alpha_{random\ coil} \qquad \textbf{(Equation 4.1)}$$

4.2.7 Structure calculation

Structure calculations were done with CYANA 2.1 program [96]. Small peptides do not adopt a unique, highly stable structure. Instead, they adopt an ensemble of preferred, similar and modestly stable conformers, which are in equilibrium with low-populated conformers. To characterize the preferred conformers, medium and long range NOEs were selected. NOE integrated cross-peaks were translated into distance restraints, and the Φ and Ψ dihedral angle restraints were obtained using TALOS-N [97]. Typically, 100 structures were calculated using a standard protocol. The lists of distance constraints were checked with the corresponding NOESY spectra, ambiguous constraints were relaxed or removed in order to generate a final list used as input for a standard simulated annealing CYANA 2.1 calculation. The 20 conformers with the lowest target function values were selected. The structural ensembles were visualized and examined using MOLMOL [98] and PyMOL [99].

4.2.8 Dimer modeling and NMR-driven docking

To model dimer structures, two different approximations were carried out. First, sets of simulated annealing calculations were performed with CYANA using intermolecular NOEs detected and assigned during the assignment processes (**Table 4.5**, see Results **Section 4.3.4**).

Standard protocols with 82 linker pseudo-residues between monomers were used. The size of the linker was selected based on two reasons. First, the number of pseudo-residues was long enough to allow each monomer to fold and interact individually. Second, with this number, sequence numbering of the second monomer is the same as the first one plus 100. This facilitates the construction of NOE lists and restrictions, and at the end simplifies the calculation process of the dimer. These linker pseudo-residues have zero mass and zero van der Waals radii, thus the linker can freely pass through the structure during simulated annealing.

Also, docking protocols with PG-C6C15 were performed using the Haddock program [36]. Residues with identified intermolecular NOEs were chosen as active in the docking interface. The structures with the best Haddock docking scores were selected as representative conformers.

4.3 Results

4.3.1 Peptide design

To analyze the relevance of disulfide bonds in the formation of the β-hairpin structure found for wild type protegrin (**Figure 4.7**), and their possible interactions with micelles, three variants have been designed (PG-T4, PG-C6C15 and PG-C8C13, **Table 4.2**). Given that Thr amino acids have a high tendency towards β formation [246–248], the linear version of the PG-1 have all its Cys residues substituted with Thr, PG-C6C15 have only two Cys residues changed to Thr (residues 8 and 13) and PG-C8C13 posses only Cys residues in positions 8 and 13, being residues 6 and 15 changed to Thr.

4.3.2 CD data of PG-1 variant peptides

As shown in **Figure 4.8**, differences found on the CD spectra of PG-T4, PG-C6C15 and PG-C8C13 in H_2O and DPC micelles demonstrates that the peptides modify their structure in the presence of DPC micelles. In H_2O, a minimum of negative ellipticity is located between 195 and 205 nm for the three peptides, characteristic of a disordered or random coil

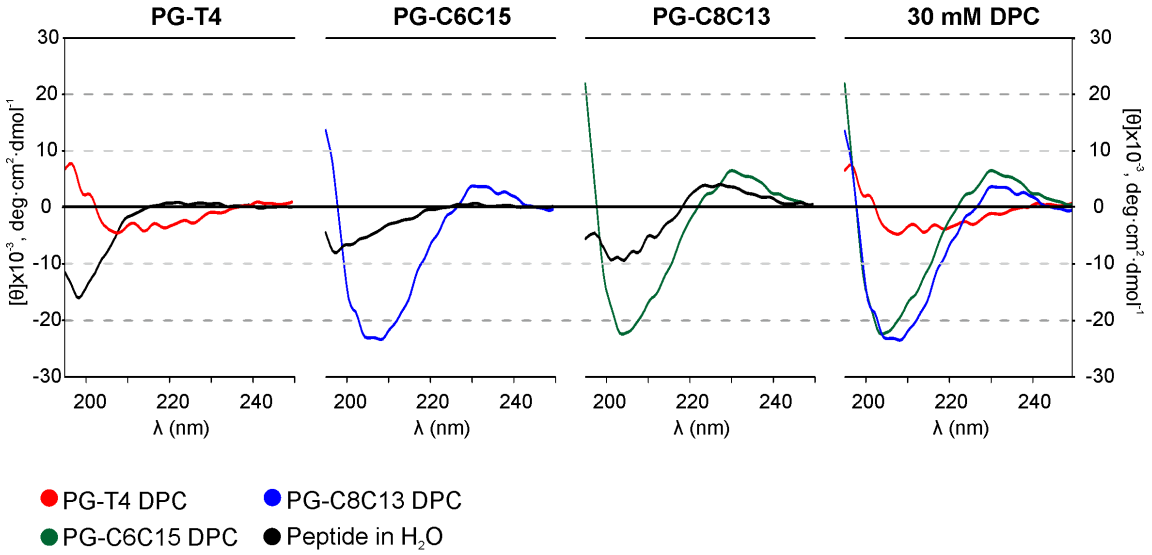


Figure 4.8. CD spectra of 30 µM peptides PG-T4, PG-C6C15 and PG-C8C13 in H_2O and 30 mM DPC micelles, 5 ºC, pH 3.5. The significant differences in the spectra indicate that DPC exerts a change in the global structure for all peptides and towards a β-sheet structure in the cases of PG-C6C16 and PG-C8C13. Even when PG-T4 changes after the addition of DPC micelles, a specific structure could not be defined through CD.

conformation. After the addition of DPC micelles, the spectra change completely for PG-C6C15 and PG-C8C13 peptides, in which the CD minimum get displaced towards 210 nm and its intensity increases considerably. This result is compatible with the stabilization of a β structure. However, the change in the CD spectra of the linear peptide PG-T4 is smaller, so that it is not possible to deduce the nature of the structure, if any, formed in DPC micelles.

As mentioned in the next sections, some of these peptides show a clear tendency to oligomerize. We tested by CD the spectral behavior of the three peptides at different concentrations (20, 30, 45 and 60 μM) in either H_2O or DPC micelles. However, CD spectra were not affected in the μM range (the one accessible for this spectroscopy), suggesting that larger concentrations are necessary for the association process (see below).

4.3.3 NMR data reveals peptide β-hairpin tendency in all variants

To better characterize the structural properties of the three peptides, we performed their NMR study in the two solvent conditions used for CD: H_2O and DPC micelles. In PG-C6C15 and PG-C8C13 peptides, the oxidized state of the Cys side chain groups is explicitly revealed by the chemical shift of the Cβ resonances (about 40 ppm; **Section C.1**, **Section C.2** and **Section C.3**). In addition, the two Hβ protons of both Cys residues are discernable when oxidized, in contrast with the almost identical values expected when Cys are in their reduced form [101]. After having assigned the NMR spectra, as described in Method's **Section 4.2**, we proceeded to analyze the profiles of chemical shift deviations for the Hα protons (**Figure 4.9**). In both solvents, H_2O and DPC, the conformational shifts of PG-C6C15 and PG-C8C13 are similar, and show the sign and distribution which are characteristic of a β-hairpin structure (**Figure 4.9**), but the magnitudes is different. Specifically, the segments spanning G2-R9 and F12-G17 present negative Δδ Hα values (typical of β-strands) and deviations of opposite sign in the connecting turn between residues R10 and R11. In the PG-C6C13 and PG-C8C15 peptides, values are larger (with respect to the random coil values) in DPC than in H_2O, meaning that the presence of DPC micelles is favoring the β-hairpin conformation. This can be observed in **Figure 4.10**, where the chemical shift values for the Hα in DPC were subtracted by those measured in H_2O. In these plots, the regions corresponding to the β strands gave positive values (**Figure 4.10**). Interestingly, the differences corresponding with the turn are low or negligible. This indicates that the population of the turn, already present in water, does not change substantially in micelles, but it increases considerably in the β strand regions. However, Δδ values in DPC are larger than those expected for monomeric β-hairpins [253, 100] (**Figure 4.9**, **Appendix C.2** and **Appendix C.3**), suggesting that these peptides can be involved in supra molecular interactions.

In the case of the linear PG-T4 peptide, the profiles of conformational shifts of the Hα protons are very small in magnitude (mostly within the random coil range: $|\Delta\delta| < 0.02$ ppm), but the plot of DPC minus H_2O values shows the same pattern that PG-C6C13 and PG-C8C15 peptides, though the magnitude is very small. This indicates that the linear PG-T4 peptide presents a low tendency to adopt an ordered conformation, but still a β-hairpin structure gets stabilized in DPC micelles, as seen in **Figure 4.10**.

The complete assignment of the ^{13}C chemical shifts was not achieved in the presence of DPC because signals are broader in these conditions, particularly in the less sensitive ^{13}C-HSQC spectra, which was acquired at natural abundance of ^{13}C. Nevertheless, the profiles of conformational shifts for the assigned 13Cα carbons are in concordance with the conclusions derived from those of Hα protons. The broadening effect was also important for the NOE analysis of the PG-C8C13 peptide. Therefore, in order to solve this problem, other type

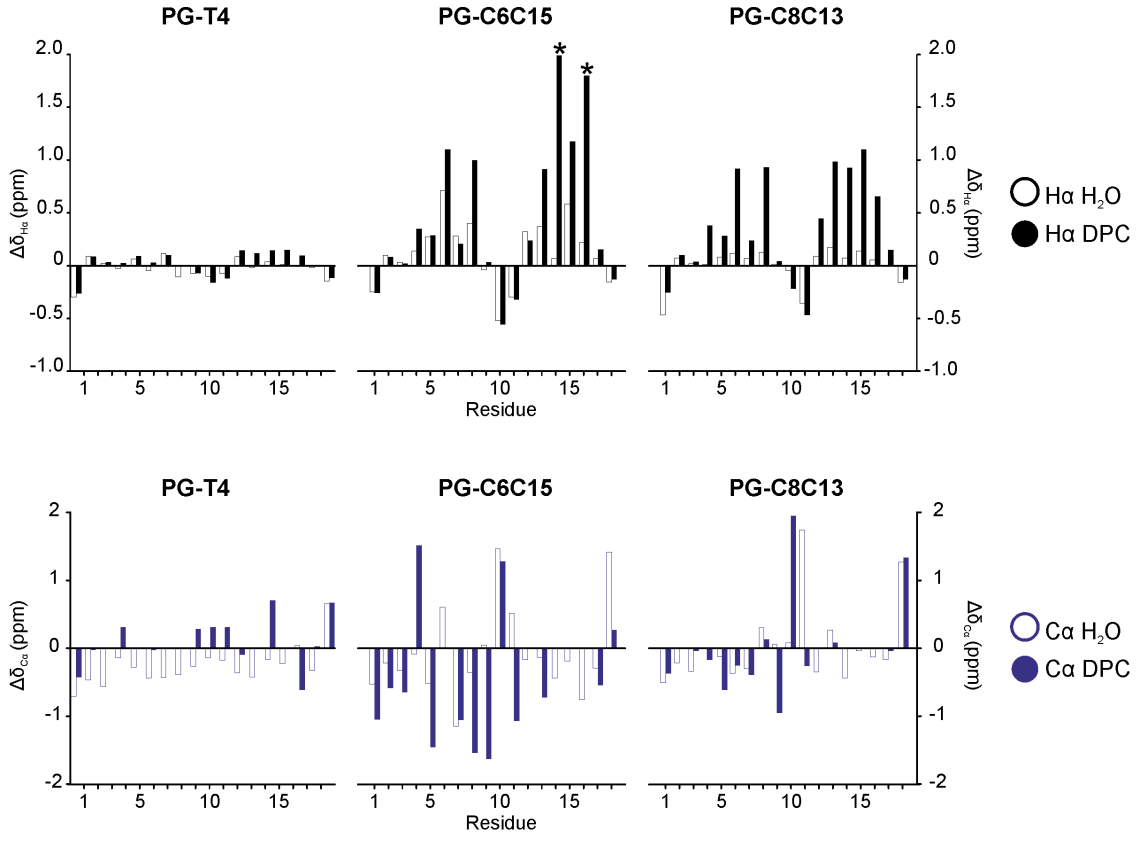


Figure 4.9. Barplot representation of the chemical shift deviations of α protons (black bars) and carbon resonances (blue bars) relative to random coil values [101] for peptides PG-T4, PG-C6C15 and PG-C8C13 in H_2O (5 ºC) (open bars) and 30 mM DPC micelles (35 ºC) (filled bars), pH 3.5. The profile displays a high level of β-sheet structure stabilization in DPC micelles for PG-C6C15 and PG-C8C13, the difference is more striking for the peptide PG-C6C15.
* Ambiguous assignment, values can be exchanged.

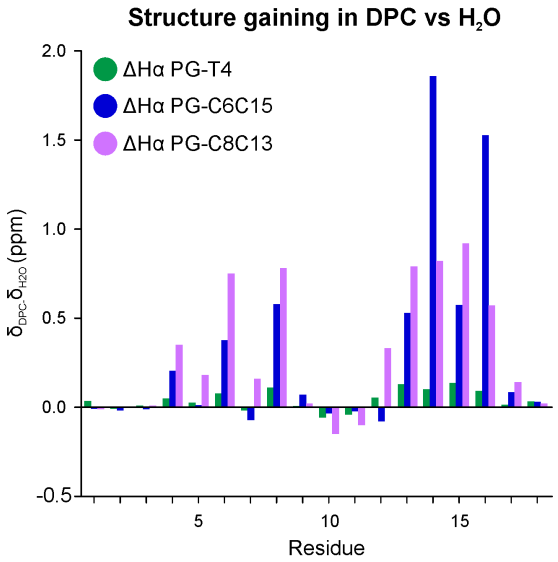


Figure 4.10. Barplot diagram of the Hα chemical shifts of the peptides. Hα chemical shift values in DPC are substracted with the Hα values obtained in H_2O.

of micelles was tested. However, PG-C8C13 in SDS micelles also presents a large signal broadening, making it impossible to fully assign all the spin systems (**Figure 4.11**).

As seen on **Table 4.4** (see further pages), long-range NOE connectivities in accordance with the formation of a β-hairpin structure were found in PG-C6C13 and PG-C8C15 peptides in H_2O and in DPC micelles. The number of these long-range NOEs is higher in DPC,

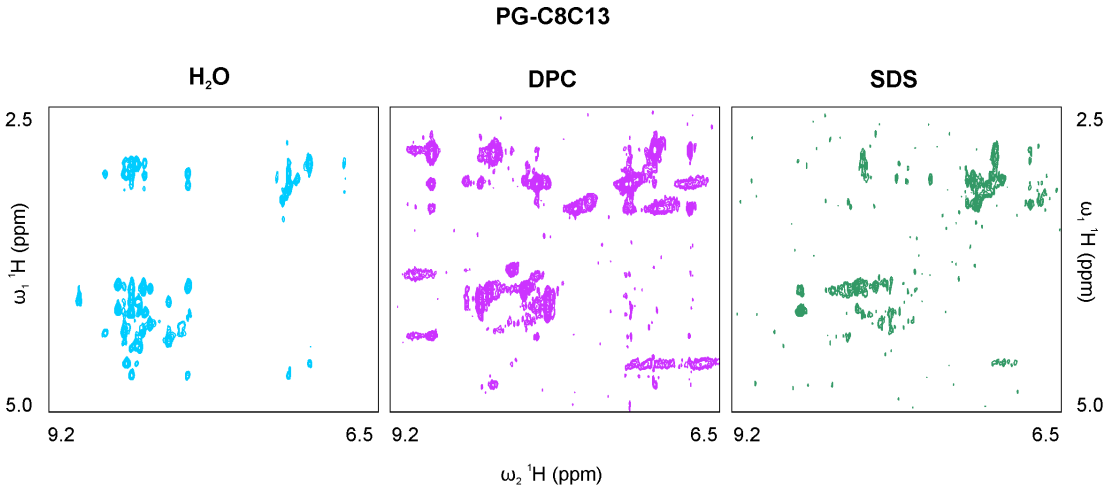


Figure 4.11. Region of the TOCSY spectra for the peptide PG-C8C13 in H_2O 5 ºC, 30 mM DPC 35 ºC and 80 mM SDS, 35 ºC, pH 3.5. The presence of DPC and SDS micelles produce signal broadenings that might avoid a complete assignment of the spin systems.

which corroborates the higher percentage of ordered structures. In the case of the linear peptide PG-T4, no long-range NOE was found in H_2O, but there are medium-range NOEs, particularly at and around the turn region. This suggests that the linear peptide have some local ordered structure at the turn region, as seen in the calculated structure (see next section). In the presence of DPC, some long-range NOEs indicative of β-hairpin formation were found (**Table 4.4**). These results point out that the micelle media stabilizes the β-hairpin formation even in the absence of the native disulfide bonds.

4.3.4 Structure calculation

To visualize the features of the adopted β-hairpins, we performed structure calculations. Distance and angular constraints were used in the calculations, and additional constraints corresponding to disulfide links were added pertinently. All structures formed a β-hairpin in both solvents, except PG-T4 in H_2O. As no long-range NOEs were found in this case, the final ensembles consist of two β strands apart from each other linked by a central turn. For PG-C8C13, as the disulfide bond is close to the central turn, N- and C-termini show some fraying. Meanwhile, the disulfide bond between cysteines 6 and 15 of PG-C6C15 stabilizes the formation of the β-hairpin in H_2O by bringing the strands closer, with little or no difference compared with the structure obtained in DPC. A summary of these results is displayed in **Table 4.4**.

The relatively high RMSD values for the backbone atoms for the whole sequence (2.4, 1.7 and 2.4 Å for PG-T4, PG-C6C15 and PG-C8C13, respectively) and the low values between the regions spanning positions 6 to 15 (0.3, 0.4 and 0.3 Å for PG-T4, PG-C6C15 and PG-C8C13, respectively) in H_2O, indicate that the regions closer to the N- and C-termini are more dynamic (**Table 4.4**).

As seen in **Figure 4.12** and **Figure 4.13**, in all variants, the hydrophobic residues (green and magenta) are clustered in the center of the hairpins, while the hydrophilic ones (in blue) get disposed in the β-turn and the N- and C-termini. The hydrophobic cluster gets more ordered in DPC micelles, and the side chains get faced.

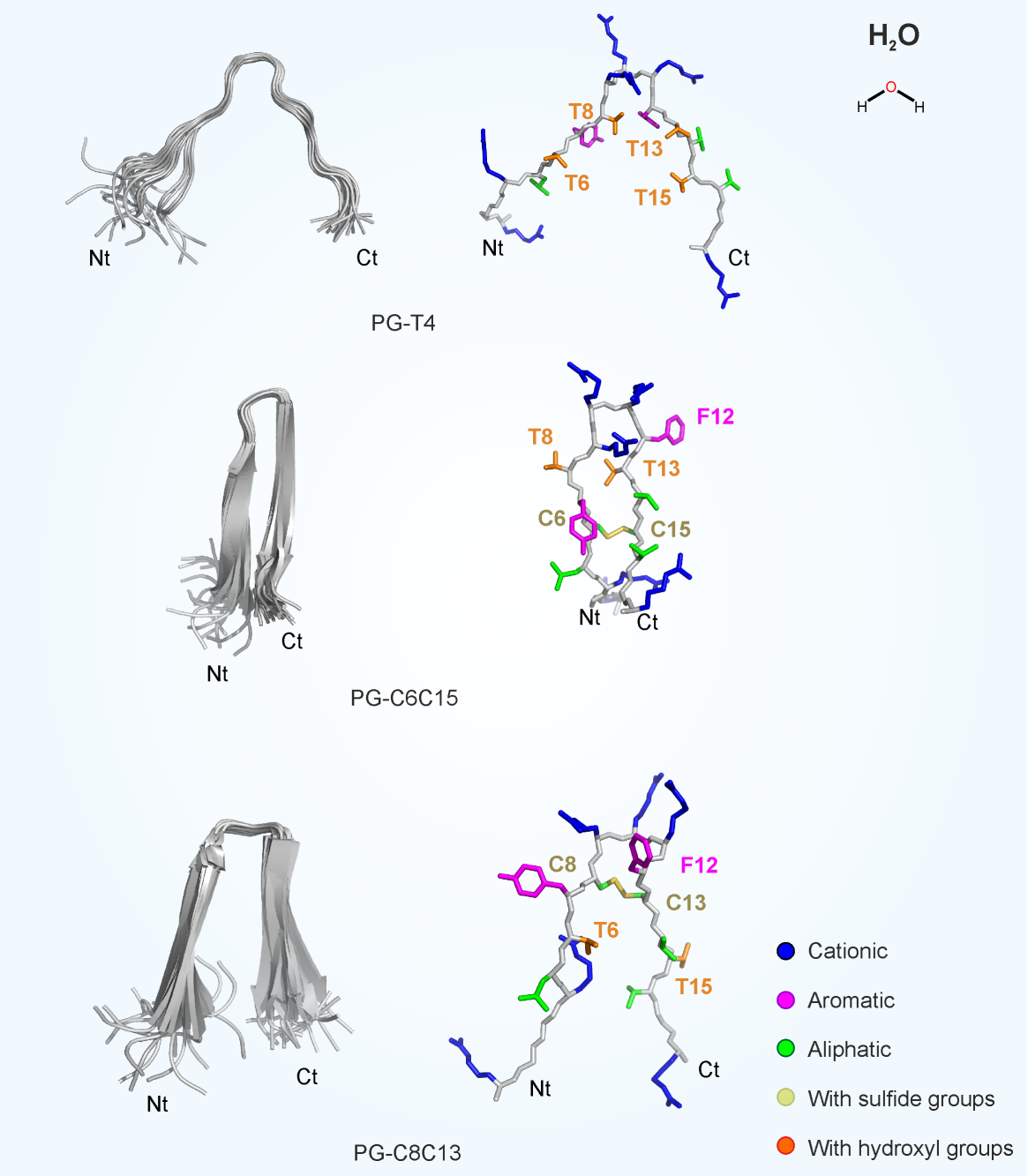


Figure 4.12. Ensemble of the 20 structures with the lowest target function values of PG-T4, PG-C6C15 and PG-C8C13 in H_2O, 5 °C, pH 3.5. The superposition of the backbone of the 20 lowest target function structures in each family is represented as ribbon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Cationic: blue, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

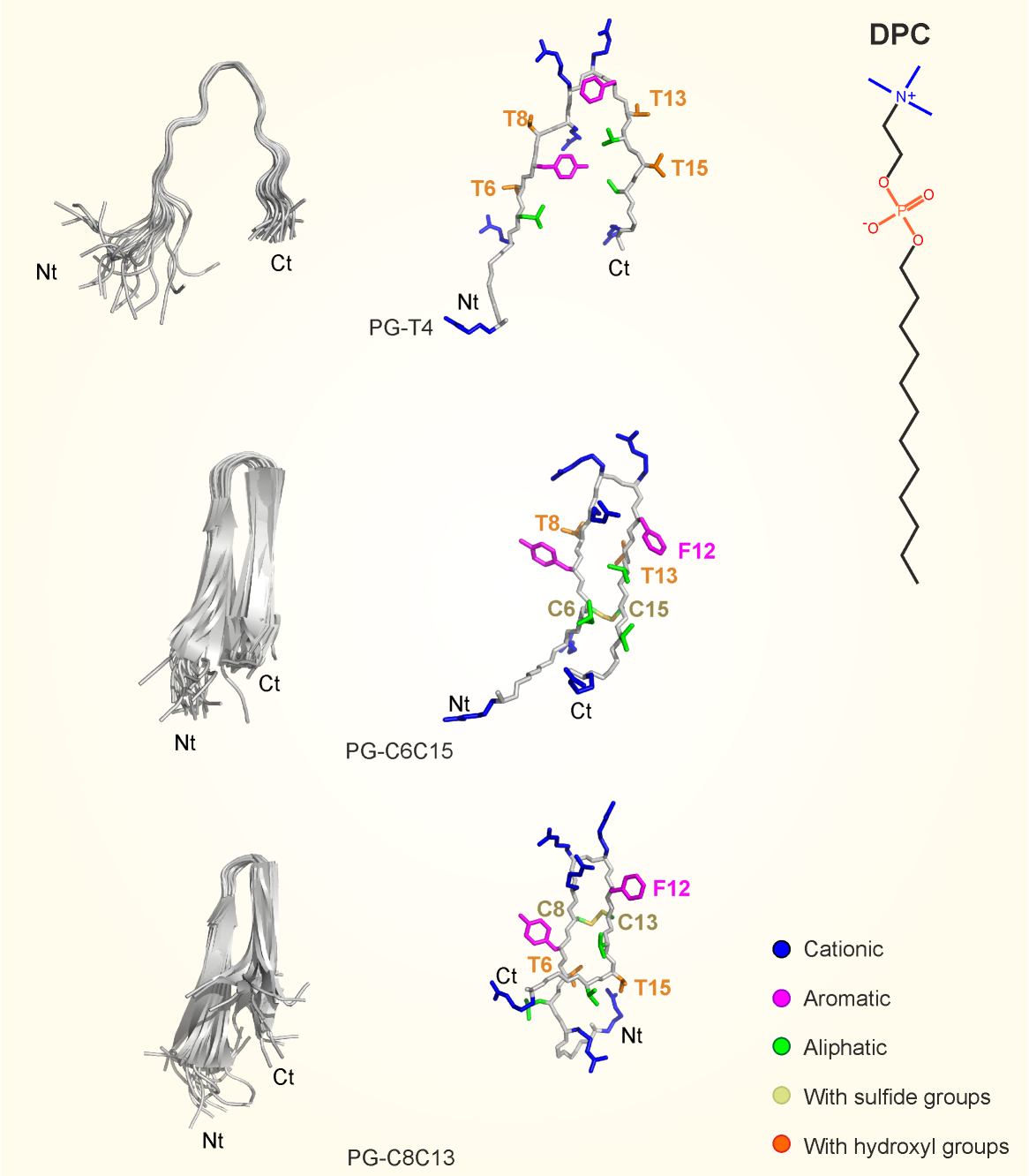


Figure 4.13. Ensemble of the 20 structures with the lowest target function values of PG-T4, PG-C6C15 and PG-C8C13 in 30 mM DPC, 35 °C, pH 3.5. The superposition of the backbone of the 20 lowest target function structures in each family is represented as ribbon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Cationic: blue, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

Table 4.4. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for PG-T4, PG-C6C15 and PG-C8C15 in H_2O 5 ºC and 30 mM DPC, 35 ºC, pH 3.5

	H_2O			30 mM DPC		
	PG-T4	**PG-C6C15**	**PG-C8C13**	**PG-T4**	**PG-C6C15**	**PG-C8C13**
Upper Limit distance restraints (from NOEs)						
Total	83	42	64	85	72	109
Short-range $\|i-j\| \leq 1$	53	28	37	49	35	54
Medium-range $1 < \|i-j\| < 5$	30	7	23	30	28	40
Long-range $\|i-j\| \geq 5$	0	7	4	6	9	15
Disulfide bonds		C6-C15	C8-C13		C6-C15	C8-C13
H bonds					Y7-V14	
φ/ ψ Dihedral angle constraints (from chemical shifts)	11/11	12/11	11/11	13/7	4/2	8/7
Average CYANA target function value	0.04 ± 0.1	0.02 ± 0.1	0.01 ± 0.1	0.04 ± 0.1	0.16 ± 0.1	0.04 ± 0.1
Averaged maximum violation per structure						
Distance (Å)	0.01 ± 0.1	0.08 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1
Dihedral angle (º)	0.48 ± 0.2	0.87 ± 0.1	0.49 ± 0.2	0.48 ± 0.2	1.38 ± 1.4	0.41 ± 0.4
Pairwise RMSD (Å)						
Backbone atoms	2.4 ± 0.7	1.7 ± 0.6	2.4 ± 0.7	2.4 ± 0.7	1.5 ± 0.4	3.0 ± 0.8
All heavy atoms	4.0 ± 0.8	3.0 ± 0.6	3.9 ± 0.8	3.6 ± 0.9	3.2 ± 0.6	4.3 ± 0.9
Region 6-15						
Backbone atoms	0.3 ± 0.2	0.4 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.6 ± 0.3	0.9 ± 0.4
All heavy atoms	1.8 ± 0.5	1.6 ± 0.3	2.0 ± 0.4	1.3 ± 0.2	2.2 ± 0.4	2.1 ± 0.4
Ramachandran plot (%)						
Favorable	82.8	83.8	85.9	80.6	46.6	72.5
Allowed	11.9	13.4	5.0	16.2	46.2	22.5
Outlier	5.3	2.8	9.1	3.1	7.2	5.0

4.3.5 Peptide oligomerization

In the presence of DPC, PG-C6C15 and PG-C8C13 show some additional NOE signals incompatible with a monomeric β-hairpin that lead to distance and angular violations during the structure calculation of the monomeric form (**Table 4.5**). These include strong signals between the side chains of Val16 and Phe12 and Leu5 and Cys8.

The additional NOEs, summarized in **Table 4.5**, were used as input for the calculation of dimeric structures in the micellar media and no distance violations were found in this case. We should mention that these data sets are probably not complete due to the broadening effects in DPC that prevent the observation of weak intermolecular NOEs. This might be the motive for the failure of the simulated annealing calculations in the torsion angle space done with CYANA to calculate the dimer structure (see Method's **Section 4.2**). For this reason, a docked structure based on the experimental NMR data is the convenient approach to get a molecular model.

PG-C6C15 spectra showed a better quality than the PG-C8C13 ones in DPC, and much more data was obtained that facilitate the docking strategy. As seen in **Figure 4.14** and **Figure 4.15**, the docking procedure generated two types of dimer models arranged in a head-tail disposition. In both cases, they are composed by two antiparallel β-hairpins forming an arrangement of four parallel/antiparallel β-strands. In model A, residues spanning positions

Table 4.5. Intermolecular NOEs found in peptides in 30 mM DPC 35 ºC, pH 3.5.
* Q letters refer to centrally located pseudoatom as a reference point for the distance restraint.

Residue	Proton		Residue	Proton
PG-C6C15				
Leu 5	HN	→	Thr 8	Qγ
Leu 5	Qδ	→	Thr 8	Hα
Phe 12	Qβ	→	Val 16	QQγ
Phe 12	Qδ	→	Val 16	QQγ
Phe 12	Qε	→	Val 16	QQγ
Phe 12	Qβ	→	Gly 17	HN
Phe 12	HN	→	Arg 18	Hα
Val 14	Hα	→	Val 16	Hα
PG-C8C13				
Phe 12	Qβ	→	Val 16	QQγ
Phe 12	Qδ	→	Val 16	QQγ
Phe 12	Qε	→	Val 16	QQγ

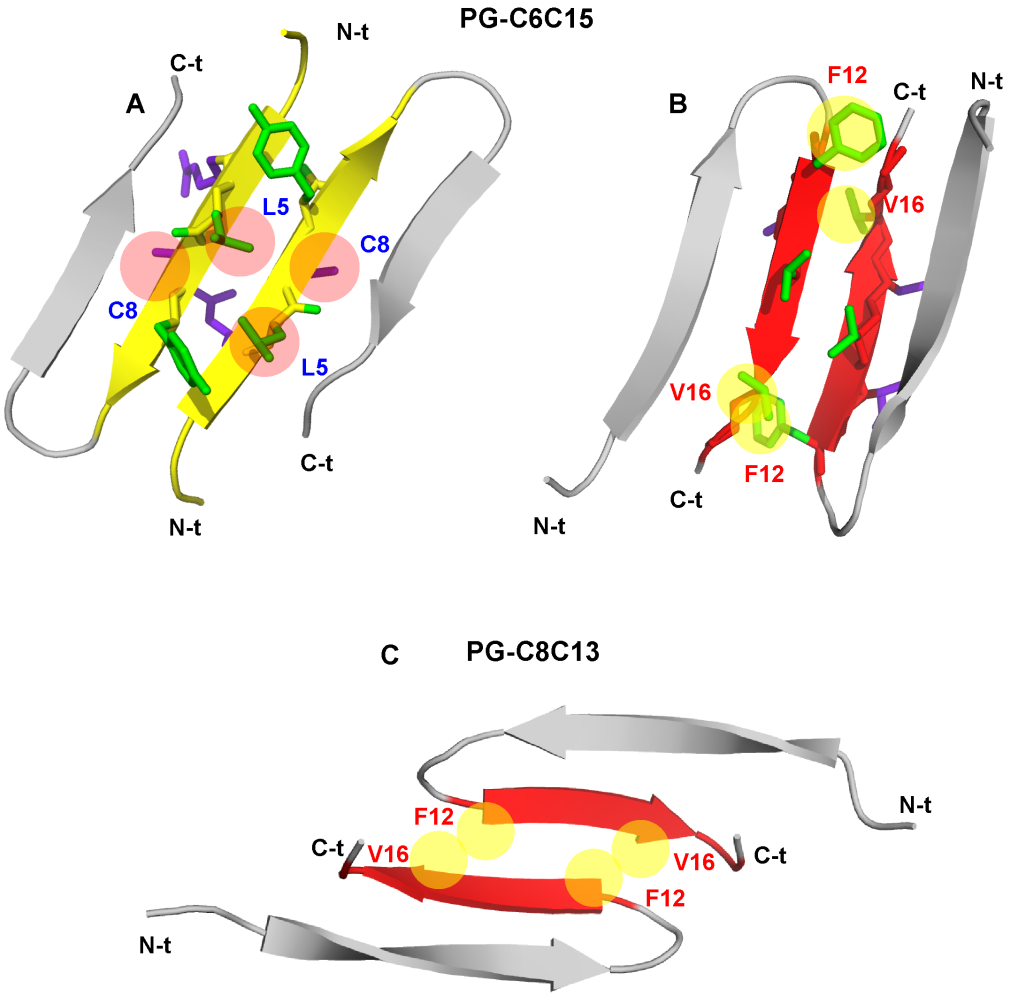


Figure 4.14. Structural 3D computed models of PG-C6C15 (A and B) and PG-C8C13 (C) obtained by docking. Residues in the same face of the sheet are colored either in green or purple. Residues used as input for the docking protocols are highlighted in yellow and red circles. Interacting N-terminus proximal regions are shown in yellow and C-terminus proximal in red.

Gly3 and Arg9 interact with their cognates of the other monomer; while in model B, residues between Phe12 and Gly17 interact with their homolog residues. NOE data reveals that both models take place when DPC micelles are present.

Due to signal broadening caused by the presence of DPC micelles, the spectra of PG-C8C13 peptide did not provide enough intermolecular NOE signals to calculate both types of dimers. The observed intermolecular NOEs corresponded to the hydrophobic residues located in the same face of the β-sheet. Therefore, a single docking structure could still be addressed with the interactions found between Phe12 and Val16 (**Figure 4.14 C**). This

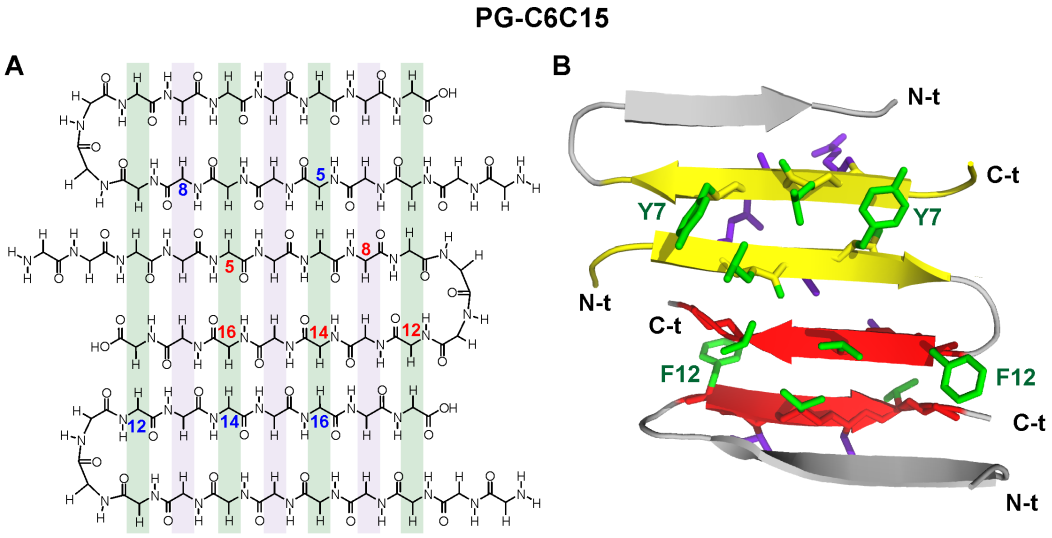


Figure 4.15. Schematic representation (A) and combined 3D model (B) of possible oligomerization for the peptide PG-C6C15 in 30 mM DPC micelles at 35 ºC. Backbone of the interacting side chains are indicated in red and yellow and residues facing in the same side are grayed out in green or purple.

information suggests that dimerization also arises in the presence of DPC micelles in a similar way occurred in PG-C6C15.

On the basis of the two possible dimers found in DPC, a 3D model of a possible oligomer for PG-C6C15 was constructed. In **Figure 4.15**, the combination of both Haddock docking results is joined in one model in which two different faces can be clearly differentiated. In one face, the side chains of hydrophobic residues such as Leu5, Phe12, Val14 and Val16 and the aromatic ring of Tyr7 are interacting. In the other face, the side chains of the cationic and polar residues (lysines and threonines) are in the same plane. Simultaneously, certain curvature can be appreciated in the dimers with the hydrophilic residues pointing outwards the structure. This slight curvature, and the orientation of these residues are compatible with the formation of the core of a possible β-barrel oligomer.

4.4 Discussion

AMPs are becoming promising molecules against resistant pathogens [202], but there are still several challenges to overcome, such as their low specificity, manufacturing cost and potential toxicity against animal cells [189]. As their mechanisms of action still remain unclear and to overcome the current problems, a full understanding of the structure-function relationships of these peptides is essential to design new functional analogs. In this regard, our study uses NMR and CD spectroscopies to analyze the structural relevance of the disulfide bonds of PG-1 through the substitution of Cys residues with Thr.

Is has been proposed that the β-hairpin formed by protegrins is essential for their propensity to interact with the microbial membranes and cause cell disruption [242, 254]. The β-turn in the center of the hairpin consists of a cationic Arg-rich region. Mutants on these residues indicate that the presence of Arg might be essential towards pore formation, suggesting that the cationic Arg residues drag out the polar heads of the lipids while the peptide gets inserted through the bilayer [254].

4.4.1 Structures of PG-1 variants, the role of Cys residues

The total depletion of Cys residues in PG-T4 does not allow it to adopt a compact β-sheet structure in water, however, the β-turn is still present (**Figure 4.12**) and the strands get approximated and stabilized in the presence of DPC micelles. In contrast, previous studies with peptides in which all Cys residues were substituted by Ala showed that they do not fold and they adopt a random coil conformation, thus stating that the Cys residues are essential for the β-hairpin formation [249, 252]. Here, our results indicate that Cys are actually not essential for hairpin formation (**Figure 4.13**) and that their substitution with Thr (instead of Ala) is enough to populate the hairpin structure.

Also, the results obtained here with PG-T4 indicate that the interaction with the DPC micelles is present in the absence of Cys and it is responsible for the observed conformational change. Thus, the micelles exert a stabilizing effect in the peptide structure of PG-T4 even in the total absence of disulfide bonds. However, even when the interaction with micelles still occurs, further evidences of intermolecular interactions between the monomers were not found.

In the case of the peptides C6C15 and C8C13, the Cys residues are not located at H-bonding sites (**Figure 4.7**). According to previous studies [255], the presence of a disulfide

bond in these positions stabilizes the β-hairpin structures mainly by decreasing the loss of entropy while folding. Simultaneously, and as seen in **Figure 4.9**, conformational chemical shifts are larger in PG-C6C15 than in PG-C8C13, showing that the population of ordered structures (both in H_2O and DPC) also depends on the position of the disulfide bridge. This is in agreement with previous studies showing that the disulfide stabilization effect depends on the distance with the turn region [255], being stronger in peptides with the disulfide bond located farther from the turn. These entropic effects, combined with the packing of the hydrophobic side chains may contribute for the peptide C6C15 to present a more stabilized β-hairpin structure than its counterpart. These effects are more important in DPC than in water solutions and can be clearly appreciated from the comparison of their 3D structures (**Figure 4.12** and **Figure 4.13**). For PG-C8C13, as the disulfide bond is closer to the turn, the structure is less stable, C- and N-termini of the β-strands are more prone to be separated in water. However in DPC micelles, they get closer and the β-hairpin becomes more stabilized as has been noted for C6C15.

4.4.2 Insights into the mechanism of action

Probably, the most interesting result we have found is that the hairpins of PG-C8C13 and PG-C6C15 associate in the presence of DPC to form stable supramolecular entities. These aggregates are observed at the concentration levels used in NMR, while in the μM range of the CD experiments could not detect any sign of oligomerization, suggesting that it occurs at higher peptide concentrations not reached by this spectroscopy. The presence of these specific dimers was not unexpected, as oligomerization processes of other related peptides have been described in the literature [242–245]. These processes are key for the interaction with membranes and determine their mechanism of action.

Looking the antecedents of some other AMP results, different models of interaction with the membranes can be established (**Figure 4.16**). First, it has been previously proposed that the barrel-stave model might not be plausible for this kind of AMPs [202], and molecular dynamics simulations do not support this formation [256] due to the electrostatic repulsion present in the model. Therefore, two other models such as the toroidal pore or the carpet-like model are in principle feasible in this family of peptides.

In the toroidal pore model, we can propose that the different properties of both faces of the model (explained in the Results' **Section 4.3.5**) can play a role in the interaction with membranes. Thus, the hydrophilic residues located in one side of the sheets (Lys and Thr) of each peptide, might keep the contacts with the polar heads of the lipids. In the other side of

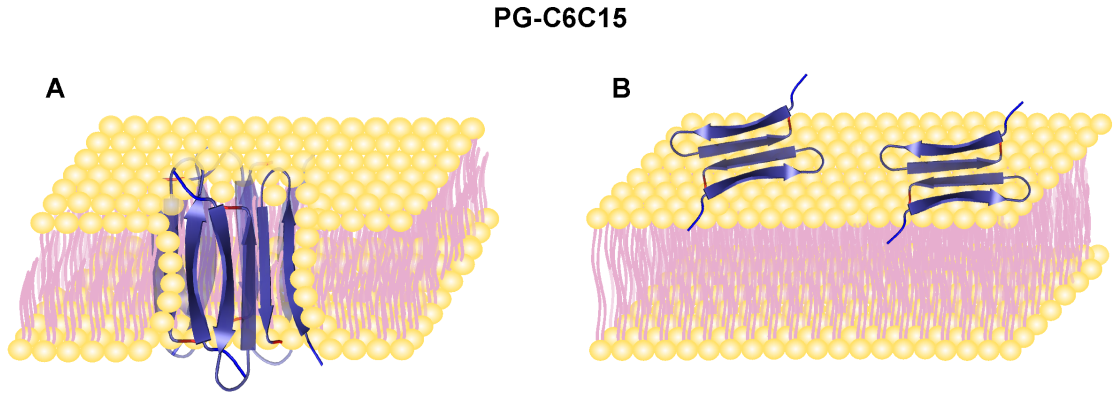


Figure 4.16. Models of possible interactions between the PG-C6C15 peptide and biological membranes based on NMR data and 3D docking simulations. Peptidic C-termini are indicated in red. A: Toroidal pore formation: PG-C6C15 aggregates and introduces into the lipid bilayer, the hydrophilic face of the peptides get in contact with the polar heads of the lipids. The hydrophobic residues in the other face of the peptides get in contact and form a pore in the core. B: Aggregation occurs upon the contact with the membranes and pores are formed as described previously in the carpet-like model.

the sheet, the hydrophobic residues might maintain the structural packing, forming a pore in the core of a possible multi-molecular association. Also, the slight curvature found in our calculated dimer of dimers suggests a predefined organization of the aggregate. This pre-arrangement can be energetically favorable and drive to a final β-barrel pore. The model presented here for the peptides studied in the present study, is in agreement with previous solid-state NMR (ssNMR) experiments [242], as well, with dynamic simulations done with PG-1 in POPE/POPG bilayers that strongly support the formation of a toroidal pore [256].

In this regard, two structural studies related with the association of AMPs can be also considered. A previous crystallographic study of the human AMP defensin HNP-3 revealed that it associates as an hexamer stabilized by hydrogen bonds and hydrophobic contacts between the peptides [257]. On the other side, simulated lipid environments suggest that each pore of PG-1 is made by eight monomers [256]. Given that the PG-1 peptide structurally resembles to the human defensin HNP-3 and that the present results with PG-1 mutants are compatible with an oligomerization in DPC micelles, the model of action may consist of a barrel formed by certain number of dimers (**Figure 4.16 A**), but the exact quantity could not be determined on the basis of our results.

The carpet-like model previously suggested for PG-1 and other related AMPs [242, 258] can be also applied to our peptides. In this model, the aggregation should occur upon the contact with the membranes, as we have demonstrated for the PG-C8C13 and PG-C6C15 peptides in DPC. The hydrophobic residues located at the central part of the strands might

form a carpet with the hydrophilic regions contacting the polar heads of the lipids. Because of the flexibility of the C- and N-termini, these residues might be within the polar heads, allowing the hydrophobic ones to be anchored between the hydrophobic tails of the lipids (see **Figure 4.16 B**).

It is also important to mention that ssNMR experiments of PG-1 in POPE/POPG and POPE/cholesterol membranes did not contradict the present proposals [242], though, when cholesterol is present in the POPC bilayer, ssNMR experiments suggested that the peptide disposition resemble a carpet-like model, as the sterol rings reduce the membrane elasticity and avoids the possible insertion of the PG-1 peptides. Nevertheless, it has to be taken into account that cholesterol is not naturally present in the bacterial membrane [259]. However, it has been reported that cholesterol can be exogenously incorporated in certain strains [260] and also, there are bacterial hopanoids that act as cholesterol analogs [261]. This means that the carpet-like model proposition should be also taken into account for our system, but with certain caution and much less probability.

Finally, we must say that we do not have enough evidences to distinguish or take sides for one or other model proposition, but it seems clear that the association process previous or concomitant with the membrane interaction is present, that it is important for the first steps of membrane disruption and that it can be clearly detected by NMR in the presence of membrane mimetics as DPC.

4.5 Conclusions

1. The results of the present study support the hypothesis in which the Cys residues are highly important for the stabilization and rigidity of the β-sheet structure at least in H_2O, but are less crucial towards the formation of the hairpin in DPC. In this media, the presence of long-range NOEs in the PG-T4 variant indicates that the strands are close in space (less than 5 Å).

2. In accordance with the NMR results, CD assays reveal the overall changes suffered by the mutants, as they are either placed in water or micelle media. In H_2O, the lack of certain disulfide bonds gives more disordered structures; nonetheless, the absence of the disulfide bond between residues 8 and 13 (the bond closer to the turn) did not affect significantly to the overall backbone in both H_2O and DPC micelles (**Figure 4.17**).

3. CD spectroscopy revealed a mainly random conformation for the PG-T4 peptide, though, by NMR it has been proven that a certain degree of ordered conformation can be described. It confirms that through NMR, more detailed information regarding the structure is given.

4. The present NMR data reveals that while the position of the cysteine residues is quantitatively relevant for the structural stability, the lack of cysteine residues in PG-T4 does not avoid its possible interaction with DPC micelles.

5. In the presence of DPC micelles, the NMR data obtained are in accordance with the formation of a toroidal oligomer with their antiparallel monomers additionally stabilized by hydrogen bonds. This is in agreement with diverse models of PG-1 oligomerization previously reported.

6. Our results can also be compatible with a carpet-like model, however, and according to previous literature, to favor this formation over the toroidal pore model, sterols should be present in the bilayer and higher peptide concentrations are needed to elicit a lytic effect.

7. Finally, several approaches should be addressed to obtain more complete models such as other biophysical techniques employing peptides embedded in membranes or bicelles. As well, biological assays are needed to analyze their efficiency against the target of these promising therapeutic molecules.

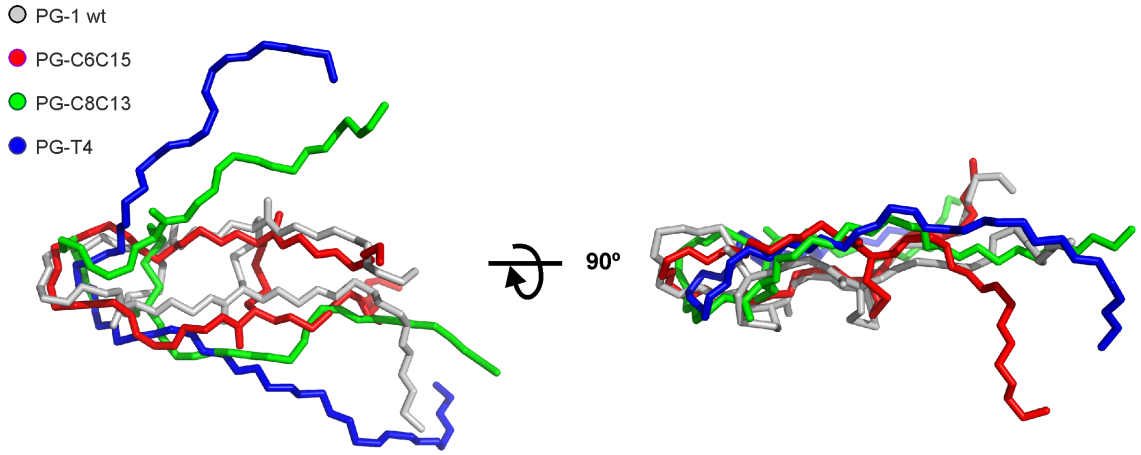


Figure 4.17. 90º rotated views in the horizontal axis of the backbone superposition of one conformer for each peptide: PG-1 wt (gray), PG-T4 (blue), PG-C6C15 (red) and PG-C8C13 (green) in H_2O. PG-C6C15 maintains a similar structure compared with the wild type PG-1. The depletion of all Cys residues in PG-T4 changes dramatically the stability of the β-sheet structure in H_2O. PG-C8C13 still conserved a similar turn, but the position of the only disulfide bond closer to the turn allows the peptide to adopt more disordered termini regions.

5

Structure and molecular interactions of the α-mating pheromone from *Fusarium oxysporum* in the presence of membrane mimetics and other cosolvents

5.1 Introduction and objectives

5.1.1 Morphological and physiological traits in ascomycetes

Ascomycetes is a fungi division that belongs to the clade dykarya, this means that their components are characterized by having a portion of their life cycle with paired nuclei. In the same way, ascomycetes are believed to belong to a monophyletic (one common ancestor) group consisting on three main different subphyla: Taphrinomycotina, Saccharomycotina and Pezizomycotina, in which the majority of the filamentous fungi (fungi organized as a network of filaments) belong to the last one. Ascomycetes are the largest and most diverse group of fungi; it includes unicellular yeasts such as *Saccharomyces* sp. (**Figure 5.1**), and complex filamentous fungi that can be found establishing symbiotic associations with algae and superior plants in the form of mychorrhizae. Moreover, they can be found as saprophytes and as animal and plant pathogens [262].

The term ascus in ascomycetes refers to the shape of the appendage that holds the spores, which is like a sac [263]. During their biological cell cycle, ascomycetes may produce from four to eight sexual spores that are held by the ascus (**Figure 5.1**). Subsequently and depending on the species, this ascus can be integrated into a fruiting body known as ascocarp, which can contain millions of asci and a network of hyphae [264]. Though, these fungi can also undergo a fast asexual life cycle and produce big amounts of asexual spores called conidia. This capacity is what makes this type of fungi highly invasive and destructive [265].

In order to colonize new habitats, filamentous fungal pathogens and symbionts experiment hyphae growth towards determined chemical signals coming from either other fungi or host plants [266–268]. In this regard, different chemical signals guide hyphae growth through the stimulation of diverse routes, such as lateral branching, the fusion with other hyphae or the elongation of the polarized tip [269]. These processes create a vegetative fungal colony known as mycelium, which is a complex and wide network made by several hyphae components [269].

Root penetration is one of the diverse infective pathways performed by filamentous fungi in the host plant and it is hypothesized that hyphae might sense some specific chemical signals from the host plant and take advantage of the natural openings within the junctions of epidermal cells [270]. However, the mechanisms underlying these chemotropic responses remain unclear in most filamentous fungi, as the traditional model organism employed to study these processes up to date is the unicellular fungus *Saccharomyces cerevisiae* [270].

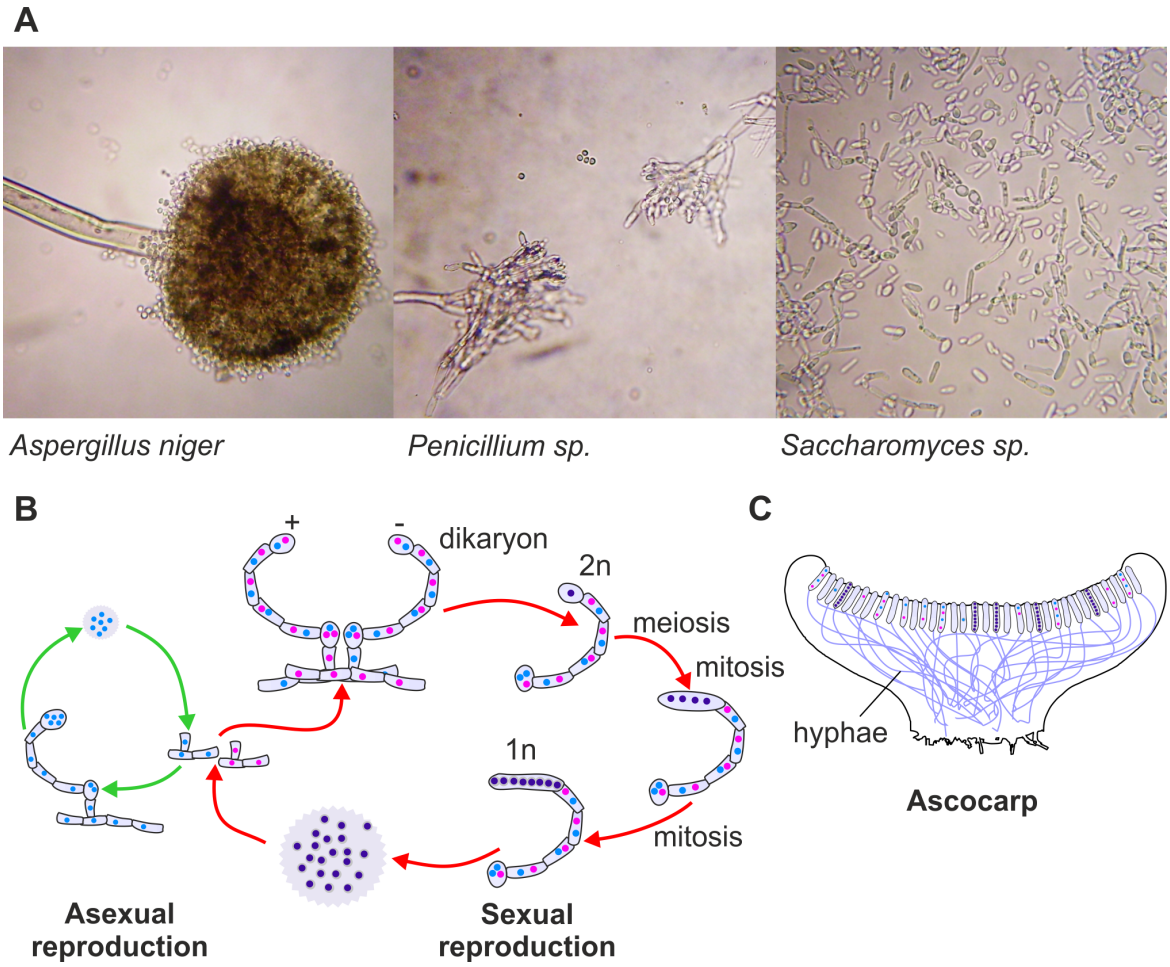


Figure 5.1. Biological cycle of ascomycetes fungi. A: Some members of the ascomycetes division. B: Sexual and asexual reproductive processes. During sexual reproduction, the ascogonium (+) and the antheridium (-) mate, and the resulting diploid ascus (2n) starts meiosis, after meiosis, the ascs divide during the mitosis process and give different recombinant ascospores (1n). In asexual reproduction, one mating type form one type of spores starting from one mycelium. C: Ascocarp containing asci with different states of maturation.

5.1.2 General characteristics of *Fusarium* and filamentous fungi

The genus *Fusarium* was first described in 1809 by Link [271]. Later in 1935 it was defined as a group of various species and varieties [272] and were finally unified as a unique fungal genus in 1940 by Snyder and Hansen [273]. *Fusarium* spp. belong to the division Ascomycota and produce mainly three different types of spores known as macroconidia, microconidia and chlamydospores (**Figure 5.2**). Given that some *Fusarium* species can produce the three types and others do not [274], this feature is used as an aid to identify their different species and varieties [275, 274].

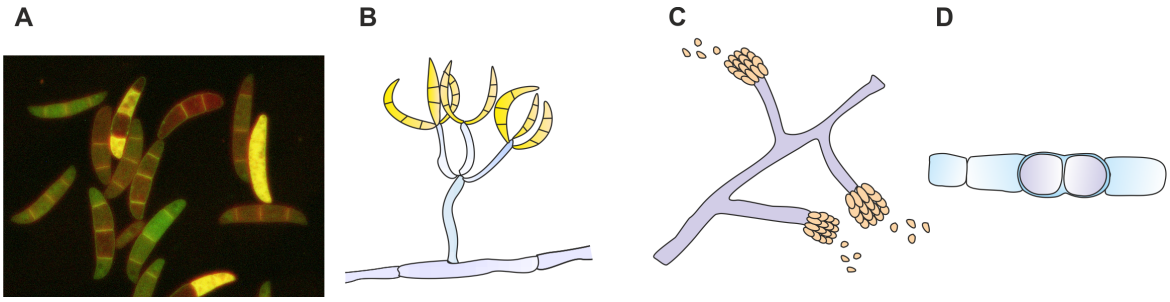


Figure 5.2. A schematic diagram of the spores produced by ascomycetes. Macroconidia spores from the Genus *Fusarium* labeled with a fluorescent analog of phosphatidylcholine [280] (A). The different fluorescent colors are due to the excitonic interactions between the fluorescent group BODIPY*. Schematic representation of the three main different spores found in the division Ascomycota: macroconidia (B), microconidia (C) and chlamydospores (D).
* Photo donated by Dr. U. Acuña from the Institute of Physical Chemistry "Rocasolano".

Fusarium oxysporum belongs to the family *Nectriaceae* and as other ascomycetes, are predominantly found as benign saprophytes [276, 277], plant symbionts [278, 279], and at the same time, certain strains can be found as animal and plant pathogens as well. Although it is uncertain if all ascomycetes present sexual cycles, various genomic analysis of *Fusarium oxysporum* revealed the presence of putative proteins involved in mating and chemotropic processes that are also found in other ascomycetes species like *S. cerevisiae* [270].

5.1.3 Economic and health impact of *Fusarium* species

Fusarium is one of the most widely known pathogenic fungi. It produces several different toxic compounds that can affect both plants and animals [281]. Diverse species and strains belonging to the genus *Fusarium* have been identified as the main causes of human and animal poisoning through the ingestion and direct contact of the mycotoxins produced by the fungus (**Table 5.1**) [281]. *Fusarium* toxins (see **Table 5.1**) can also be found in prepared animal feeds and human food products, including beer and processed foods [281, 282]. In plant industry, it is the cause of massive crop damages and millions of dollars are lost annually worldwide. The three main crops produced in the world (maize, wheat and rice) are severely affected by this pathogen and suffer the greatest impact of *Fusarium* micotoxins [283, 281]. Additional plants affected include bananas, asparagus, figs, forage grasses, soybean and other legumes, spice plants, medicinal plants, nut crops, tomato, melon, lettuce, and others [284]. It is unknown the exact number of toxigenic species id unknown, but it is known that the quantity of *Fusarium* species (including the non-toxic ones) can be greater than 300 [285–287].

Table 5.1. Main toxins found in *Fusarium* spp.[281]

Mycotoxin	Mechanism of action	Effects on animals and humans	Effects on plants
Trichothecenes	Inhibition of ribosomal protein synthesis, DNA and RNA biosynthesis, and mitochondrial function	Growth retardation, reproductive disorders, immune system suppression, anorexia, vomiting, hemorrhaging, diarrhea, and death.	Reduced seed germination, stunting of coleoptiles, roots, and shoots, chlorosis, wilting, and necrosis.
Zearalenone	Estrogenic activity	Reproductive abnormalities, enlarged mammary glands and genitalia, atrophy of ovaries or testes, infertility, reduced litter size, and reduced weight of offspring.	Unknown.
Fumonisins	Interference with sphingolipid biosynthesis	Liver and kidney lesions, apoptosis, necrosis and cell proliferation. It is also identified as the cause of certain cases of esophageal cancer in humans.	Inhibit the growth of shoots and roots, wilting, chlorosis, and necrosis.
Fusaric Acids	Inhibition of DNA synthesis	Changes in brain weight and body weight un rats.	Wilt symptoms.
Moniliformin	Inhibition of protein synthesis, cytotoxicity, and chromosome damage	Rapid death in poultry and swine.	Undetermined phytotoxicity.
Fusarins	Mutagenic	Disruption of chromosomes in mammalian cell cultures.	Undetermined phytotoxicity.
Beauvericin and Enniatins	Not determined	Antimicrobial activity, toxic to several insect species, and cytotoxic in laboratory studies with cell lines of insects and humans.	Unknown.
Fusaproliferin	Not determined	Toxic to brine shrimp larvae, cytotoxic to insect and human cell lines. Teratogenic to chicken embryos	Undetermined phytotoxicity.

Panama disease is a wilt disease produced by *Fusarium oxysporum* located at the roots of banana plants and other crops. In the decade of 1950's caused the worldwide decline of the production of the most popular banana in that time: the Gros Michel variety. This obligated to switch the production towards another resistant but less tasty variety known as Cavendish, which is the one currently commercialized [288]. However, *Fusarium* spp. is

evolving towards more virulent forms, and the Cavendish production is starting to experience the same outcome as other variants [288, 289]. The annual loss produced by the fungi in the United States has been estimated to be between 1 and 20 million US dollars per year. Nevertheless, during outbreak years, the annual losses can increase up to 30 or 46 million US dollars [289], without taking into account the economical loss in affected farm livestock and humans [281].

As a human pathogen, *Fusarium* spp. might be the second most common nosocomial fungal pathogen after *Aspergillus* [290]. Various plant pathogens are also known to affect humans, such as *F. solani*, *F. oxysporum*, *F. verticillioides* and *F. proliferatum* [291]. These species have been identified to be multiresistant and opportunistic, and are the main cause of fungal keratitis in the world. Simultaneously, it is believed to be the cause of 3 % to 5 % of the invasive fungal diseases in humans [292–295, 291].

The mortality rate of disseminated fusariosis in immunocompromised humans might be over 75 % [296–298, 294]. In Brazil, it is the leading source of invasive infections caused by molds; and in Spain, *Fusarium* is present within 1.2 % of the tissue biopsies and has been found to be more resistant than *Aspergillus* spp. as well [297]. Given its high mortality rate and invasiveness, there is an increasing clinical interest towards its control and treatment [290].

5.1.4 Fungal sex pheromones trigger hyphal growth

Experiments in *Saccharomyces* showed that yeast cells are sensitive to sexual pheromones trough a transmembrane G protein-coupled receptor (GPCR). Thus, when subjected to a gradient of peptide pheromones from the opposite sex, they produce cell projections (called shmoos) towards the stimuli. Within these sex pheromones, fungi produce two known peptide types termed a and α, which seem to be highly conserved within the ascomycota division [269, 299, 267, 270, 300].

In this context, α-pheromones of *Saccharomyces* species have been used as a paradigm in the study of the interaction of small peptides with their cognate GPCRs [301, 270]. Fungal GPCRs are transmembrane receptors, which once bound to their ligands, prompt various cellular responses such as cell cycle arrest, chemotropic growth towards the pheromone, shmoo projections and transcriptional reprogramming [269, 302, 267].

During the sexual cycle of yeast cells, there are two identified mating types named MATa and MATα; the recognition between them occurs when the MATα partner secretes α-pheromones and are recognized by the GPCR Ste2 of MATa cells. It also takes place the complementary process, in which the a-pheromone secreted by the MATa mating partner

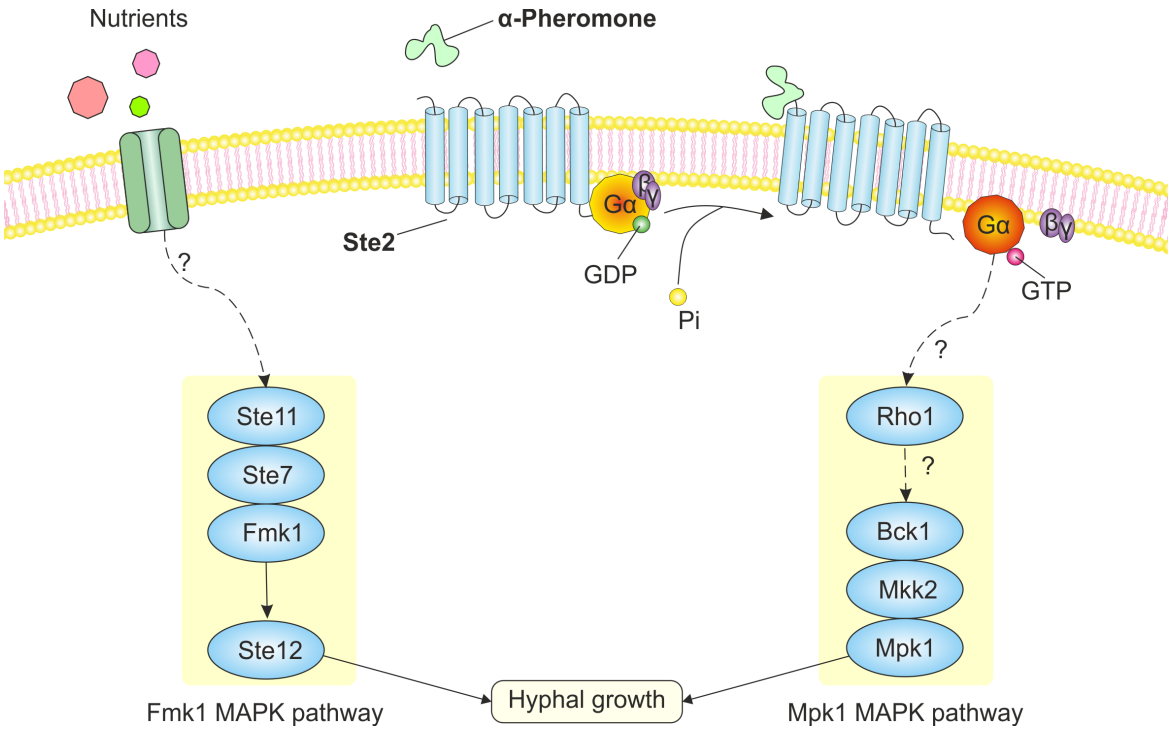


Figure 5.3. Chemotropic signaling routes in *Fusarium oxysporum*. Nutrients can act as chemical signals and activate the Fmk1 MAPK pathway. When the G-protein coupled receptor Ste2 recognizes the α-pheromone, the Gα subunit dissociates from the receptor and from Gβ/γ subunits. Then, Gα exchanges GDP with GTP and through different unknown processes, the Mpk1 MAPK pathway starts. Both Fmk1 and Mpk1 pathways induce hyphal growth [270].

binds to its cognate GPCR Ste3 of MATα. This prompts the approach of the haploid cells of each mating type to promote sexual reproduction [269, 302, 303, 267].

However, there is still little known about these processes in *Fusarium oxysporum* and other filamentous fungi, but there are new evidences supporting the hypothesis in which within this genus, similar pathways previously seen in *Saccharomyces* may also take place. As seen in **Figure 5.3**, the receptor Ste2 of *Fusarium oxysporum* can recognize either plant peroxidases or its α-pheromone, and activate the Mpk1 MAPK (Mitogen-Activated Protein Kinase) pathway; which triggers the directed hyphal growth [270].

5.1.5 Structural insights on the fungal α-pheromone of *S. cerevisiae*

In α mating cells, the locus MAT codes for the α factor and its postraductional modifications result in a mature α-pheromone. The α factor gene is first translated into a 165 residue-long polypeptide, which is glycosylated in the endoplasmic reticulum. Once in the Golgi apparatus, it is subjected to several modifications, including proteolysis processes that bring the mature

13 residue-long α-pheromone [304–306, 301]. In contrast with α-pheromones, a-pheromones contain both lipid and peptide components. The postraductional modification of a 36 or 38 (corresponding to the MFa1 and MFa2 genes, respectively) residue-long precursor occurs in the cytosol, where it is submitted to several modifications (including farnesylation and proteolysis) and gives the final mature a-pheromone, which is composed by a 12 amino acid-long peptide covalently bound to a phospholipid [307, 308, 305, 309, 310].

Although no 3D structure is available for the mature α-pheromone of *S. cerevisiae* (**Figure 5.4**), it is known that there is a loop composed by the central residues Pro8 and Gly9 that form a Type II β-turn, and seems to be highly conserved within yeast fungi. Diverse mutational studies have been done to identify the regions that might trigger specific functions on the cell. For example, D-Ala replacement of the central residues Pro8 and Gly9 leads to a decline of the pheromone affinity with its receptor. Simultaneously, D-Ala replacement of residues closer to the N- and C-termini lead to less activity and binding affinity, respectively. This means that residues closer to the N-terminus play an important role towards the Ste2 activation, residues located at the C-terminus are essential for receptor recognition and binding, while the loop is crucial to orient the signaling and binding regions of the pheromone towards its receptor [312, 313].

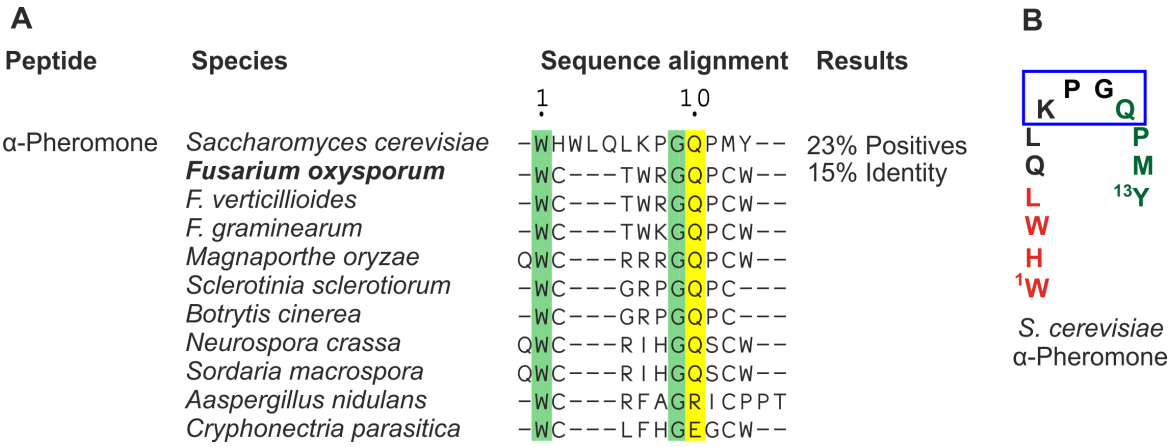


Figure 5.4. (A) Alignment of the sequences of *S. cerevisiae*, *F. oxysporum* α-pheromones and its orthologues from ascomycete fungi. The pheromone studied in this Thesis is depicted in bold. The positives percentage indicates the number and fraction of residues for which the alignment scores have positive values (yellow), indicating similar residues, the identity reflects the percentage of residues that are identical (green). Identical residues are listed on the bottom rows. Analysis generated with ESPrint [109]. (B) Conformation adopted by the *S. cerevisiae* α-pheromone as defined by [311]. Red residues correspond with the signaling region, green residues belong to the binding region and loop residues are boxed.

However, concerning the structural features of the α-pheromone in other filamentous fungi such as *Fusarium*, there is no information currently available and given their low sequence similarity with the yeast α-pheromone (**Figure 5.4**), it is uncertain whether ascomycete fungi's α-pheromones behave in the same way. In this regard, new information about structure-function relationship for this type of α-pheromones would be extremely valuable.

5.1.6 Surfactants as mimic molecules of the pheromone's receptor environment

As mentioned in the Introduction chapter, diverse lipids compose the cell membrane and can vary depending on the biological system: phospholipids, sphingolipids and sterols. For instance, the main phospholipids that can be found in *S. cerevisiae* are phosphatidylinositol (representing the 27.7 % of the membrane composition), phosphatidylcholine (17 %), phosphatidylethanolamine (14 %) and phosphatidylserine (3.8 %) [314, 315]. Given that the Ste2 receptor of the α-pheromone is a transmembrane protein that surpasses the lipid bilayer seven times, it is biologically relevant to study the possible natural environment of the pheromone prior, during or after its binding with its cognate receptor.

Due to the small size of α-pheromone sequences and their hydrophobic composition, their study in solution has some experimental difficulties. In this regard, NMR spectroscopy might be the best methodology to study the structural features with atomic resolution in a biological relevant environment. Nonetheless, the size limits shown by solution NMR make necessary the use of simplified systems that substitute the big lipid bilayers. In this aspect, DPC micelles might be a good system, as it is a surfactant widely used in NMR studies, and its choline group is shared with one of the most abundant phospholipids in fungal membranes. Also, as previously described in the Introduction chapter, surfactant micelles mimic the membrane interface due to the presence of a polar head group and a hydrophobic tail on each monomer. Though, three main considerations should be taken into account in relation with the use of DPC in solution NMR experiments: i) high concentrations of surfactants are needed and perdeuterated samples are usually necessary, ii) their high critical micelle concentrations (CMC) make them disadvantageous for NMR studies, as their monomers denature most globular proteins [316], and iii) sometimes it is difficult to determine if the protein or peptide is interacting with either the micelle o with the individual monomers present in the equilibrium [49].

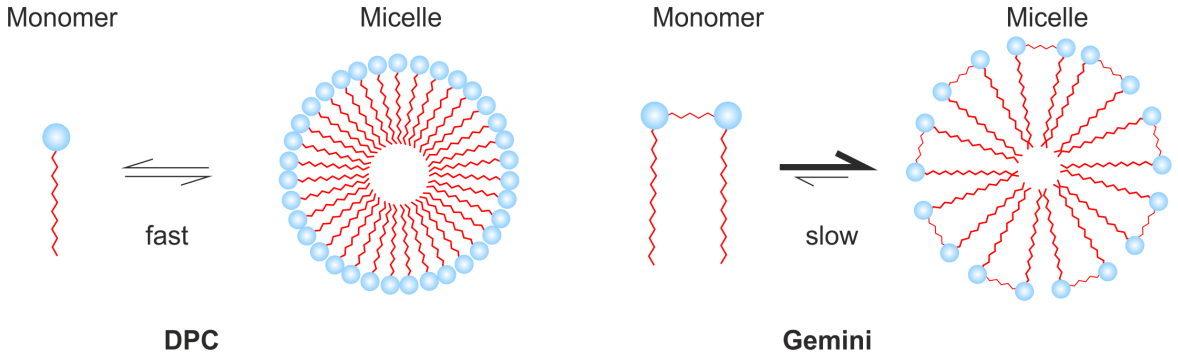


Figure 5.5. DPC and Gemini micelles as membrane mimetics. DPC and other surfactants widely used in NMR have high CMC values and present fast micelle ⇌ monomer kinetics. The concentration of the monomer in the equilibrium is hight and it is difficult to determine if the molecule is either interacting with the monomer or the micelle. Gemini surfactants have a much lower CMC values and less free monomeric surfactant is present in the media [317].

To surpass the present limitations, Gemini surfactants might be a good alternative solution. They present a low CMC, slow millisecond monomer to micelle kinetics and smaller micelle sizes (**Figure 5.5**) [50]. Hence, their properties make them excellent for NMR study. Given that low Gemini concentrations are enough to trigger micelle formation [49, 317], no deuteration is required for NMR studies. In addition, the low percentage of monomers in the equilibrium guarantees that the peptide intermolecular interactions (if present) established are with the micelle and not with the surfactant monomers. Nevertheless, the main problems for its use are that they are not usually commercially available and that these micelles experience a high curvature due to their small size.

Taking into account all the aforementioned information, it will be interesting to understand the molecular mechanism underlying the interaction between α-pheromones and their receptors near the membrane and the subsequent activation of the corresponding Ste GPCR. To reach this goal, new data about the structure-function relationships for this type of α-pheromones within membrane environments would be mandatory.

5.1.7 Objectives

The main objective of this study is to obtain the structural and functional insights of the α-mating pheromone of the highly destructive ascomycete plant pathogen *Fusarium oxysporum* employing diverse biophysical techniques. Given that there is no structural information concerning these systems in filamentous fungi, we have employed NMR spectroscopy as the central tool to solve the 3D structure and characterize possible interactions of the pheromone with a substitute of the fungal cell membrane. In this particular case, NMR is the technique of choice given the highly hydrophobic nature and short sequence of the peptides that limit their structural characterization by other biophysical methods. However, as explained before, due to the current size limitations of liquid NMR spectroscopy, simplified versions of the cell membrane have to be used.

The specific objectives proposed are:

- To design α-pheromone analogs to determine the residues implicated in the structure and functional properties.
- To fully characterize these peptides (the wt sequence and three designed analogs) in H_2O and TFE (non-polar solvent) through biophysical methods and define if there are residues and/or localized regions essential for its structure formation.
- To elucidate the regions and residues that might be crucial for specific functions like the chemoattractant and growth inhibitory activity. This objective has been proposed in collaboration with the teams of Prof. Antonio Di Pietro and Álvaro Martínez del Pozo.
- To determine the structural relevance of the disposition of each residues and their implications with possible membrane interactions employing DPC and Gemini micelles as membrane mimetics.
- To provide a final model of the possible peptide-membrane interaction based on the experimental atomic data.

5.2 Materials and methods

5.2.1 Chemicals

The deuterated compounds [D_3] TFE (2,2,2-trifluoroethanol) (99 %), [D_{38}] DPC (dodecylphosphocholine) (98 %), and D_2O (99.9 %) were obtained from Cambridge Isotope Laboratories (USA). The percentages of deuteration are indicated in parenthesis. Non-deuterated DPC was from Sigma. Gemini surfactant (pentanediyl-1,5-bis(hydroxyethylmethylhexadecylammonium bromide)) was synthesized by Dr. Razieh Amiri, University of Isphahan, Iran. The estimated CMC of the Gemini surfactant is about 2 - 3 μM at 25 ºC [318], and that of DPC about 1 mM [319].

5.2.2 Peptide synthesis

Peptide sequences are listed in **Table 5.2**. Synthetic *F. oxysporum* wt α-pheromone, and a scrambled (Scr) version thereof were synthesized by Caslo ApS (Denmark), and di-alanine substituted analogs D-Ala^{1,2} and D-Ala^{6,7} were obtained from GenScript (Piscataway, NJ). Fmoc-solid phase synthesis protocols were followed and purification was performed by reverse-phase HPLC using a C18 column until reaching 97 % or more purity. The identity of the peptides was confirmed by MALDI-TOF MS.

α-Pheromone: WCTWRGQPCW RP-HPLC: tR = 6.40 min, purity 97.71 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 30-42 % B buffer gradient in 12 min). HRMS: Theoretical molecular mass: 1322.54 (m/z); $[M+H]^+$ found: 1322.88 (m/z).

Table 5.2. Sequences of the peptides studied.
* a: D-isomer of alanine.

Peptide	Sequence	Theoretical pI
α-Pheromone	WCTWRGQPCW	8.07
Scrambled	WRWPCCWGQT	8.07
D-Ala^{1,2}	aaTWRGQPCW	8.29
D-Ala^{6,7}	WCTWRaaPCW	8.07

Scrambled: WRWPCCWGQT RP-HPLC: tR = 19.41 min, purity 99.68 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2 (v/v); buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 20-45 % B buffer gradient in 25 min). HRMS: Theoretical molecular mass: 1322.54 (m/z); $[M+H]^+$ found: 1322.84 (m/z).

5.2.3 Dynamic light scattering

Particle size and polydispersity index of aged samples of the wt and Scr peptides (61 μM) in H_2O, 20 mM DPC and 2 mM Gemini were determined by dynamic light scattering (DLS) with a DynaPro MS/X (Wyatt Inc) spectrometer. Twenty acquisitions of 10 seconds at 25 ºC were obtained. Water was used as blank for data analysis; 20 mM DPC and a fresh wt α-pheromone solution were used as control.

5.2.4 Mass spectrometry

* Mass spectrometry assays were performed by the MS facility of the Institute of Physical Chemistry "Rocasolano", Spanish National Research Council (CSIC).

Soluble fractions of the α-pheromone (aged NMR samples prepared in H_2O) were characterized through a MALDI-TOF Voyager-DE PRO (Applied Biosystems) mass spectrometer with an acquisition mass range between 500 and 10000 Da embedded in a 2,5-dihydroxybenzoic acid matrix.

5.2.5 NMR sample preparation

NMR samples were prepared at 0.1 - 0.5 mM concentration in H_2O, H_2O/TFE (7/3 v/v), 20 mM [D_{38}] DPC, 20 mM [D_{38}] DPC/non-deuterated DPC (1:1 v/v) or 2 mM Gemini in H_2O/D_2O (9:1 v/v). For ^{13}C-HSQC natural abundance acquisition, the samples were lyophilized and dissolved in 100 % D_2O. The pH was measured with a Hamilton glass micro-electrode and adjusted to 5.0 by adding minimal amounts of NaOD or DCl. All samples were placed in 5 mm NMR tubes, and contained DSS as the internal reference for ^{1}H chemical shifts.

5.2.6 NMR spectra acquisition

NMR spectra were recorded on a range of 5 ºC and 40 ºC with Bruker Avance spectrometers, operating at 600 and 800 MHz (^{1}H), both equipped with triple resonance, z-field gradient cryoprobes.

Membrane-like assays were performed with both fully or half-deuterated DPC. Gemini was only used in the protonated form. In order to analyze sample evolution in H_2O and micelles, a set of measurements were performed over time spanning several weeks (from 0 to 260 days) depending on the media.

Phase-sensitive two-dimensional total correlated spectroscopy (TOCSY) spectra were obtained by using 20 and 60 ms mixing times. The nuclear Overhauser enhancement

Table 5.3. NMR spectra conditions of the peptides studied.

Peptide	**Solvent**	**Temperature (ºC)**
α -Pheromone	H_2O	5, 25
	D_2O	5, 25
	30 % TFE/H_2O	5
	30 % TFE/D_2O	5
	20 mM DPC/H_2O	5, 25, 35
	20 mM DPC/D_2O	5, 25, 35
	20 mM 50 % deuterated DPC/D_2O	25
	2 mM Gemini/H_2O	5, 25
	2 mM Gemini/D_2O	5, 25, 35, 40
Scrambled	H_2O	5, 25
	D_2O	5, 25
	30 % TFE/H_2O	5
	30 % TFE/D_2O	5
	20 mM DPC/H_2O	5, 25, 35
	20 mM DPC/D_2O	5, 25, 35
	20 mM 50 % deuterated DPC/D_2O	25
	2 mM Gemini/H_2O	25
D-Ala^{1,2}	H_2O	5, 25
	D_2O	5, 25
	30 % TFE/H_2O	5
	30 % TFE/D_2O	5
D-Ala^{6,7}	H_2O	5, 25
	D_2O	5, 25
	30 % TFE/H_2O	5
	30 % TFE/D_2O	5

spectroscopy NOESY mixing time was 150 and 50 ms to discard NOE signals from spin diffusion. ^{1}H-^{13}C single quantum coherence (HSQC) spectra were recorded at ^{13}C natural abundance. ^{13}C δ-values were indirectly referenced by using the IUPAC-IUB recommended ^{1}H/^{13}C chemical shift ratio (0.25144953) [101]. Water signal was suppressed by either presaturation or by using a 3-9-19 pulse sequence. The 2D data matrices were multiplied by a square-sine-bell window function with the corresponding shift optimized for every spectrum and zero-filled prior to Fourier transformation. Baseline correction was applied in both dimensions. Data were processed with the standard TOPSPIN program (Bruker Biospin, Karlsruhe, Germany). NMR and sample conditions are summarized in **Table 5.3**.

5.2.7 NMR spectra assignment

Assignments of the ^{1}H spectra were done following the sequential assignment protocols [94] with the help of the SPARKY software [95]. The ^{13}C resonances were identified based on the correlations between the hydrogen and the bound carbon atom present in the ^{1}H-^{13}C-HSQC spectra. Peptides exhibit limited solubility in H_2O and are generally more soluble in TFE, DPC and Gemini media. Thus, in some cases, the evaluation of weak NOE signals and the complete assignment of the minor isomers (Xxx-Pro *cis, trans*) were difficult or impossible. All assigned NMR chemical shifts are listed in the **Appendix D**. DPC and Gemini spectral assignment were performed based on their previously reported chemical shift list [49, 158].

5.2.8 Structure calculation based on NMR restraints

Structure calculations of the peptides in H_2O, TFE, DPC and Gemini were done with the CYANA 2.1 program. Structure calculation was performed from the distance and dihedral angle constraints derived from NMR parameters using the manual or the standard iterative procedure for automatic NOE assignment [96]. Distance constraints were obtained from the cross-peaks present in 150 ms 2D [^{1}H-^{1}H]-NOESY spectra, which were integrated using the standard SPARKY integration sub-routine [95], and the dihedral angle restraints for Φ and Ψ angles were derived from ^{1}Hα, ^{13}Cα and ^{13}Cβ chemical shifts using the TALOS-N webserver [97]. Typically, 100-200 structures were calculated using standard protocols. Since no ambiguous constraints were found upon examination of the list of distance constraints, the 20 lowest target function structures were selected as the final structure ensembles. The structural statistics data for these structures are provided in **Table 5.5** and **Table 5.6**. The structural ensembles were visualized and examined using MOLMOL [98] and PyMOL [99].

5.2.9 Micelle interaction evaluation

Micelle-peptide interactions were identified based on the ^{1}H assignment of both components and evaluated from intermolecular NOEs found in the partially or non-deuterated micelle media. For the analysis, the intensities of all NOEs between the micelle and the protons of each residue were added and scaled with respect to a known distance and well-separated signal, the intraresidual Trp Hδ1-Hε1 NOE cross peak. The resulting bar plot, together with the intermolecular NOE intensities, were employed to identify the peptide residues and their more frequent location within the distinguished micelle regions (hydrophobic core or polar surface). This qualitative analysis was used to build a graphical model of the peptide-membrane interaction.

5.3 Results

5.3.1 Peptide design and cosolvent selection

To demonstrate the role of the different residues in the α-pheromone structure and activity, we have studied the wt sequence and three designed variants (**Table 5.2**). Based on previous experiments [311], sequences with two D-Ala substitutions of contiguous residues in positions 1 and 2 (D-Ala^{1,2}), and 6 and 7 (D-Ala^{6,7}) were used to analyze the relevance of the N-terminus and the central residues Gly6 and Gln7, respectively. A non-natural peptide consisting on a randomly scrambled sequence with the same residues of the α-pheromone was used as a control. To design this control peptide, 10 different scrambled α-pheromone sequences were generated using the Scrambled software (http://www.mimotopes.com/peptideLibraryScreening.asp?id=97). Among these, we selected a sequence that does not carry the conserved N-terminus Trp1-Cys2 and the central Gly6-Gln7 dipeptide, and has two contiguous Cys residues at positions 5 and 6. All the peptides were studied in H_2O and TFE solvents.

Taking into account that the corresponding α-pheromone's GPCRs appear embedded within the plasmatic membrane, it is biologically relevant to analyze the potential influence of a lipid environment over the structural properties of the *F. oxysporum* α-pheromone. Given that it is not feasible to employ solution NMR to study further interactions with lipid bilayers and vesicles, it is necessary to employ lipid membrane mimetics [49, 320–322]. As mentioned in the introduction section, traditional DPC, together with Gemini micelles were used as cosolvents for the wt and the scrambled peptides in this study.

5.3.2 NMR characterization of α-pheromone related peptides

A first qualitative analysis was performed based on the NMR data of the peptides under different conditions and solvents (**Table 5.3**). All the assigned chemical shifts can be seen in the **Appendix D**.

It is very well accepted that the chemical shift deviation analysis can bring qualitative information regarding to the secondary structure of peptides as a function of the sequence [81, 132, 101]. However, in the present examples, Δδ Hα, Δδ Cα and Δδ Cβ, were not informative as no classical or recognizable profiles were obtained. This might be due to the chemical shift anisotropy caused by the aromatic rings of the Trp residues [323]. Thus, in the case of the α-pheromone, scrambled and D-Ala^{6,7} sequences, 3 out of 10 residues

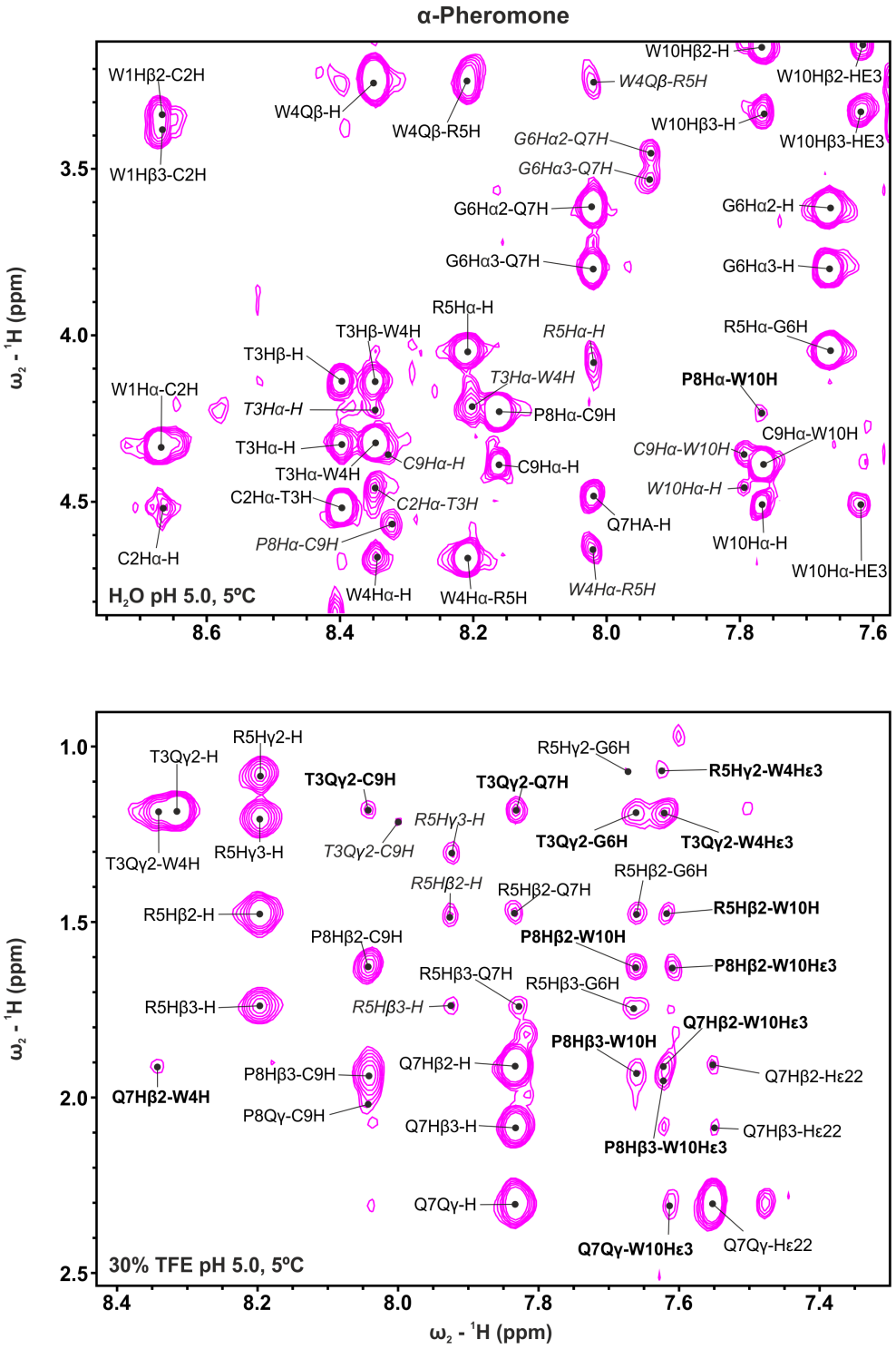


Figure 5.6. Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of α-pheromone in H_2O (upper panel) and 30% TFE (lower panel), 5 ºC, pH 5.0. Non-sequential NOEs are labeled in bold, and signals belonging to Gln-Pro *cis* conformation are in italics

correspond to Trp. In the D-Ala[1,2] variant, there are 2 Trp out of 10 amino acids in the sequence. Therefore, the evaluation of the secondary structure must rely entirely on the NOE data, as the chemical shifts are not informative given that the mentioned effects dominate their values.

Interestingly, several sets of signals corresponding to individual entities were found in all NMR spectra (water, TFE and micelles). Two of these groups of signals correspond with the *trans* and *cis* isomers of the Xxx-Pro bond, which are in conformational equilibrium (**Figure 5.6**). The other two sets of signals, as will be described below, correspond to the presence of the either reduced or oxidized states of the α-pheromone peptide in the micelle media (**Figure 5.7**). This means that for the α-pheromone, the sets of signals correspond as

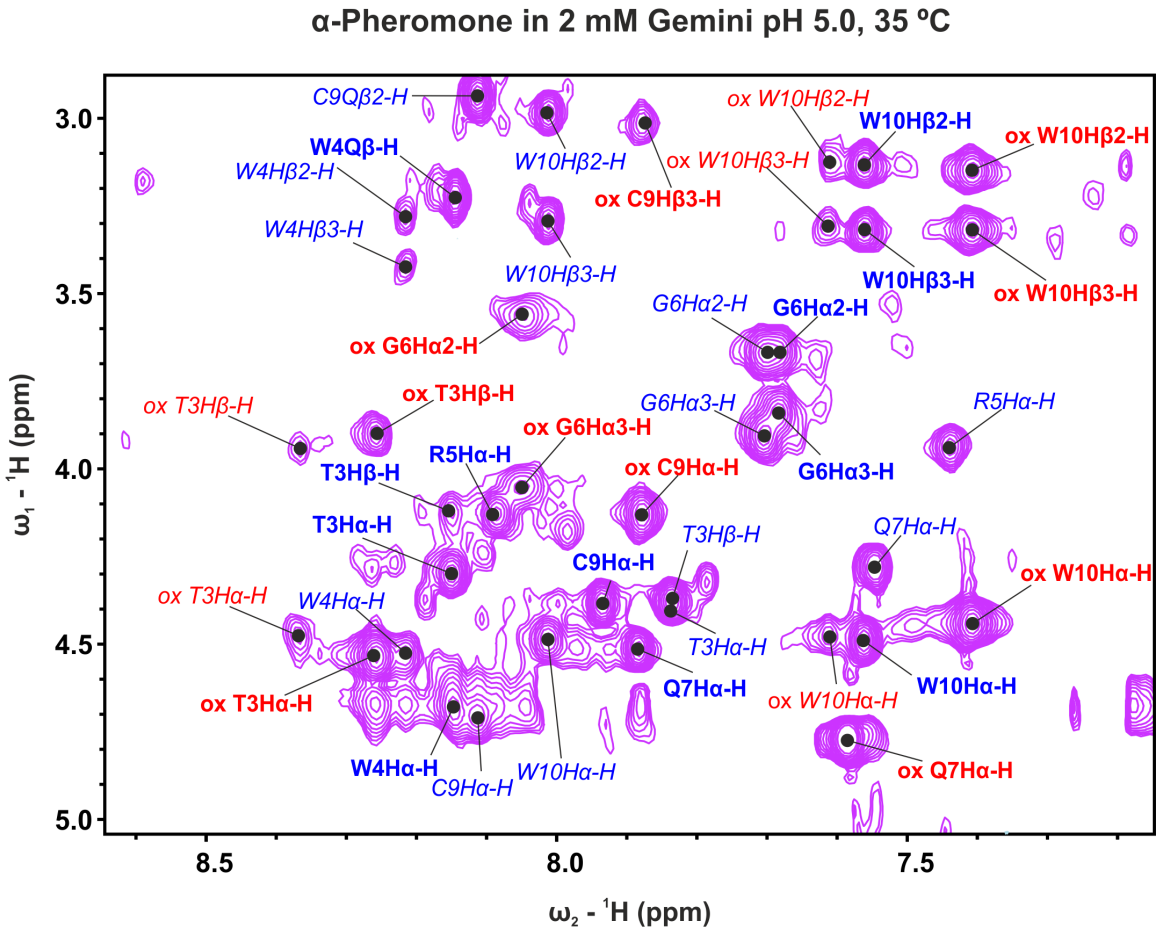


Figure 5.7. Selected region of the 2D ^{1}H-^{1}H NOESY spectra of α-pheromone in 2 mM Gemini micelles, 35 ºC, pH 5.0, showing the presence of four sets of signals. Signals corresponding with the reduced pheromone are in blue, signals corresponding with the oxidized state are labeled as ox, Gln-Pro *cis* conformation are in italics and Gln-Pro *trans* are in bold.

follow: reduced Gln-Pro *trans*, oxidized Gln-Pro *trans*, reduced Gln-Pro *cis* and oxidized Gln-Pro *cis* states. Given the relative population for each form, NOE signals were not enough to perform a successful structure calculation for the minor species.

The chemical shift difference between the Pro ^{13}Cβ and ^{13}Cγ (δ Cβ - δ Cγ ≈ 4.5 ppm) and the sequential NOEs found between the Hα of the X (i - 1) residue with Hδ and Hδ' protons of the Pro(i) residues, confirmed the *trans* rotamer as the major species for all peptides in all conditions, except for the scrambled in Gemini micelles. The proportion of the *cis/trans* ratio was not equal for all peptides and the relative intensities of the different Cβ and Cγ of Pro signals could not be properly evaluated in all conditions due to signal broadening and overlapping situations. In the α-pheromone and the D-Ala^{1,2} analog in H_2O and TFE, the ratio was approximately 7:1 towards the *trans*-Pro form. This corresponds to the expected *trans/cis* ratio for flexible linear peptides [324]. However, in the scrambled and the D-Ala^{6,7} analog the relationship change to around 1.5:1 (for the Trp-Pro bond in Scr) and 2:1 (for D-Ala-P bond in D-Ala^{6,7}), respectively. A similar effect was observed in other linear peptides when a Pro residue is preceded by an aromatic amino acid [324, 325], being this is the case of the scrambled sequence. Also, in the native state of proteins, there is a marked trend for an aromatic amino acid to precede a *cis* proline. In all four peptides tested, the equilibrium in water was not significantly affected in the presence of 30 % TFE.

5.3.3 Ageing peptides yield different species depending on the environment

The presence of four different species found in the spectra of the α-pheromone in DPC and Gemini micelles (**Figure 5.7**), showed the existence of the *cis/trans* forms for the peptides in two different states. In each new set of acquired spectra, additional signals appeared and gained intensity, while the original ones observed in the freshly prepared sample decreased and disappeared as a function of time.

In pure water, the evolution of these changes was visualized through weekly recordings of 1D spectra along 81 days. As seen in **Figure 5.8**, the 1D spectrum at day 23 starts to show new sets of signals. Simultaneously, the solubility of the peptide present in solution decays over time and the overall intensities are reduced significantly at day 81, suggesting the formation of some soluble aggregates of high molecular weight not observable in our NMR conditions. That sample was later subjected to mass spectrometry (MS) and dynamic light scattering (DLS) analysis. A final TOCSY, NOESY and ^{13}C-HSQC NMR spectra were recorded to obtain data of the molecular entities present in the aged sample in water. The

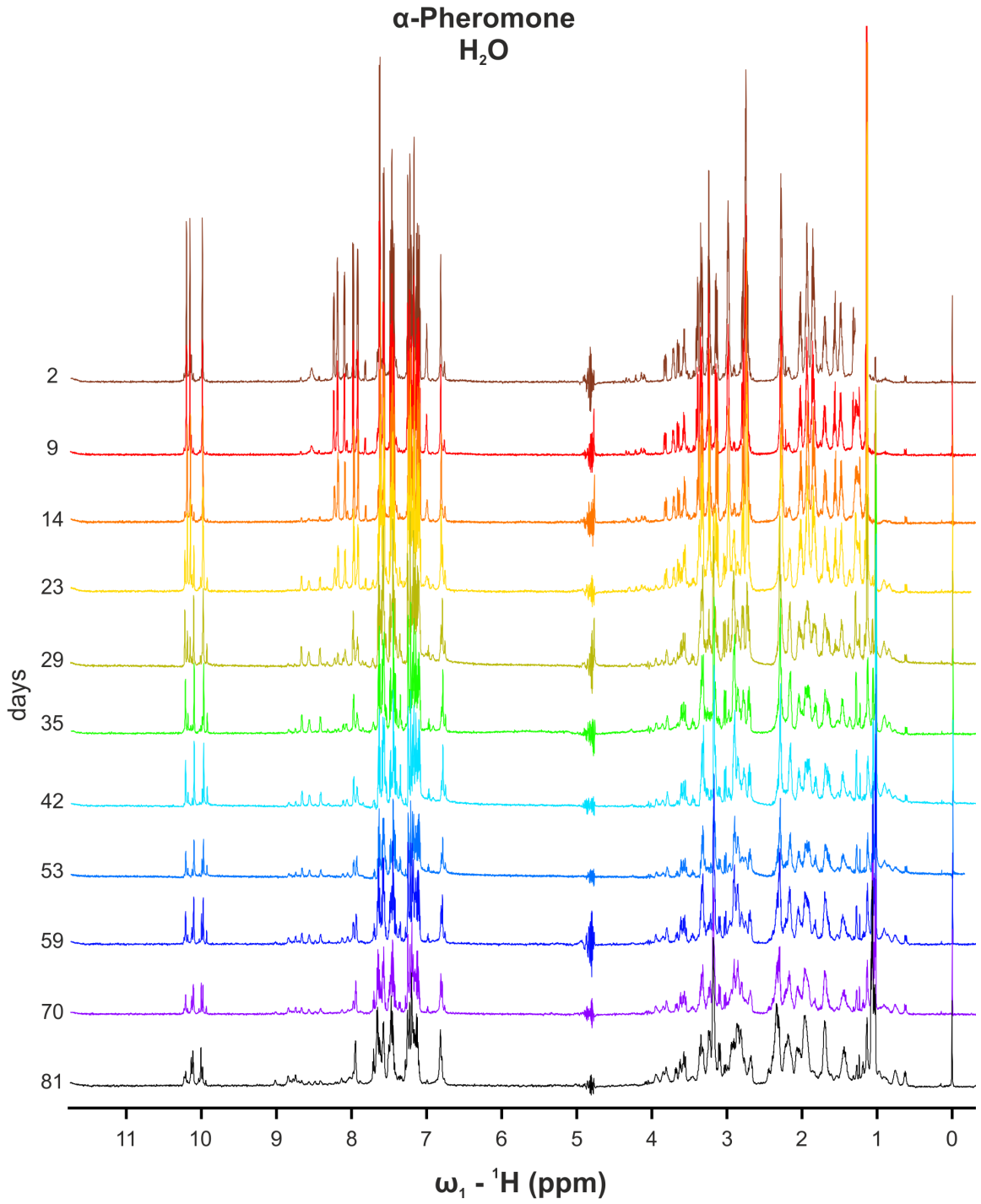


Figure 5.8. Overlapping 1D NMR spectra of α-pheromone in H_2O, 25 ºC, pH 5.0, as a function of time. Days are indicated. 1D spectrum at day 23 clearly starts to show new sets of signals, and the solubility starts to decay. By day 81, the intensities of the signals are significantly reduced.

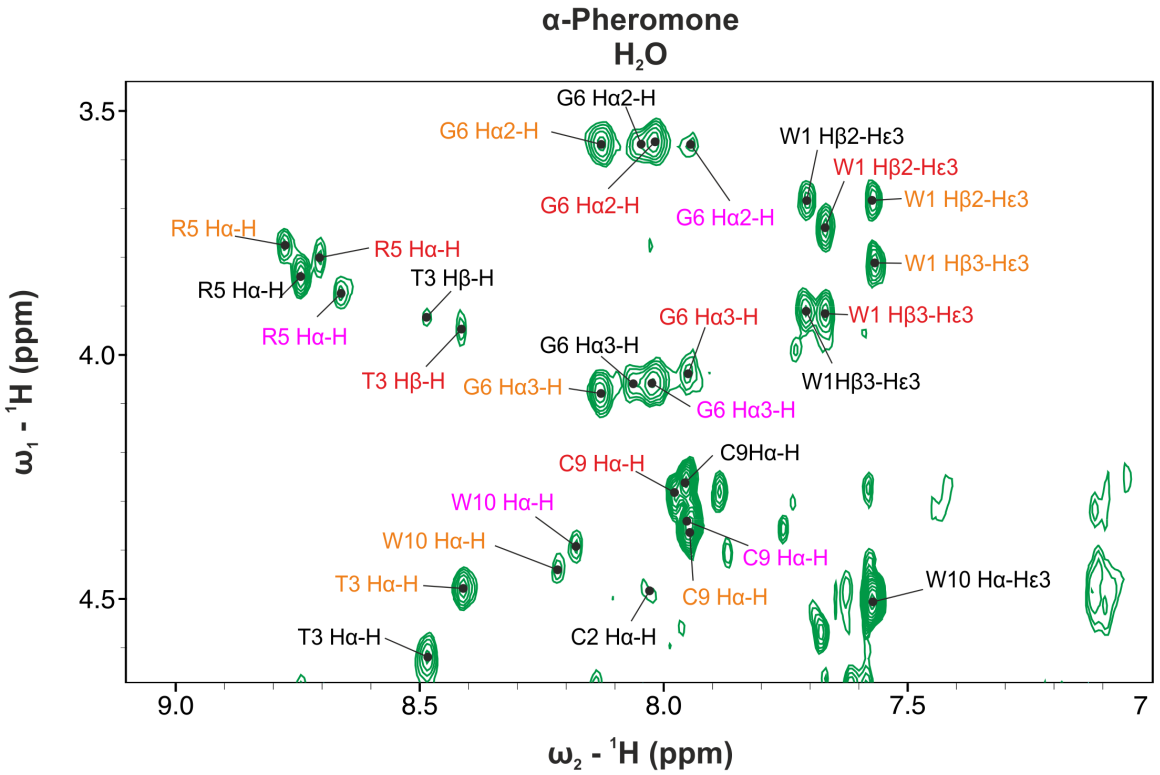


Figure 5.9. Selected regions of the 2D ^{1}H-^{1}H TOCSY spectra of α-pheromone in H_2O, 25 °C, pH 5.0. Different sets of signals are depicted in different colors (red, magenta, black and orange) could not be assigned to a specific molecular species.

analysis of the TOCSY (**Figure 5.9**) data show the presence of at least four different spin systems for each amino acid of the sequence. Unfortunately, the NOESY and HSQC spectra did not provide us enough information for the unambiguous assignment to each one of the specific species present in the solution, due to their complexity, different populations and low concentration of the detected forms.

The evolution of the α-pheromone and the scrambled sequence in the presence of micelles was also followed by NMR spectroscopy. Under these conditions, the samples maintained their solubility, suggesting that the micelles exert certain effect preventing peptide oligomerization. There is important information in the ^{13}C-HSQC spectra regarding to the oxidation state of the Cys side chains. ^{13}Cβ chemical shift values around 28 ppm indicate that the Cys is reduced, and they are around 41 ppm for its oxidized state (**Figure 5.10**). In the same way, Hβ nuclei chemical shift values corresponding to the reduced Cys are usually found around 2.93 ppm, and have values of 3.25 and 2.99 for the oxidized forms [101].

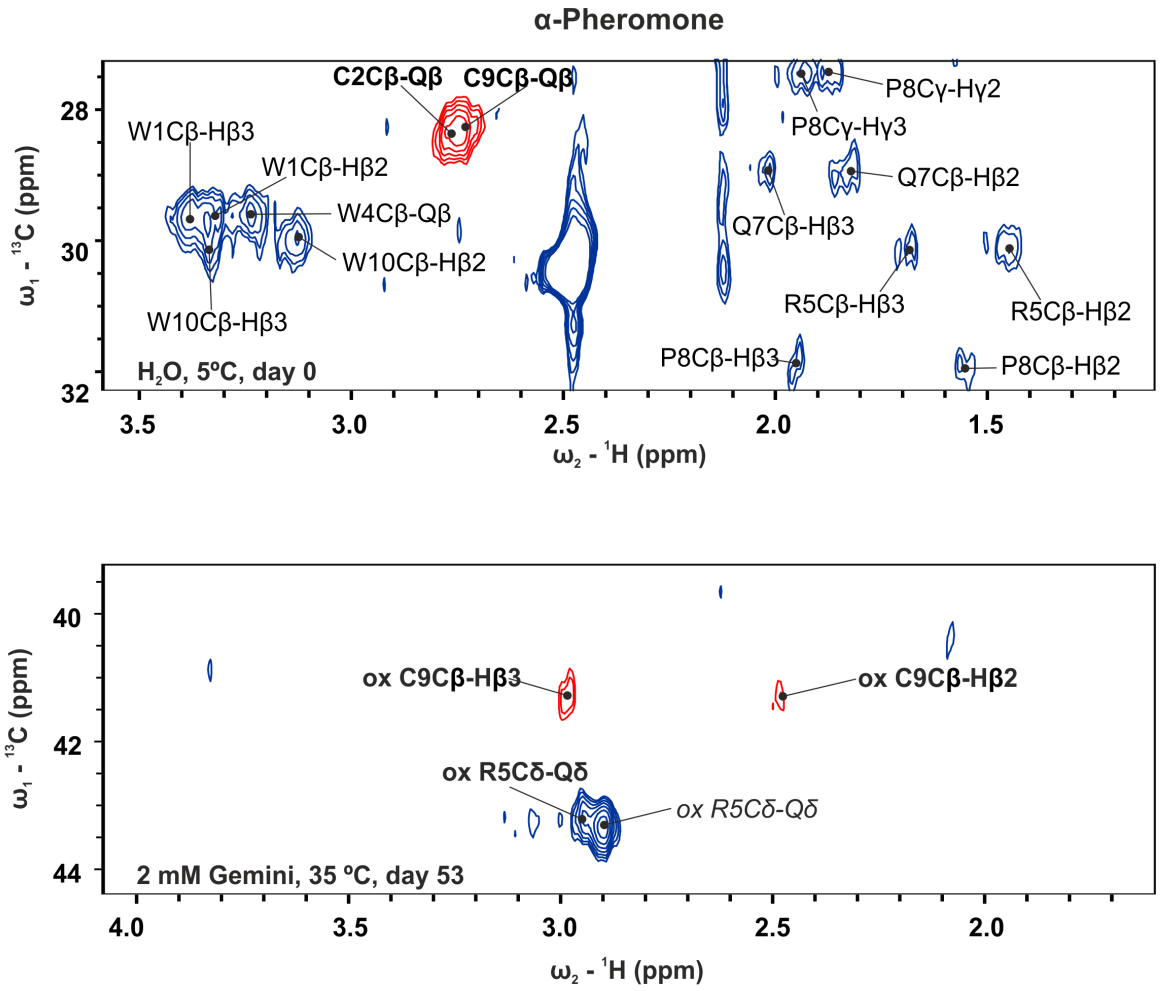


Figure 5.10. Selected regions of the ^{1}H-^{13}C HSQC spectra of reduced α-pheromone in H_2O (top panel), and oxidized α-pheromone in Gemini micelles media after 53 days from preparation (lower panel). Signals corresponding to the *cis* form are in italics. Signals corresponding to Cys residues are in bold and the oxidized Cys Cβ are shown in red.

Following this statement and analyzing each set of NMR spectra (TOCSY, NOESY and ^{13}C-HSQC) acquired at different times after sample preparation, it was confirmed the emergence of the α-pheromone oxidized monomeric form, involving Cys2 and Cys9. Starting on day 0, the initial reduced form was prevalent until the day 260, when only the signals corresponding to the oxidized form were found for both DPC and Gemini micelles. This oxidation between Cys2 and Cys9 renders an intramolecular S-S bond, which provides a dramatic change of the NMR parameters, including all backbone atoms and most side-chain protons (**Figure 5.11** and **Figure 5.12**). The oxidation of the Cys SH groups is explicitly revealed by changes in the chemical shift of the Cβ resonances, from 28 ppm in the reduced state to 41 ppm in the oxidized form (**Figure 5.10 B**). In addition, the two Hβ protons of

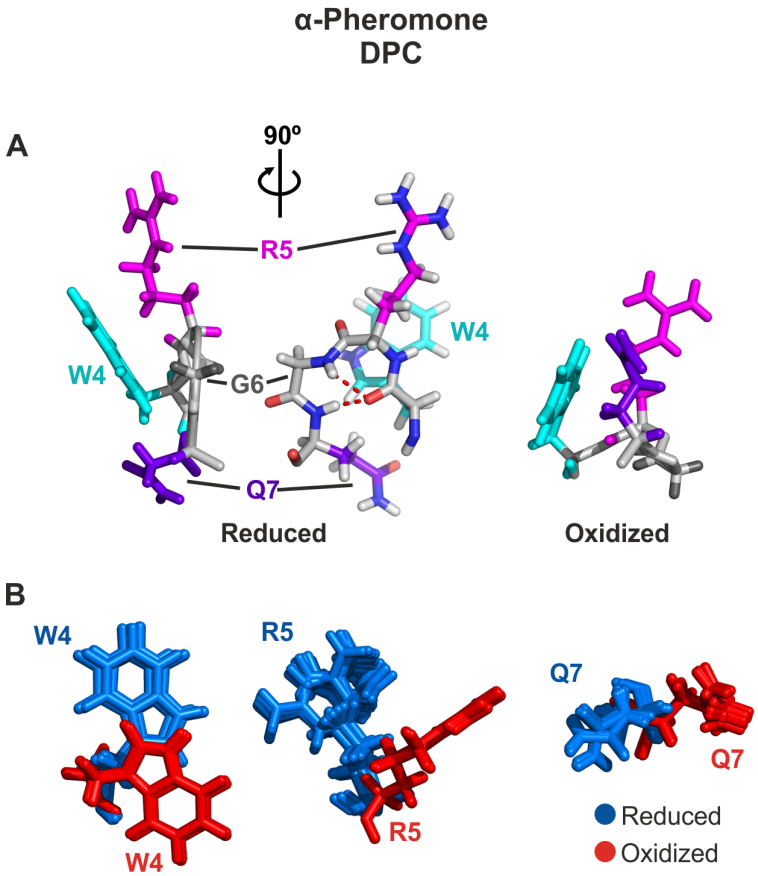


Figure 5.11. Differences in the side chains of α-pheromone between reduced and oxidized states in DPC. (A) Relative position of the residues located in the β-turn and stabilizing hydrogen bonds (red dashes) of the reduced and oxidized α-pheromone. (B) Comparison of the reduced and oxidized pheromone between the aligned amino acids Trp4, Arg5 and Gln7.

both Cys residues are discernable when oxidized, in contrast with undistinguishable values in the reduced form (**Figure 5.7**). Regarding to the remaining peptide signals, HN chemical shift values of Trp4, Arg5, and Gly6, were strongly affected by the oxidation and displaced to larger chemical shift values (**Figure 5.12**), whereas the Hα of Arg5 and Gly6 showed lower chemical shift values. These changes, in many cases larger than 0.5 ppm, both in the backbone and in the side chain protons, indicate an important conformational change due to the S-S bond formation. Interestingly, no oxidation was found during the same time period for the scrambled peptide, which contains two Cys at the adjacent positions 5 and 6.

Dynamic light scattering (DLS) was used to determine the presence or not, of high aggregates within the sample (**Table 5.4**). This analysis, when performed at day 102, revealed that the 70.6 % of the suspended mass belonged to a multimodal polydisperse distribution corresponding with molecules with radius larger than 196.6 nm (**Table 5.4**), indicating

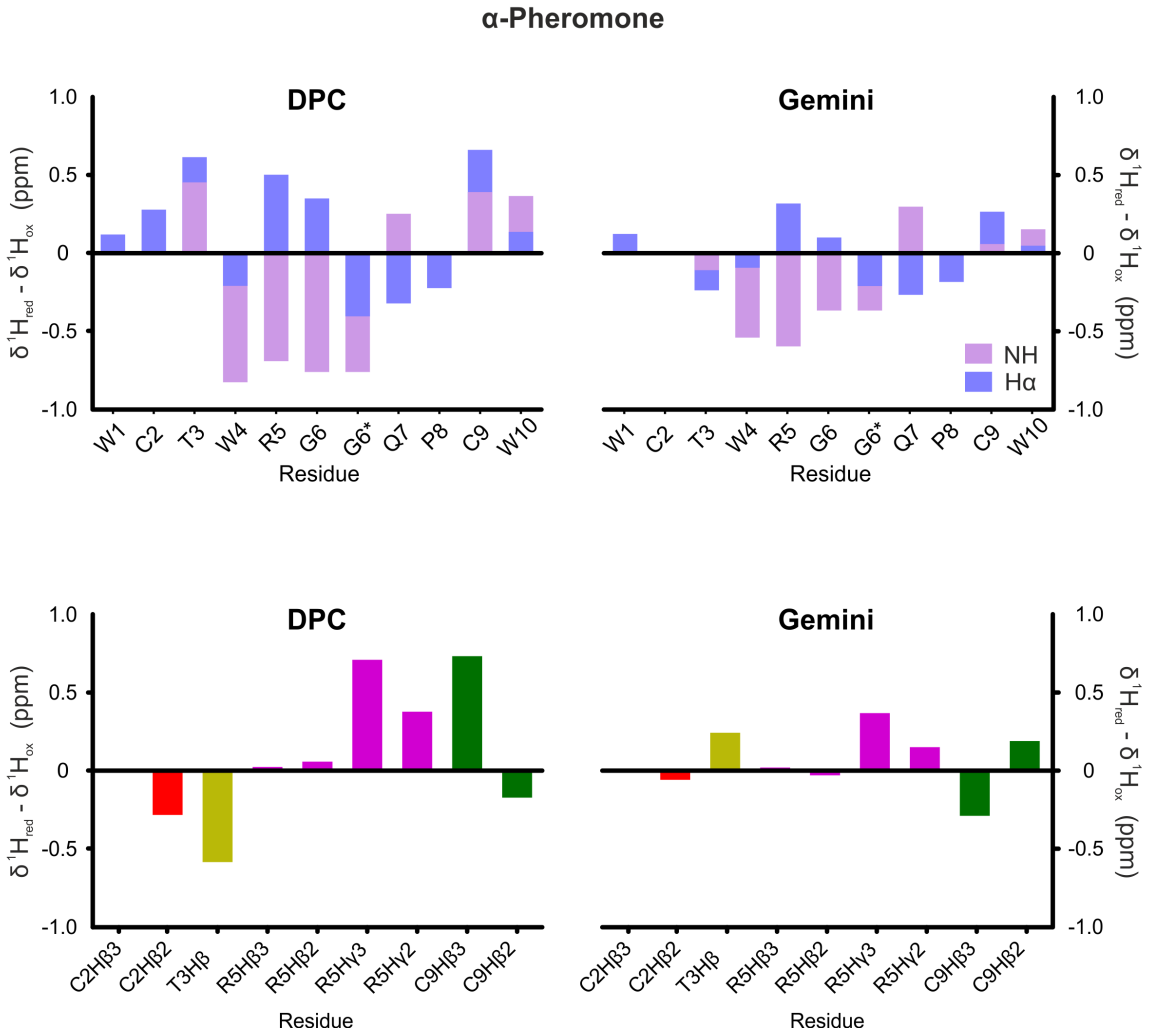


Figure 5.12. Proton chemical shift differences of the backbone HN and Hα (top) and selected side chains (bottom) between reduced and oxidized forms of α-pheromone in DPC and Gemini. The HN values of Trp4, Arg5, and Gly6 are strongly affected (differences over 0.5 ppm) by the oxidation and displaced to larger chemical shift values (upper panel). The side chains of Cys2, Thr3, Arg5 and Cys9 are also strongly affected (lower panel).
* Two bars are for Gly6, as glycine has two Hα.

the formation of different non-specific aggregates. The remaining 29 % corresponded to monomodal monodisperse aggregates of around 38.6 nm. These results differ from the ones obtained for the fresh monomeric form of the wt α-pheromone, with a single monodisperse peak equivalent to a radius of 2.3 nm in water (**Table 5.4**).

Finally, as shown by DLS analysis, it is also remarkable how both peptides in micelle media (α-pheromone and Scr sequences) showed either a monodisperse or polydisperse monomodal distribution, with radii between 2.3 and 3.1 nm, which remained constant over

Table 5.4. Dynamic Light Scattering analysis of α -pheromone and scrambled peptides in H_2O, DPC and Gemini, 25 ºC, pH 5.0.

Molecule	Solvent	Radius (nm)	Polydispersity (%)	Mass (%)
20 mM DPC		2.4	11.7	100
α -Pheromone	H_2O (0 d)	2.3	0	99.5
	H_2O (102 d)	> 196.6	> 16.8	70.6
		38.6	0	29.4
	DPC (249 d)	2.3	30.1	100
	Gemini (305 d)	3.1	13.6	100
Scrambled	H_2O (0 d)	14.2	0	100
	H_2O (674 d)	6.5	0	100
	DPC (449 d)	2.4	17.6	100
	Gemini (305 d)	2.7	15.2	99.8

time (**Table 5.4**). This observation was in clear contrast with the results obtained in water solution, showing that in the presence of micelles, the aggregation process was not taking place, or at least the process was much slower.

To support these findings, the sample that remained soluble of the aged α-pheromone was subjected to mass spectrometry analysis. The results showed two main peaks with 1322.60 and 1320.59 mass units (**Figure 5.13**): the first one would correspond to the molecular mass

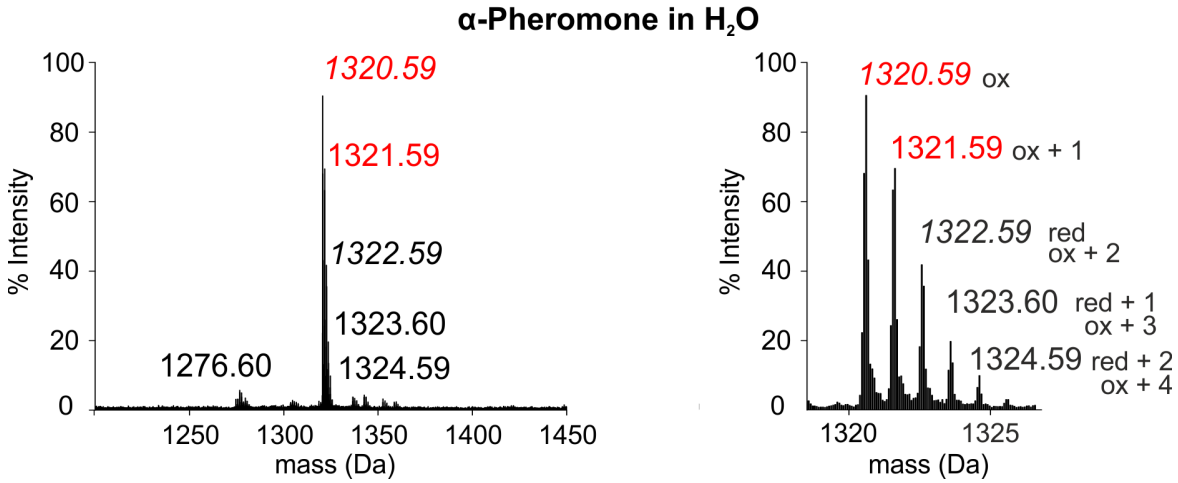


Figure 5.13. Mass spectrum of the 84 days old α-pheromone in H_2O. Selected region (left) showing the different redox states and the degradation of the pheromone, and zoomed region (right) showing the peaks corresponding to the reduced and oxidized species. Monoisotopic masses are depicted in italics, peaks corresponding to the oxidized pheromone are in red, and peaks corresponding to the reduced states in black, these peaks can be the sum of different isotopic species of the oxidized and reduced forms.

of the monomeric peptide, still present in the mixture. The second one (MW minus 2 mass units) should correspond to an oxidized form arising from an S-S bond formation between of Cys2 and Cys9 side chains of the α-pheromone. In the mass range studied, there was no evidence that supports the formation of aggregates through covalent bonds and the additional peaks found of lower molecular mass indicates the degradation of the aged pheromone.

5.3.4 3D structure of α-pheromone related peptides

The 3D structures of the α-pheromone and its derived peptides (*trans* or *cis* Xxx-Pro major form) in water, TFE, DPC and Gemini micelles were calculated based on the NMR data. **Table 5.5** and **Table 5.6** show the main structural statistics of the calculations for the peptides in H_2O and TFE, and in micelle media, respectively.

Regarding the α-pheromone wt in H_2O and TFE, due to oligomerization processes in aged samples, only the structure of the reduced forms could be obtained, while in DPC and Gemini micelles both reduced and oxidized forms could be calculated. In H_2O, the reduced α-pheromone wt adopted a β-turn structure spanning residues between Trp4 and Gln7 (**Figure 5.14**). The turn is very well defined with a RMSD of 0.4 Å for the backbone atoms. In contrast, the N- and C-termini (Trp1-Cys2-Thr3 and Pro8-Cys9-Trp10), particularly the side-chains, were more flexible and disordered, and failed to adopt defined conformations. Also, no side-chain interactions were found to stabilize the turn.

In TFE (**Figure 5.15**), the NOE data of the reduced α-pheromone wt are compatible with a main globular conformation. This preferred fold was a well-packed β-hairpin like structure, composed by two strands comprising residues Trp1 to Thr3 and Pro8 to Trp10 linked through a β-turn formed by residues spanning Trp4 to Gln7. The backbone and all side chains were very well defined (**Table 5.5**) and all the Φ and Ψ angles were located within the permitted regions of the Ramachandran plot. The structure is stabilized by a medium distance hydrogen bond between the backbone HN of Gln7 and the O of Trp4. Interestingly, π-π interactions between the aromatic rings of Trp1 and Trp10 also appeared to play an important role in peptide stabilization. Concerning electrostatic interactions, the positive Arg5 side chain charge (in the middle of the β-turn) was oriented towards the solvent and away from other peptide groups in water.

NOE data shown by the reduced α-pheromone in DPC was compatible with a globular conformation and, as shown in TFE cosolvent, the preferred folding is a well-packed β-turn centered in Gly6 and Gln7 residues (**Figure 5.16**). In Gemini micelles (**Figure 5.17**), the reduced α-pheromone populated a more disordered ensemble. This is probably due to the

poorer quality of the NMR data, attributable to signal broadening, resulting in poor quality of the structural information (**Table 5.6**). Somehow expected, the oxidized α-pheromone wt in both DPC and Gemini exhibited changes in the torsion of some backbone angles, resulting from the presence of the disulfide bond, which also alters the positions of the side chains, as shown in **Figure 5.11**.

As seen on **Figure 5.11**, hydrogen bonds between the O of Trp4 and the HN of Gly6 stabilize the turn of the reduced α-pheromone in DPC, and between the O of Trp4 and the HN of Gln7 as well. In contrast with the structure obtained in DPC micelles, the reduced α-pheromone in Gemini presents a more disordered conformation.

About the scrambled peptide, intramolecular oxidation processes were not found. As seen in **Figure 5.14** and **Figure 5.15**, in H_2O and TFE, the peptide was found in a major *trans* Trp-Pro form and failed to adopt a completely folded structure. In H_2O, the structure was mainly random with some tendency to form a β-turn centered at Cys5 and Cys6. Compared with the wild type peptide, the conformational ensemble of the scrambled peptide in water was less ordered (**Table 5.5**), the RMSD values were higher and the hydrophobic side chains projected towards the solvent into a higher degree. The scrambled peptide maintained the preferred conformation for the *trans*-form in DPC. In this instance, when compared with its structure in water, it also displayed a more compact and ordered conformation. In the presence of Gemini micelles, the *cis*-form was the preferred conformation, but in this case, signal broadening made impossible to obtain the structure from the NOE data.

The structures of the D-Ala analogs, which contain substitutions in important residues for the activity (see below), the D-Ala^{1,2} analog in H_2O, showed a turn centered between Trp4 and Gln7, and was more compact in TFE, even though it lacks the Trp amino acid at position 1 (**Figure 5.14**). In this case, the structure was only stabilized by the hydrogen bond of the turn between the backbone atoms of Trp4 and Gln7.

By contrast, the D-Ala^{6,7} analogs showed a strikingly different behavior. This sequence could not bend and behaved as a linear unfolded peptide, without any secondary structure in either H_2O or TFE solvents (**Figure 5.14** and **Figure 5.15**). The tendency of this analog to adopt a linear conformation is in accordance with previous studies reporting that substitution of two adjacent L-amino acids of a potentially helical peptide by the corresponding D-isomers resulted in an increased water accessibility and flexibility [326].

In summary, the central β-turn, although populated to varying degrees, was detected in all tested peptides, except for the D-Ala^{6,7} analog, where the two consecutive D-Ala residues at positions 6 and 7 prevented bending. In general, all peptides tested were more ordered in TFE and DPC media, suggesting that the cosolvent stabilizes the preformed conformations.

Importantly, these results indicate that *F. oxysporum* α-pheromone could adopt different structures, being more extended or flexible in the aqueous conditions.

Table 5.5. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for α -pheromone, scrambled, D-Ala^{1,2} and D-Ala^{6,7} in H_2O and 30 % TFE.

	H_2O				**30 % TFE**			
	α -Pheromone	**Scrambled**	**D-Ala^{1,2}**	**D-Ala^{6,7}**	**α -Pheromone**	**Scrambled**	**D-Ala^{1,2}**	**D-Ala^{6,7}**
Total	57	59	88	73	109	78	125	99
φ/ ψ Dihedral angle constraints (from chemical shifts)	4	2	14	11	11	4	9	15
Average CYANA target function value	0.01 ± 0.1	0.01 ± 0.1	0.02 ± 0.1	0.11 ± 0.1	0.02 ± 0.1	0.11 ± 0.1	0.28 ± 0.1	0.08 ± 0.1
Averaged maximum violation per structure								
Distance (Å)	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.01 ± 0.1
Dihedral angle (º)	0.01 ± 0.1	0.01 ± 0.1	0.40 ± 0.1	1.28 ± 0.1	0.16 ± 0.1	0.01 ± 0.1	1.52 ± 0.1	0.54 ± 0.1
Pairwise RMSD (Å)								
Backbone atoms	1.7 ± 0.6	2.9 ± 1.0	0.7 ± 0.3	2.0 ± 0.8	0.2 ± 0.1	2.2 ± 0.9	1.0 ± 0.4	0.8 ± 0.3
All heavy atoms	3.4 ± 0.9	4.4 ± 1.0	1.6 ± 0.5	3.2 ± 0.7	0.6 ± 0.2	4.2 ± 1.2	1.9 ± 0.5	2.2 ± 0.5
Ramachandran plot (%)								
Favorable	62.5	62.5	66.2	83.8	77.5	90.8	76.2	75
Allowed	37.5	37.5	33.8	16.2	22.5	9.2	23.8	25
Outlier	0	0	0	0	0	0	0	0

Table 5.6. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for α-pheromone and scrambled in 20 mM DPC and 2 mM Gemini.

	20 mM DPC			**2 mM Gemini**	
	α -Pheromone red	**α -Pheromone ox**	**Scrambled**	**α -Pheromone red**	**α -Pheromone ox**
Upper Limit distance restraints (from NOEs)					
Total	167	77	95	117	84
φ/ ψ Dihedral angle constraints (from chemical shifts)	12	8	11	9	12
Average CYANA target function value	0.52 ± 0.1	0.01 ± 0.1	0.18 ± 0.1	0.32 ± 0.1	0.68 ± 0.1
Averaged maximum violation per structure					
Distance (Å)	0.03 ± 0.1	0.01 ± 0.1	0.01 ± 0.1	0.02 ± 0.1	0.01 ± 0.1
Dihedral angle (º)	1.06 ± 0.1	0.25 ± 0.1	1.28 ± 0.1	0.48 ± 0.3	1.76 ± 0.2
Pairwise RMSD (Å)					
Backbone atoms	0.0 ± 0.1	0.3 ± 0.2	0.4 ± 0.3	0.6 ± 0.2	0.4 ± 0.2
All heavy atoms	0.6 ± 0.2	1.2 ± 0.5	1.0 ± 0.3	1.9 ± 0.6	1.5 ± 0.6
Ramachandran plot (%)					
Favorable	75	84.4	45.6	75	84.4
Allowed	16.9	15	54.4	20	15.6
Outlier	8.1	0.6	0	5	0

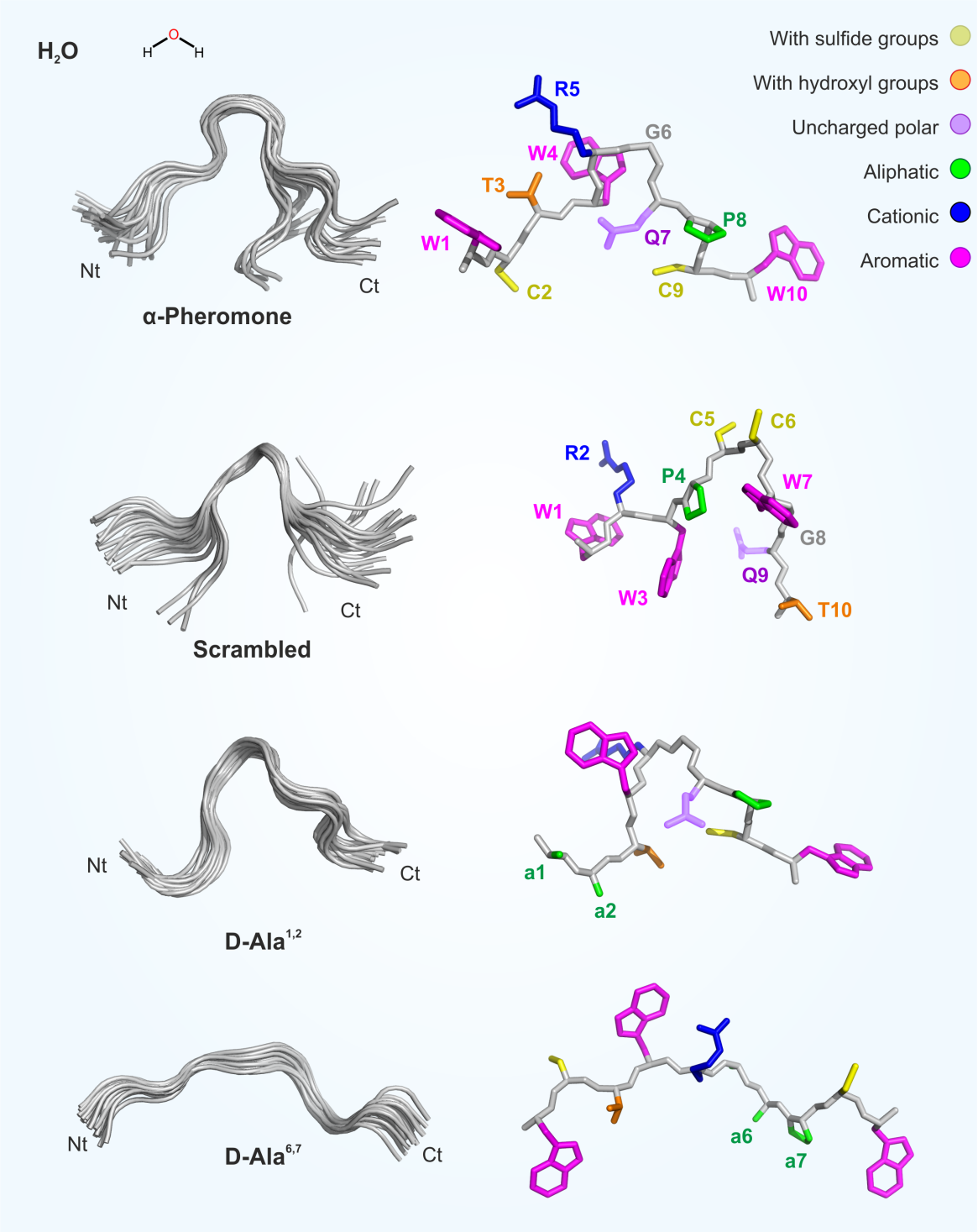


Figure 5.14. Ensemble of the 20 structures with the lowest target function values of α-pheromone, scrambled, D-Ala^{1,2} and D-Ala^{6,7} in H_2O, 5 °C, pH 5.0. The superposition of the backbone of the 20 lowest target function structures in each family is represented as cartoon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

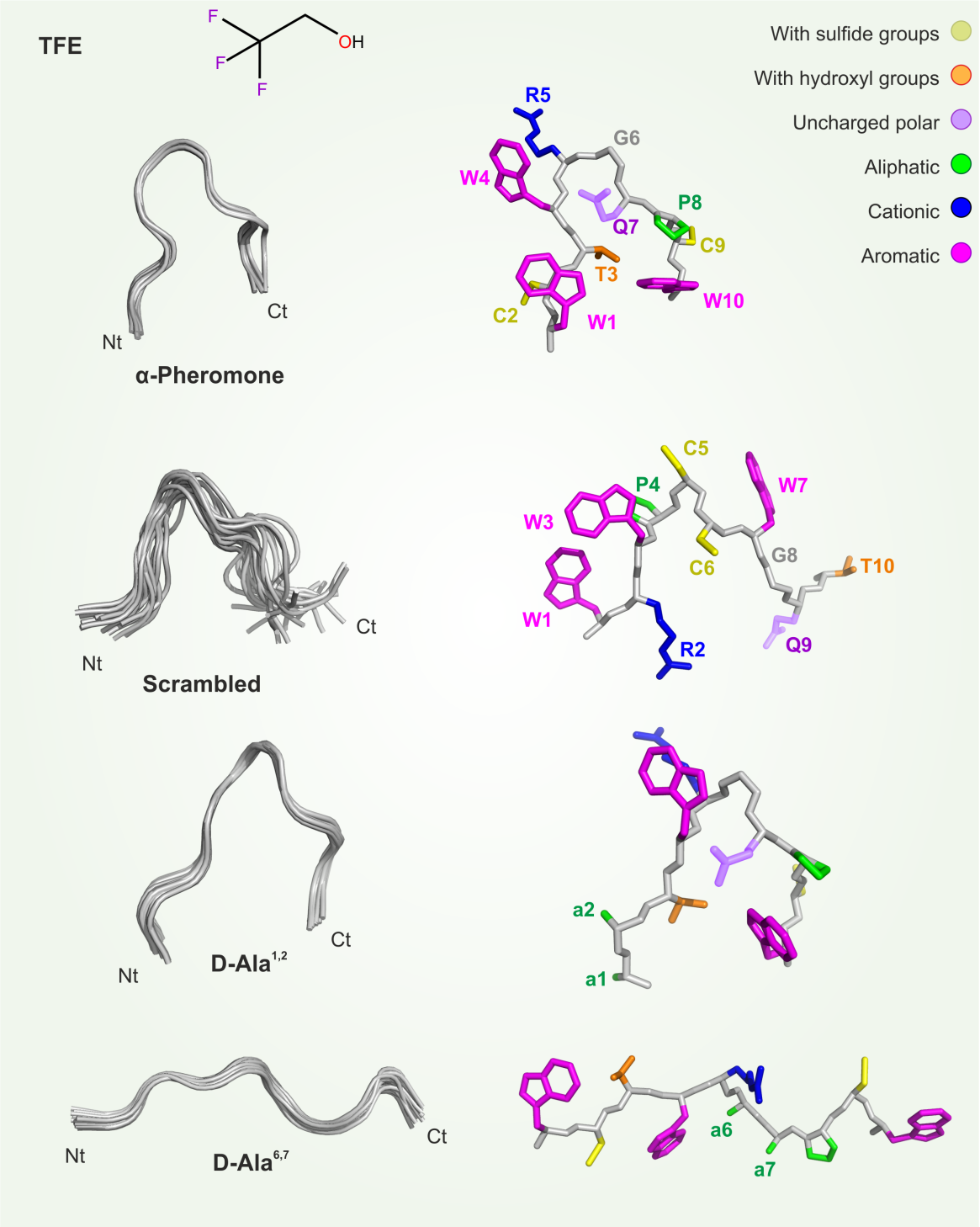


Figure 5.15. Ensemble of the 20 structures with the lowest target function values of α-pheromone, scrambled, D-Ala^{1,2} and D-Ala^{6,7} in 30 % TFE, 5 °C, pH 5.0. The superposition of the backbone of the 20 lowest target function structures in each family is represented as cartoon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

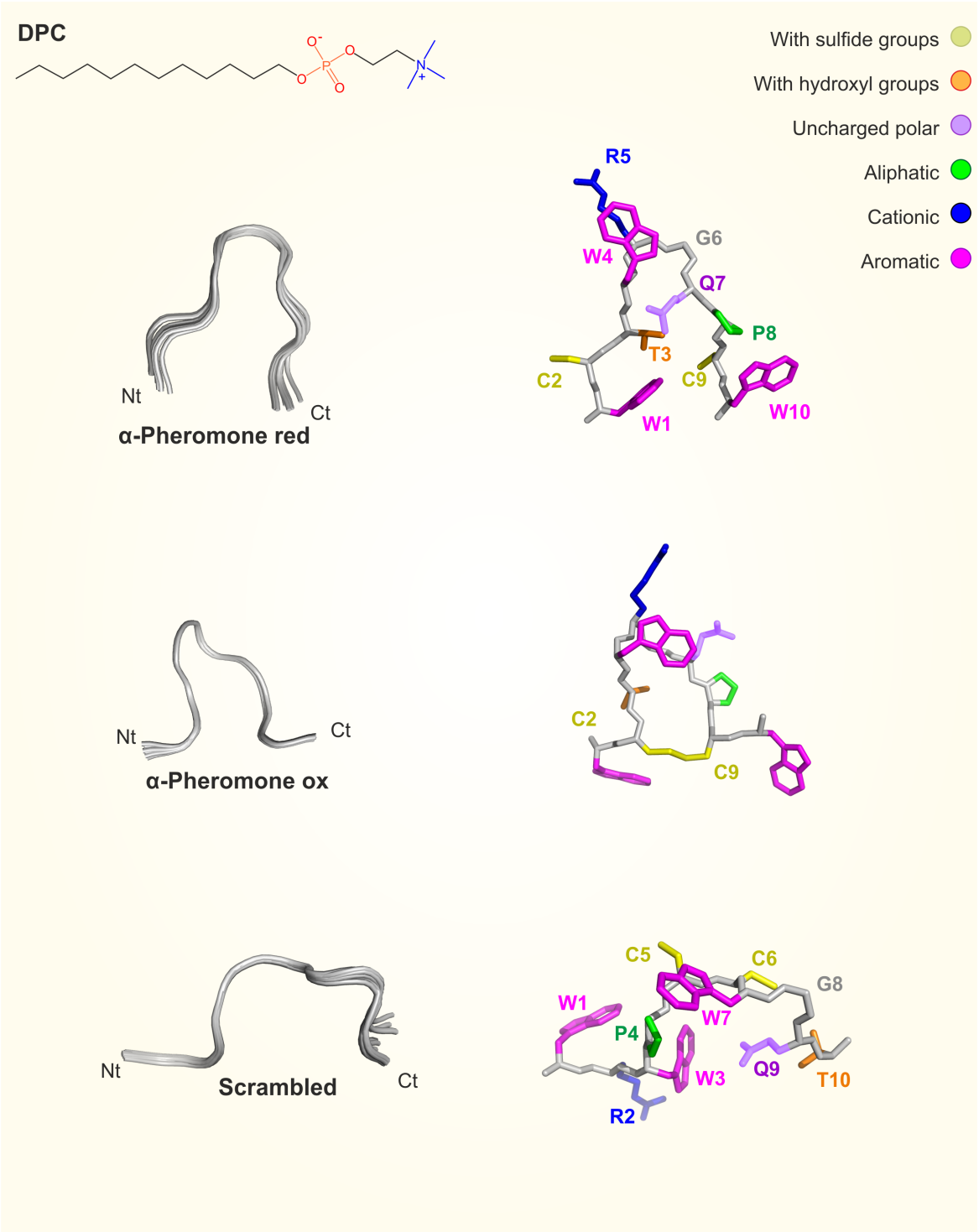


Figure 5.16. Ensemble of the 20 structures with the lowest target function values of α-pheromone and scrambled in 20 mM DPC, 25 °C, pH 5.0. The superposition of the backbone of the 20 lowest target function structures in each family is represented as cartoon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

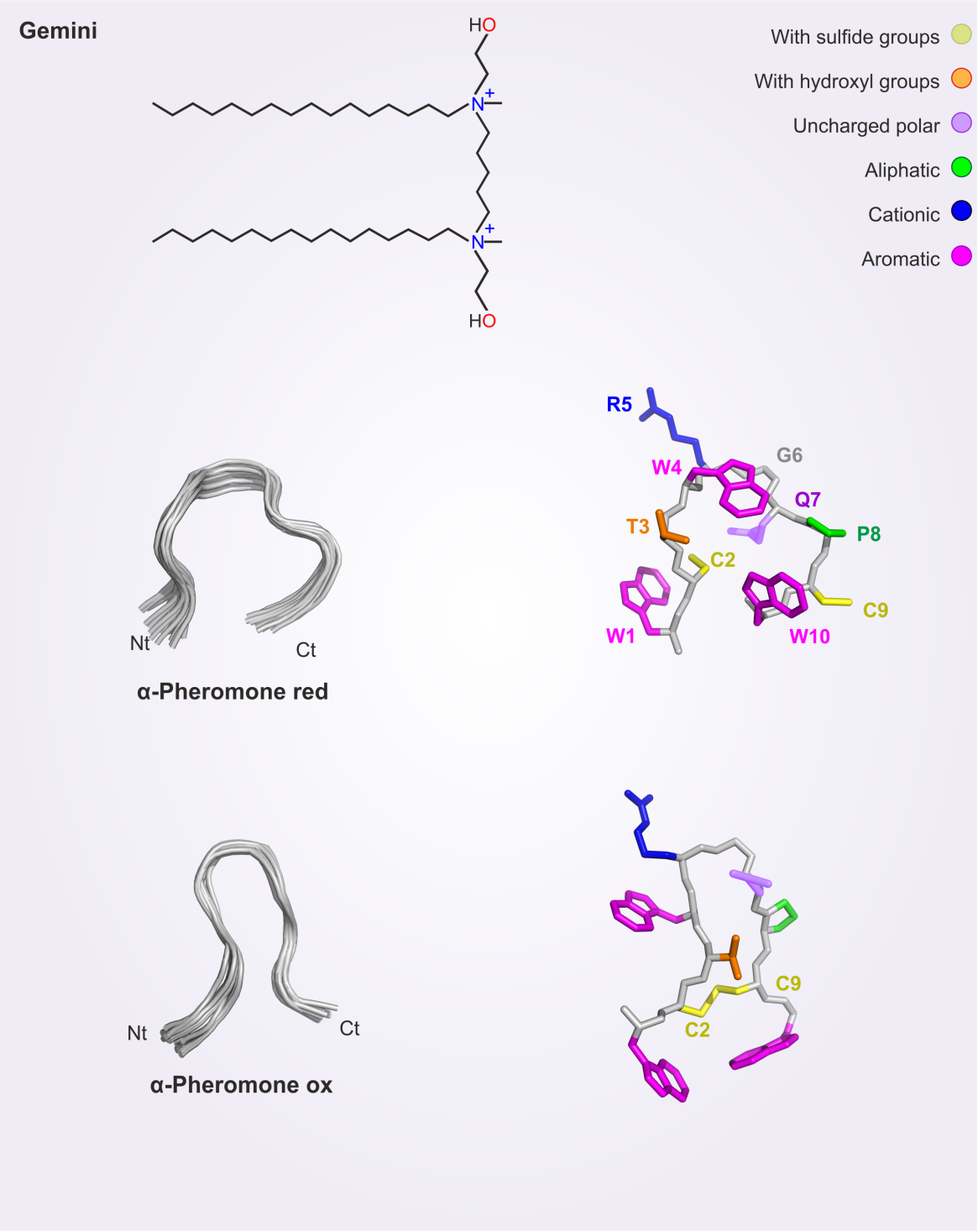


Figure 5.17. Ensemble of the 20 structures with the lowest target function values of α-pheromone and scrambled in 2 mM Gemini, 25 °C, pH 5.0. The superposition of the backbone of the 20 lowest target function structures in each family is represented as cartoon. Side chains of the lowest target function structure in solution are depicted in different colors depending on the type of amino acid.
Cationic: blue, uncharged polar: violet, aromatic: magenta, aliphatic: green, with sulfide groups: yellow and hydroxyl groups: orange.

5.3.5 Peptide interaction with micelles

To test if possible intermolecular interactions with micelles take place, 2D ^{1}H-^{1}H-NOESY spectra of the wt α-pheromone and Scr peptides (**Figure 5.18**) were analyzed, and the identified intermolecular NOEs were evaluated and scaled with respect to a known distance as explained in the experimental section (**Figure 5.19**).

Intermolecular NOEs proved the establishment of direct interactions between both peptides and micelles. In the wild type, the indole rings corresponding to Trp1, Trp4 and Trp10 were crucial for the interaction with the hydrophobic moiety of the DPC and Gemini micelles. However, the complex appears as dynamic since their contacts go from the micelle core (B in DPC and Gemini) up to the external protons (G in DPC and H in Gemini) as seen in **Figure 5.19**. Generally, it also seems that Trp10 took a major role within the interaction when the peptide became oxidized. Apart from the aromatic groups, Gln7 and Arg5 also establish contacts with the micelle and for Arg5, a snorkel model-type interactions common for Arg and Lys side chains [327], seems feasible.

In the scrambled peptide, the interaction is also preferably with the hydrophilic regions of the DPC and Gemini micelles through the aromatic rings of Trp1, Trp3 and Trp7. In this case, also a different contribution of these rings can be seen depending on the media, and Trp7 and Trp1 became more relevant in the interaction with Gemini.

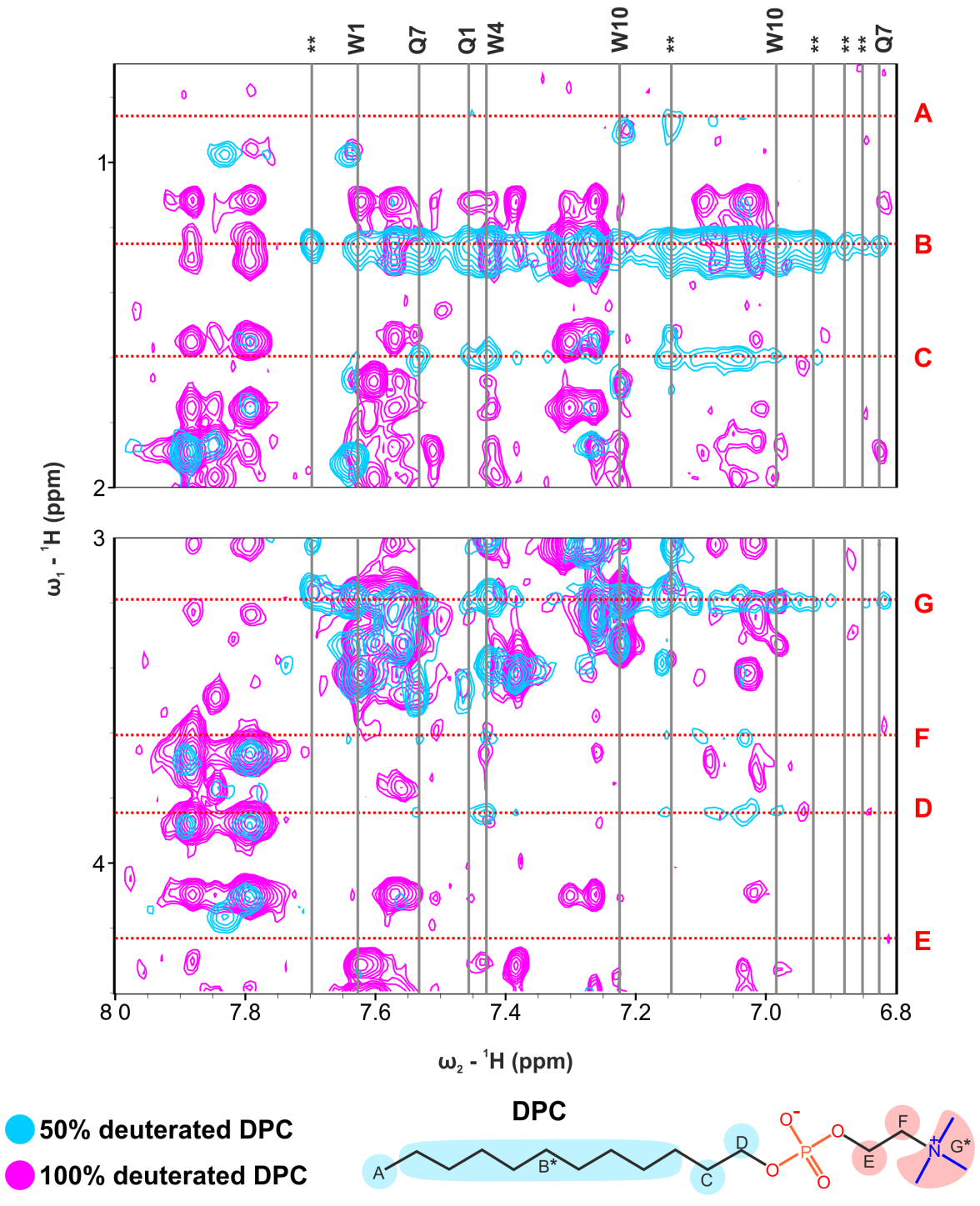


Figure 5.18. Selected regions of the 2D ^{1}H-^{1}H NOESY spectra of reduced α-pheromone in 20 mM DPC/DPC-d38 1:1, 35 ºC, pH 5.0. DPC chemical shifts are indicated with dashed lines. Intermolecular NOE signals correspond with the blue peaks in the partially deuterated DPC media that are not present in the fully deuterated DPC spectra. Differences on the signals are also due to different oxidation ratios in the timescale. ** Ambiguous assignments from the minority *cis* species. * Protons with the same chemical shift.

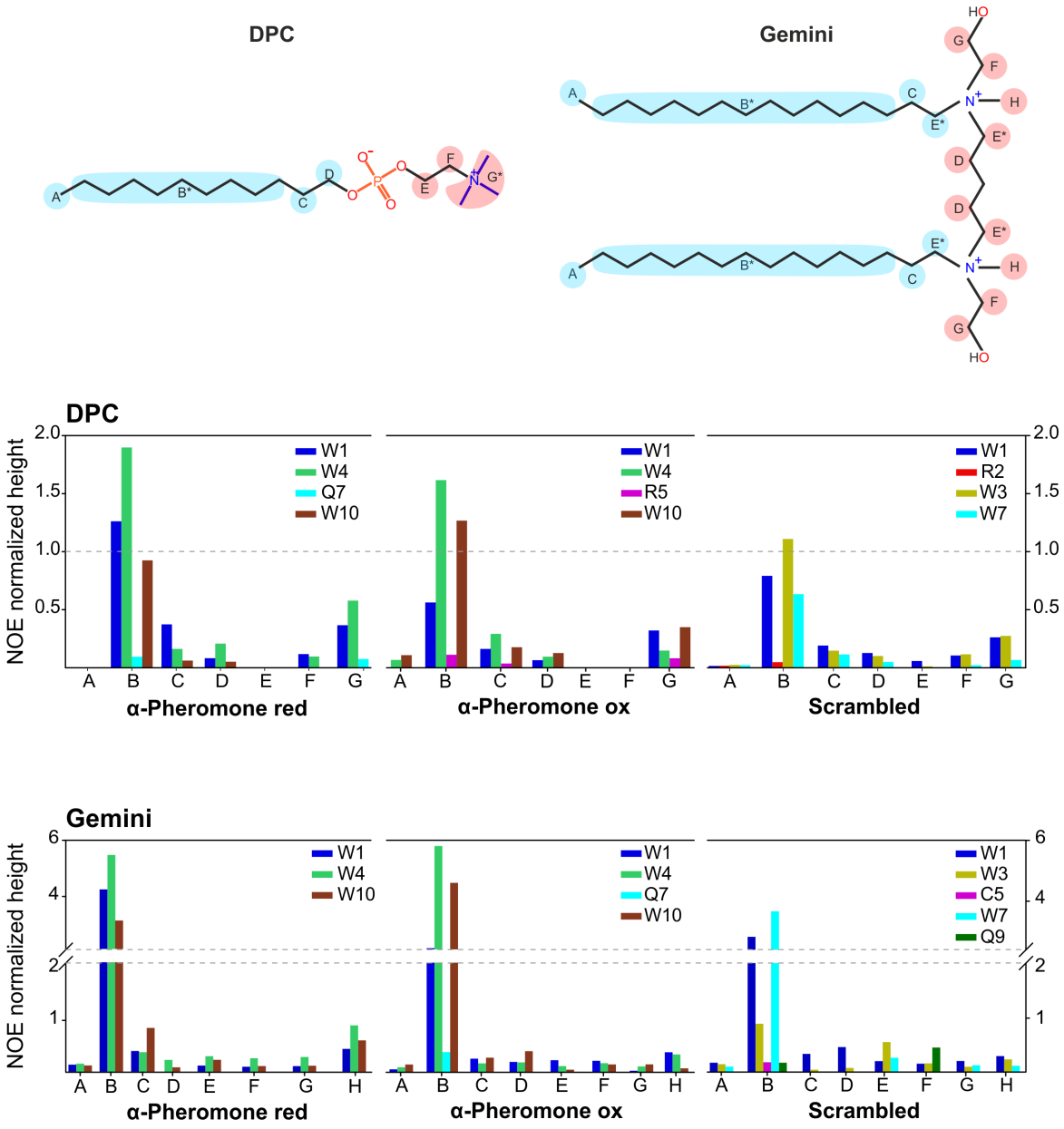


Figure 5.19. Barplot representing the scaled intermolecular NOEs for each residue found in reduced and oxidized α-pheromone and scrambled peptides. 20 mM DPC (upper panel), and 2 mM Gemini (lower panel) were selected as membrane mimetics.
* Protons with the same chemical shift.

5.4 Discussion

Previous to this study, a 13 residue long peptide from *Saccharomyces cerevisiae* has been used as paradigm in the study of the interactions between small peptides and their cognate G protein-coupled receptors (GPCR) [328, 329, 312]. Studies by NMR of the α-pheromone from *S. cerevisiae* reveals that the residues close to the N-terminus might be important for the receptor activation, while the ones at the C-terminus should be involved in the receptor binding [312] (**Figure 5.4**). These studies confirm that not only large proteins are important for biological roles and that even short sequences can depict different and localized functions. However and regarding this kind of systems applied for other types of filamentous fungi, which share no sequence similarity with its homolog, there was no structural information defined at atomic resolution [330, 270], but they still share comparable functions.

In this study, in collaboration with the groups of Prof. Di Pietro from the University of Cordoba and Prof. Martínez del Pozo from the Complutense University of Madrid, a complete analysis of the structure and function relationship was done for the α mating pheromone of *Fusarium oxysporum*. The importance of this study arises from the role of this peptide in the activation of its cognate GPCR Ste2, which has an essential implication during its chemotropic growth and therefore, throughout plant infection [267, 270]. This process, as mentioned in the introduction section, has a high economic cost worldwide.

5.4.1 *Fusarium* α-pheromone's 3D structure

Despite their difference in length and sequence, *S. cerevisiae* and *F. oxysporum* α-pheromones share a number of common features. Both are cationic peptides with a pI around 8. Both have, at the center of the ordered β-turn, a charged residue (Arg in *F. oxysporum* and Lys in *S. cerevisiae*) whose side chains point towards the solvent (as directly seen for *Fusarium* α-pheromone, and intuited for the *Saccharomyces*'s pheromone), suggesting that these residues may act as a molecular antenna, playing a key role in regulating and driving potential intermolecular interactions. An interesting feature that differentiates *F. oxysporum* α-pheromone from its orthologous in *S. cerevisiae* is the presence of two Cys residues, both of which were found to be reduced in the fresh samples *in vitro*. Although it is currently unknown whether this fact reflects the biological context of the pheromone-receptor interaction, together with the high reactivity and known biological functions of the thiol groups, this opens the intriguing hypothesis that *F. oxysporum* α-pheromone may have also an additional unreported function in redox regulated processes, while *S. cerevisiae* pheromone

would not due to its lack of Cys residues. Cysteines can easily function as nucleophiles, as they are also key elements in the catalytic triad for cysteine proteases [331, 332] and thus could form covalent adducts with different molecules such as lipids or ADP [333, 334], a hypothesis which might be of interest for future investigations.

Regarding the relevance of the cell membrane towards the regulation of these processes, we have compared the reduced structures in H_2O and H_2O/TFE [335] and those in micelles. We saw that DPC and Gemini micelles favor the establishment of different Trp-Gln interactions. In the Scr structure, Gln9 and Trp3 side chains have a tendency to stay close in space with distances below than 5 Å in most conformers of the structural ensemble. Also, Gln7 in the wt pheromone is closer in space with Trp1, Trp4 and Trp10 in micelles. The difference is that in this last case, very short distances (about 3.0 - 3.5 Å) are measured for some of these contacts between Trp-Gln side chains. Assuming that the structure adopted by these peptides in micelles also represents the structure in vesicles, these short Trp-Gln contacts can explain the important decrease of the Trp fluorescence emission observed in the wt, by our colleagues (A. Martínez del Pozo, personal communication).

Interestingly, we have found that a non-enzymatic oxidation process occurs in a time dependent manner *in vitro*. According with MS and NMR results, an intramolecular and not intermolecular disulfide bond between Cys2 and Cys9 gets formed in the α-pheromone in both water and micelle media. In pure water, this process leads to a massive oligomerization of the peptide, while the intramolecular bond formation was a slower process in micelles. However, and very interestingly, aggregates were not then detected in micelles, suggesting that the micellar structure protected the monomer scaffold, which probably might be the biological active form. From a structural/mechanistic point of view, this could be explained by a conformational change in the peptide due to the interaction with the micelle. In this situation, the two SH groups, that were apart in the free state, can be accommodated in the correct arrangement for the intramolecular redox reaction to take place. From a biological point of view, it can be inferred the existence of an enzyme (probably near the membrane) that might oxidize the pheromone. On the other hand, and suggesting the biological significance of these results, the Scr peptide, with two consecutive Cys residues (positions 5 and 6), was not oxidized and neither formed aggregates, independently of the presence, or not, of surfactant micelles. On the basis of these results, it is tempting to propose that the existence of an equilibrium between the oxidized and reduced forms of the α-pheromone *in vivo* might lead to different functions for this short peptide, in one redox state, the pheromone might be interacting more deeply with the membrane and the other one participates in the activation of its receptor.

In this regard, a computer-based model for *S. cerevisiae* α-factor bound to Ste2 was previously proposed based on biochemical and biophysical data [336]. The model places the α-pheromone bend around the Pro-Gly center of the peptide, with the Lys side chain facing away from the transmembrane domains and interacting with a binding pocket formed by the extracellular loops of the receptor. For *F. oxysporum*, our structural data are compatible with this type of model based on a central β-turn and a Arg cationic side chain located in similar positions.

5.4.2 *Fusarium* α-pheromone contains discrete structural and functional regions

Previous studies revealed that α-pheromone elicits a robust chemotropic response in germ tubes of *F. oxysporum*, which is dependent on the cognate GPCR Ste2 [267, 270]. Results obtained by our collaborators in Cordoba clearly show that in order to exert its chemotropical properties, the β-turn between Gly6 and Gln7 must be present [335]. Thus, the performed biological assays revealed a dramatic loss of the chemotropical activity for D-Ala^{6,7} and the scrambled sequence, the two variants without the turn in that position. This means that Gly6 and Gln7 residues are essential for the chemoattractant capacity of the pheromone. This is in line with the structural role of Gly6 and Gln7 in the maintenance of the 3D-structure, and strongly suggests that activation of the Ste2-mediated chemotropic growth depends on the secondary β-turn structure of α-pheromone rather than on its amino acid composition or pI. The idea is further supported by the finding that the N-terminus Trp1 and Cys2 residues which play no substantial role in the structure of α-pheromone, are not required for chemoattraction. Indeed, Ala substitutions at the N-terminus of the α-pheromone in *S. cerevisiae* and *C. albicans* led to increased pheromone activity, suggesting that this region could have an inhibitory function role in receptor-mediated signalling [311, 312].

As the *S. cerevisiae*'s α-pheromone can induce transient cell cycle arrest, an analogous biological test was made with *Fusarium* α-pheromone variants by the group of Prof. Di Pietro [335]. The results indicated that, similar to *S. cerevisiae* [311, 312], α-pheromone inhibits cell division in *F. oxysporum*. Interestingly, the growth inhibitory activity was abolished in the D-Ala^{1,2} and the scrambled analogs both of which lack a Trp residue at the N- or the C-termius, respectively. Aromatic residues, particularly Trp, have been reported to undergo specific interactions with lipid moieties, resulting in anchoring to the membrane in a *velcro-like* manner [133, 337]. The presence of the Trp residues Trp1 and Trp10 at the

two termini suggests that they could act as a structural clamp during a potential membrane interaction linked to the growth inhibitory activity of the α-pheromone.

5.4.3 Additional insights into the pheromone's mechanisms of action, membrane interaction

Fluorescence and isothermal titration calorimetry (ITC) experiments performed by Prof. Martínez del Pozo reveal that the α-pheromone and the scrambled peptides were only capable to interact with negatively charged PG vesicles in a weak and superficial way, suggesting the importance of the positive charged groups of the Arg5 and the N-terminus of the peptides and their implications in the formation of electrostatic interactions. However, the interaction seemed to be stronger for the scrambled peptide. This is in agreement with their structures indicating the importance of the distribution of the positive charged and hydrophobic groups in these peptides. In the Scr, positive charges and hydrophobic Trp residues are concentrated in the N-terminus region contributing to strengthen the interaction. Taking these results together with the already three-dimensional structures reported for both peptides in the presence of membrane-like environments, it was tempting to speculate that both of them would adopt a very different arrangement when interacting with the phospholipid vesicles.

These results were in accordance with the NMR analysis, although surfactants were used as membrane mimetics. The large proportion of Trp residues (3 out of 10 residues) suggested how they are probably key players in the interaction of the peptides with lipid membranes. As seen in **Figure 5.19**, intermolecular contacts were indeed more specific for α-pheromone than for the scrambled. Differences in the local position of key amino acids such as Trp and the only Arg and Gln residues are probably responsible for the different tendencies to participate in the interaction with micelles and its specificity and since the scrambled analog is inactive, membrane interactions does not suffice to the peptide being functionally active.

5.4.4 Model of the α-pheromone-membrane interaction

In order to improve our understanding of the still unclear mechanisms of *F. oxysporum* α-pheromone, and based on the calculated structures, we have integrated all the results discussed in this study to propose a molecular model for the membrane interaction process (**Figure 5.20**). In this model, the α-pheromone wt should lay parallel or slightly tilted on the membrane surface. In this disposition, Trp1 and Trp10 rings (N- and C-termini) can interact with the hydrophobic region of the lipids, and positive charges at the N-terminus and in the

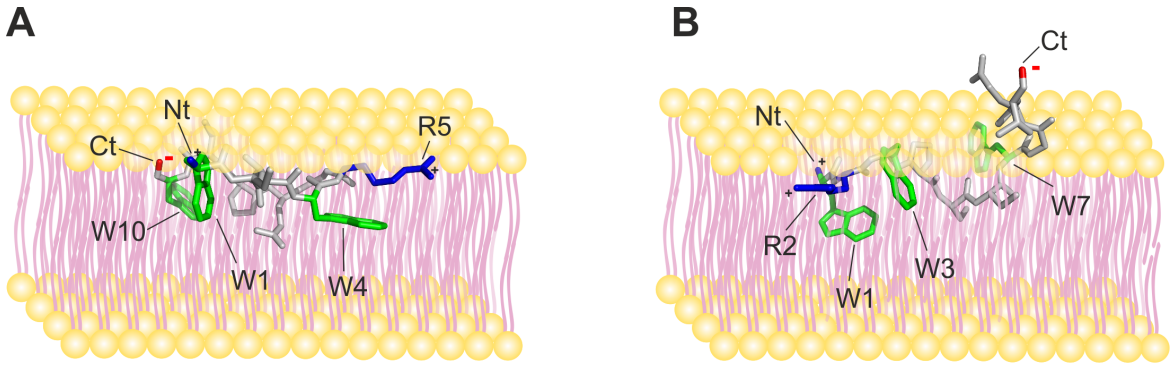


Figure 5.20. Model for the interaction between the wt α-pheromone (A) and the scrambled peptide (B) with membranes. The 3D structures of these peptides were those obtained by NMR in this study. Important residues for the interaction are labeled; Trp rings are in green, positive charges in blue and negative charges in red. Peptide backbone is in gray.

center of the β-turn (Arg5) could help in weakly anchoring the structure by both ends to the phosphate groups. This agrees with the observation that *in vitro*, the interaction with lipids occurs and it is of dynamic nature, weak and superficial. The located charge and hydrophobic amino acid distribution seems to be key players for the interaction.

The NMR-based model is also in agreement with the ITC results, which indicate that different quantity of lipid molecules are affected by the interaction, being larger in the α-pheromone than in the scrambled peptide. Thus, while this model account for the interaction of the full α-pheromone with the membrane, the accumulation of Trps and positive charges close to the N-terminus of the scrambled sequence suggests that mostly the N-terminus, and not the full peptide, is interacting with the micelle, leaving half of the sequence pointing out of the membrane (**Figure 5.20**). As a consequence, the number of lipid molecules affected by the interaction is smaller but stronger for the scrambled variant than that observed for the α-pheromone, as seen in the ITC experiments.

Once the initial contacts of the α-pheromone are formed, they can be strengthened by means of specific links that help to reorient the molecule for optimal interactions with the membrane and promote conformational changes for oxidation processes. This oxidization, which may be produced by an enzyme, is able to maintain the association. In this regard, we can advance that the membrane would have a key role favoring the active conformation of the mating pheromone and probably, modulating the activity via redox processes. Finally, another plausible role of the membrane would be to promote the sequestration of α-pheromone molecules within the cell surface, increasing the local peptide concentration and preventing the formation of large aggregates to facilitate the recognition by the membrane Ste2 GPCR, and therefore, making the recognition process more efficient.

Summarizing, our results establish that *F. oxysporum* α-pheromone adopts a defined secondary structure and, despite its short sequence, contains discrete regions involved in different biological processes, such as polarity reorientation and cell division. On the basis of these results, it is also tempting to propose that the existence of an equilibrium between the oxidized and reduced forms might lead to different functions or degree of activity for this short peptide, one of them interacting more deeply with the membrane and the other one in the activation of its receptor. We consider that even when further biological assays are needed to understand the real biological implication of the oxidized form of the α-pheromone, we contemplate that the current findings might be applied to additional α-pheromones from other ascomycetes.

5.5 Conclusions

The ascomycete fungus *Fusarium oxysporum* is a highly destructive plant pathogen and an emerging pathogen of humans. Given that *Fusarium* uses its sex pheromone receptor for chemotropic sensing and invasiveness, information regarding its structure under different biological conditions is essential to understand the mechanisms underlying these processes. According to the results described above, the main conclusions of this study are:

1. It has been proven by NMR and biological assays that despite its short sequence, the α-pheromone can adopt a defined structure and possess discrete regions associated with determined biological functions.

2. The 3D structure obtained from NMR data of the α-pheromone reveals that Gly6 and Gln7 are key in the maintenance of the structure, specifically the establishment of the central β-turn. Biological assays done by our collaborators showed that the presence of this β-turn is essential for chemoattractant activity and therefore these residues play a role in the activation of its receptor Ste2. On the other side, Trp1 and Cys2 amino acids located at the N-terminus are not relevant for structure, but essential for the regulatory processes of the α-pheromone.

3. The two Cys residues close to the N- and C-termini experimented a non-enzymatic oxidation of their SH groups, *in vitro*, in a time dependent manner. This oxidation leads to an intra- or intermolecular disulfide bond depending on the media. In water, a massive oligomerization of the peptide was found; while in micelles, aggregates were not detected and only the intramolecular S-S bond species could be characterized. This suggests that the structural change, which occurs in micelles, protects the monomer scaffold. On this basis, we propose that the existence of an equilibrium between the oxidized and reduced forms *in vivo* might lead to different functions for the α-pheromone, one state interacting more deeply with the membrane and the other one in the activation of its receptor.

4. The analysis at atomic level of the interaction of the α-pheromone with micelles demonstrates that such interaction is different in the reduced and oxidized forms, and also in the wt and scrambled peptides. Differences in the local position of key amino acids, such as Trp and the only Arg and Gln residues are probably responsible for the different tendencies to participate in the interaction and its specificity.

5. All experimental results are coherent and a model of the possible interaction of the α-pheromone with the lipid bilayer has been proposed. In this model, Trp1 and Trp10 rings can interact with the hydrophobic tails of the lipids, while the positive charges of Arg5 at the center of the turn and the N-terminus might help to anchor the structure towards the phosphate groups.

6

General conclusions

6.1 General summary and general conclusions

The present work was carried out in the group of Protein Structure, Dynamics and Interactions by NMR, belonging to the Department of Biological Physical Chemistry from the Institute of Physical Chemistry Rocasolano (CSIC), from September, 2014 to September, 2018.

Each chapter contains the results of the systems subjected to study in this Thesis and gathers the specific conclusions derived from their evaluation and discussion. Here, we include a compilation of the general concluding remarks derived from the whole work and the specific systems analyzed.

1. The results presented here corroborate that NMR spectroscopy is a potent tool that provides precise and essential information about the structure and dynamic properties involved in protein-protein and protein-membrane interactions at atomic resolution. To properly collect and characterize this kind of information is fundamental to achieve a deep understanding of the biochemical pathways that rule the biological processes of living organisms.

2. For this Thesis, we selected large and complex systems involving proteins and membranes that are difficult to be characterized at atomic detail. Our minimalist approach, in which once identified the main components of the complexes, simplified mimetic versions were designed for its analysis under a variety of biologically relevant conditions. This tactic has been proven to be very useful and can be extended to other complex systems.

3. Like in precedent studies, we confirm that the use of one single technique, even such a powerful one as NMR spectroscopy, might not suffice for the full and precise characterization of a biological system. The use of complementary experimental and computational approaches is a more successful strategy to get reliable results and put them into a wider context.

4. The careful design and selection of a simplified model has shown to be useful to gain valid results in order to be later extrapolated into a biological context. However, in agreement with previous studies, our data suggest that a caveat to have in mind when applying this excessively reductionist approach is that it may lack information from other components, which results in a loss of significant data.

5. In all cases, and following the strategies mentioned before, molecular models based on experimental results had been proposed. These help to better understand the recognition

processes that are necessary for the specific biological functionality of the different systems.

Following these global remarks, here we enumerate the main specific conclusions regarding each system studied.

- In chapter 2, in which we have studied the XTACC3-XMAP215 protein-protein interaction, it was demonstrated that the conservation of the overall structure of the different peptide variants does not necessarily imply the preservation of the interaction of the full length proteins. In the same way, it was proven that their overall structure population and interactions depend on the pH and the presence of cosolvents. We finally propose a model that can describe the initial steps of the recognition event (1:1 molar interaction) in the formation of the fully active XTACC3-XMAP215 complex (1:2).

- In chapter 3, dedicated to the glycoprotein gp41 of HIV-1, it was shown that the minimization of the full protein employing peptides is a good approach to study the interaction with membranes. Indeed, our results are in agreement with other biophysical data for peptides spanning wider regions. On the bases of our results, we propose a model of the role played by gp41 during viral infection and antibody neutralization.

- Studies on the antimicrobial peptide protegrin-1 were described in chapter 4. It was established that the presence of cosolvents (TFE and DPC micelles) stabilized the peptide-peptide interactions, at least long enough to be detected through NMR. We propose that the self-association process previous or concomitant with the membrane interaction is important for the first steps of membrane disruption and antimicrobial activity.

- In chapter 5, we studied the pheromone of the highly destructive fungus *Fusarium oxysporum*. The use of Gemini surfactants as good membrane models validates that the interactions found between the peptides and the surfactants are with the micelle (membrane mimetic) and not with the monomeric form of the detergent. All the experimental results are combined and fitted into a model, which describes the possible interaction of the α-pheromone with the lipid bilayer and proposes different biological roles of the oxidized and reduced states of the pheromone.

6.2 Publications derived from this Thesis

The obtained results are gathered in three published articles and three more are in preparation to be published soon:

- E. Rujas, J. M. Caaveiro, A. Partida-Hanon, N. Gulzar, K. Morante, B. Apellaniz, M. Garcia-Porras, M. Bruix, K. Tsumoto, J. K. Scott, M. A. Jimenez, and J. L. Nieva. "Structural basis for broad neutralization of HIV-1 through the molecular recognition of 10E8 helical epitope at the membrane interface." *Sci Rep*, 6:38177, (2016). http://dx.doi.org/10.1038/srep38177

- S. Vitale, A. Partida-Hanon, S. Serrano, A. Martinez-Del-Pozo, A. Di Pietro, D. Turra, and M. Bruix. "Structure-Activity Relationship of alpha Mating Pheromone from the Fungal Pathogen *Fusarium oxysporum*." *J Biol Chem*, 292(9):3591–3602, (2017). http://dx.doi.org/10.1074/jbc.M116.766311

- A. Partida-Hanon, M. A. Treviño, M. Mompeán, M. A. Jiménez and M. Bruix. "Structural insight into the XTACC3/XMAP215 interaction from CD and NMR studies on model peptides." *Biopolymers* 107: e23039, (2017). http://dx.doi.org/10.1002/bip.23039

- A. Partida-Hanon, M. Maestro-López, S. Vitale, D. Turrà, A. Martínez-Del-Pozo, A. Di Pietro and M. Bruix. "Study of the α-Mating Pheromone from *Fusarium oxysporum* in membrane mimetics: a molecular model for membrane recognition." *In preparation* (2018).

- A. Partida-Hanon, M. Bruix and M. A. Jiménez. "The role of disulphide bonds in peptides derived from the antimicrobial protegrin-1 for molecular association and membrane disruption." *In preparation* (2018).

- A. Partida-Hanon, E. Rujas, M. Bruix, J. L. Nieva, and M. A. Jiménez. "Structural bends of the gp41 glycoprotein of HIV-1 might contribute to the membrane fusion mechanism during infection." *In preparation* (2018).

Bibliography

[1] E. Fischer. "Einfluss der configuration auf die wirkung der enzyme". *Ber Dtsch Chem Ges*, 27(3):2985–2993, (1894).

[2] D. E. Koshland. "Application of a theory of enzyme specificity to protein synthesis". *Proc Natl Acad Sci U S A*, 44(2):98–104, (1958).

[3] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes. "The energy landscapes and motions of proteins". *Science*, 254(5038):1598–603, (1991).

[4] K. Ye, X. Tu, X. Zhang, Q. Shang, S. Liao, J. Yu, and J. Zhang. "Induced folding under membrane mimetic and acidic conditions implies undiscovered biological roles of prokaryotic ubiquitin-like protein pup". *Protein Pept Lett*, 23(8):756–62, (2016).

[5] D. D. Boehr, R. Nussinov, and P. E. Wright. "The role of dynamic conformational ensembles in biomolecular recognition". *Nat Chem Biol*, 5(11):789–96, (2009) http://dx.doi.org/10.1038/nchembio.232.

[6] G. M. Clore, C. Tang, and J. Iwahara. "Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement". *Curr Opin Struct Biol*, 17(5):603–16, (2007) http://dx.doi.org/10.1016/j.sbi.2007.08.013.

[7] R. van der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright, and M. M. Babu. "Classification of intrinsically disordered regions and proteins". *Chem Rev*, 114(13):6589–631, (2014) http://dx.doi.org/10.1021/cr400525m.

[8] S. J. Singer and G. L. Nicolson. "The fluid mosaic model of the structure of cell membranes". *Science*, 175(4023):720–31, (1972).

[9] G. Vereb, J. Szollosi, J. Matko, P. Nagy, T. Farkas, L. Vigh, L. Matyus, T. A. Waldmann, and S. Damjanovich. "Dynamic, yet structured: The cell membrane three decades after the Singer-Nicolson model". *Proc Natl Acad Sci U S A*, 100(14):8053–8, (2003) http://dx.doi.org/10.1073/pnas.1332550100.

[10] P. Verkade and K. Simons. "Robert Feulgen Lecture 1997. Lipid microdomains and membrane trafficking in mammalian cells". *Histochem Cell Biol*, 108(3):211–20, (1997).

[11] L. Fagerberg, K. Jonasson, G. von Heijne, M. Uhlen, and L. Berglund. "Prediction of the human membrane proteome". *Proteomics*, 10(6):1141–9, (2010) http://dx.doi.org/10.1002/pmic.200900258.

[12] L. Kall, A. Krogh, and E. L. Sonnhammer. "A combined transmembrane topology and signal peptide prediction method". *J Mol Biol*, 338(5):1027–36, (2004) http://dx.doi.org/10.1016/j.jmb.2004.03.016.

[13] J. Monod, J. Wyman, and J. P. Changeux. "On the nature of allosteric transitions: a plausible model". *J Mol Biol*, 12:88–118, (1965).

[14] I. M. Nooren and J. M. Thornton. "Diversity of protein-protein interactions". *EMBO J*, 22(14):3486–92, (2003) http://dx.doi.org/10.1093/emboj/cdg359.

[15] I. M. Nooren and J. M. Thornton. "Structural characterisation and functional significance of transient protein-protein interactions". *J Mol Biol*, 325(5):991–1018, (2003).

[16] J. R. Perkins, I. Diboun, B. H. Dessailly, J. G. Lees, and C. Orengo. "Transient protein-protein interactions: structural, functional, and network properties". *Structure*, 18(10):1233–43, (2010) http://dx.doi.org/10.1016/j.str.2010.08.007.

[17] G. von Heijne. "The membrane protein universe: what's out there and why bother?". *J Intern Med*, 261(6):543–57, (2007) http://dx.doi.org/10.1111/j.1365-2796.2007.01792.x.

[18] W. Cho and R. V. Stahelin. "Membrane-protein interactions in cell signaling and membrane trafficking". *Annu Rev Biophys Biomol Struct*, 34:119–51, (2005) http://dx.doi.org/10.1146/annurev.biophys.33.110502.133337.

[19] N. O. Baez, J. A. Reisz, and C. M. Furdui. "Mass spectrometry in studies of protein thiol chemistry and signaling: opportunities and caveats". *Free Radic Biol Med*, 80:191–211, (2015) http://dx.doi.org/10.1016/j.freeradbiomed.2014.09.016.

[20] R. Aebersold and M. Mann. "Mass spectrometry-based proteomics". *Nature*, 422(6928):198–207, (2003) http://dx.doi.org/10.1038/nature01511.

[21] A. H. Smits and M. Vermeulen. "Characterizing protein-protein interactions using mass spectrometry: challenges and opportunities". *Trends Biotechnol*, 34(10):825–834, (2016) http://dx.doi.org/10.1016/j.tibtech.2016.02.014.

[22] S. E. Harding and K. Jumel. "Light scattering. ". *Curr Protoc Protein Sci*, 11(7.8.1-7.8.14), (1998) http://dx.doi.org/10.1002/0471140864.ps0708s11.

[23] A. D. Hanlon, M. I. Larkin, and R. M. Reddick. "Free-solution, label-free protein-protein interactions characterized by dynamic light scattering". *Biophys J*, 98(2):297–304, (2010) http://dx.doi.org/10.1016/j.bpj.2009.09.061.

[24] S. Li, D. Xing, and J. Li. "Dynamic light scattering application to study protein interactions in electrolyte solutions". *J Biol Phys*, 30(4):313–24, (2004) http://dx.doi.org/10.1007/s10867-004-0997-z.

[25] S. M. Kelly and N. C. Price. "Circular dichroism to study protein interactions". *Curr Protoc Protein Sci*, Chapter 20:Unit 20 10, (2006) http://dx.doi.org/10.1002/0471140864.ps2010s46.

[26] M. M. Pierce, C. S. Raman, and B. T. Nall. "Isothermal titration calorimetry of protein-protein interactions". *Methods*, 19(2):213–21, (1999) http://dx.doi.org/10.1006/meth.1999.0852.

[27] A. Velazquez-Campoy, S. A. Leavitt, and E. Freire. "Characterization of protein-protein interactions by isothermal titration calorimetry". *Methods Mol Biol*, 1278:183–204, (2015) http://dx.doi.org/10.1007/978-1-4939-2425-7_11.

[28] Y. Du. "Fluorescence polarization assay to quantify protein-protein interactions in an HTS format". *Methods Mol Biol*, 1278:529–44, (2015) http://dx.doi.org/10.1007/978-1-4939-2425-7_35.

[29] A. Gijsbers, T. Nishigaki, and N. Sanchez-Puig. "Fluorescence anisotropy as a tool to study protein-protein interactions". *J Vis Exp*, (116), (2016) http://dx.doi.org/10.3791/54640.

[30] A. T. Tuukkanen and D. I. Svergun. "Weak protein-ligand interactions studied by small-angle X-ray scattering". *FEBS J*, 281(8):1974–87, (2014) http://dx.doi.org/10.1111/febs.12772.

[31] N. Jones. "Crystallography: Atomic secrets". *Nature*, 505(7485):602–3, (2014) http://dx.doi.org/10.1038/505602a.

[32] M. R. O'Connell, R. Gamsjaeger, and J. P. Mackay. "The structural analysis of protein-protein interactions by NMR spectroscopy". *Proteomics*, 9(23):5224–32, (2009) http://dx.doi.org/10.1002/pmic.200900303.

[33] K. Chen and N. Tjandra. "The use of residual dipolar coupling in studying proteins by NMR". *Top Curr Chem*, 326:47–67, (2012) http://dx.doi.org/10.1007/128_2011_215.

[34] N. V. Nucci, B. S. Marques, S. Bedard, J. Dogan, Jr. Gledhill, J. M., V. R. Moorman, R. W. Peterson, K. G. Valentine, A. L. Wand, and A. J. Wand. "Optimization of NMR spectroscopy of encapsulated proteins dissolved in low viscosity fluids". *J Biomol NMR*, 50(4):421–30, (2011) http://dx.doi.org/10.1007/s10858-011-9528-y.

[35] K. Pervushin, R. Riek, G. Wider, and K. Wüthrich. "Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution". *Proc Natl Acad Sci U S A*, 94(23):12366–71, (1997).

[36] C. Domínguez, R. Boelens, and A. M. Bonvin. "HADDOCK: a protein-protein docking approach based on biochemical or biophysical information". *J Am Chem Soc*, 125(7):1731–7, (2003) http://dx.doi.org/10.1021/ja026939x.

[37] T. Lengauer and M. Rarey. "Computational methods for biomolecular docking". *Curr Opin Struct Biol*, 6(3):402–6, (1996).

[38] G. C. P. van Zundert, J. Rodrigues, M. Trellet, C. Schmitz, P. L. Kastritis, E. Karaca, A. S. J. Melquiond, M. van Dijk, S. J. de Vries, and A. Bonvin. "The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes". *J Mol Biol*, 428(4):720–725, (2016) http://dx.doi.org/10.1016/j.jmb.2015.09.014.

[39] D. P. Frueh, A. C. Goodrich, S. H. Mishra, and S. R. Nichols. "NMR methods for structural studies of large monomeric and multimeric proteins". *Curr Opin Struct Biol*, 23(5):734–9, (2013) http://dx.doi.org/10.1016/j.sbi.2013.06.016.

[40] A. Gross, C. Hashimoto, H. Sticht, and J. Eichler. "Synthetic peptides as protein mimics". *Front Bioeng Biotechnol*, 3:211, (2015) http://dx.doi.org/10.3389/fbioe.2015.00211.

[41] E. de Alba, M. Rico, and M. A. Jiménez. "Cross-strand side-chain interactions versus turn conformation in beta-hairpins". *Protein Sci*, 6(12):2548–60, (1997) http://dx.doi.org/10.1002/pro.5560061207.

[42] M. P. Williamson. "Peptide structure determination by NMR". *Methods Mol Biol*, 17:69–85, (1993) http://dx.doi.org/10.1385/0-89603-215-9:69.

[43] J. Eichler. "Peptides as protein binding site mimetics". *Curr Opin Chem Biol*, 12(6):707–13, (2008) http://dx.doi.org/10.1016/j.cbpa.2008.09.023.

[44] R. A. Houghten, C. Pinilla, J. R. Appel, S. E. Blondelle, C. T. Dooley, J. Eichler, A. Nefzi, and J. M. Ostresh. "Mixture-based synthetic combinatorial libraries". *J Med Chem*, 42(19):3743–78, (1999).

[45] W. Li and N. B. Caberoy. "New perspective for phage display as an efficient and versatile technology of functional proteomics". *Appl Microbiol Biotechnol*, 85(4):909–19, (2010) http://dx.doi.org/10.1007/s00253-009-2277-0.

[46] A. Graf von Stosch, M. A. Jiménez, V. Kinzel, and J. Reed. "Solvent polarity-dependent structural refolding: a CD and NMR study of a 15 residue peptide". *Proteins*, 23(2):196–203, (1995) http://dx.doi.org/10.1002/prot.340230209.

[47] T. A. Cross, M. Sharma, M. Yi, and H. X. Zhou. "Influence of solubilizing environments on membrane protein structures". *Trends Biochem Sci*, 36(2):117–25, (2011) http://dx.doi.org/10.1016/j.tibs.2010.07.005.

[48] D. V. Tulumello and C. M. Deber. "Efficiency of detergents at maintaining membrane protein structures in their biologically relevant forms". *Biochim Biophys Acta*, 1818(5):1351–8, (2012) http://dx.doi.org/10.1016/j.bbamem.2012.01.013.

[49] R. Amiri, A. K. Bordbar, M. F. García-Mayoral, A. R. Khosropour, I. Mohammadpoor-Baltork, M. Menéndez, and D. V. Laurents. "Interactions of gemini surfactants with two model proteins: NMR, CD, and fluorescence spectroscopies". *J Colloid Interface Sci*, 369(1):245–55, (2012) http://dx.doi.org/10.1016/j.jcis.2011.11.062.

[50] X. Cui, X. Yang, H. Chen, A. H. Liu, S. Mao, M. Liu, H. Yuan, P. Luo, and Y. Du. "NMR investigation of the exchange kinetics of quaternary ammonium dimeric surfactants C14-s-C14.2Br". *J Phys Chem B*, 112(10):2874–9, (2008) http://dx.doi.org/10.1021/jp710784p.

[51] I. G. Denisov, Y. V. Grinkova, A. A. Lazarides, and S. G. Sligar. "Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size". *J Am Chem Soc*, 126(11):3477–87, (2004) http://dx.doi.org/10.1021/ja0393574.

[52] M. S. Terakawa, Y. Lin, M. Kinoshita, S. Kanemura, D. Itoh, T. Sugiki, M. Okumura, A. Ramamoorthy, and Y. H. Lee. "Impact of membrane curvature on amyloid aggregation". *Biochim Biophys Acta*, (2018) http://dx.doi.org/10.1016/j.bbamem.2018.04.012.

[53] M. Bornens. "The centrosome in cells and organisms". *Science*, 335(6067):422–6, (2012) http://dx.doi.org/10.1126/science.1209037.

[54] D. A. Compton. "Spindle assembly in animal cells". *Annu Rev Biochem*, 69:95–114, (2000) http://dx.doi.org/10.1146/annurev.biochem.69.1.95.

[55] N. Delgehyr, J. Sillibourne, and M. Bornens. "Microtubule nucleation and anchoring at the centrosome are independent processes linked by ninein function". *J Cell Sci*, 118(Pt 8):1565–75, (2005) http://dx.doi.org/10.1242/jcs.02302.

[56] P. Meraldi and E. A. Nigg. "The centrosome cycle". *FEBS Lett*, 521(1-3):9–13, (2002).

[57] E. A. Nigg and T. Stearns. "The centrosome cycle: Centriole biogenesis, duplication and inherent asymmetries". *Nat Cell Biol*, 13(10):1154–60, (2011) http://dx.doi.org/10.1038/ncb2345.

[58] M. Piel, J. Nordberg, U. Euteneuer, and M. Bornens. "Centrosome-dependent exit of cytokinesis in animal cells". *Science*, 291(5508):1550–3, (2001) http://dx.doi.org/10.1126/science.1057330.

[59] J. S. Andersen, C. J. Wilkinson, T. Mayor, P. Mortensen, E. A. Nigg, and M. Mann. "Proteomic characterization of the human centrosome by protein correlation profiling". *Nature*, 426(6966):570–4, (2003) http://dx.doi.org/10.1038/nature02166.

[60] L. Jakobsen, K. Vanselow, M. Skogs, Y. Toyoda, E. Lundberg, I. Poser, L. G. Falkenby, M. Bennetzen, J. Westendorf, E. A. Nigg, M. Uhlen, A. A. Hyman, and J. S. Andersen. "Novel asymmetrically localizing components of human centrosomes identified by complementary proteomics methods". *EMBO J*, 30(8):1520–35, (2011) http://dx.doi.org/10.1038/emboj.2011.63.

[61] H. Muller, D. Schmidt, S. Steinbrink, E. Mirgorodskaya, V. Lehmann, K. Habermann, F. Dreher, N. Gustavsson, T. Kessler, H. Lehrach, R. Herwig, J. Gobom, A. Ploubidou, M. Boutros, and B. M. Lange. "Proteomic and functional analysis of the mitotic *Drosophila* centrosome". *EMBO J*, 29(19):3344–57, (2010) http://dx.doi.org/10.1038/emboj.2010.210.

[62] A. Desai and T. J. Mitchison. "Microtubule polymerization dynamics". *Annu Rev Cell Dev Biol*, 13:83–117, (1997) http://dx.doi.org/10.1146/annurev.cellbio.13.1.83.

[63] A. J. Holland and D. W. Cleveland. "Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis". *Nat Rev Mol Cell Biol*, 10(7):478–87, (2009) http://dx.doi.org/10.1038/nrm2718.

[64] S. Charrasse, M. Mazel, S. Taviaux, P. Berta, T. Chow, and C. Larroque. "Characterization of the cDNA and pattern of expression of a new gene over-expressed in human hepatomas and colonic tumors". *Eur J Biochem*, 234(2):406–13, (1995).

[65] S. Charrasse, M. Schroeder, C. Gauthier-Rouviere, F. Ango, L. Cassimeris, D. L. Gard, and C. Larroque. "The TOGp protein is a new human microtubule-associated protein homologous to the *Xenopus* XMAP215". *J Cell Sci*, 111 (Pt 10):1371–83, (1998).

[66] D. L. Gard and M. W. Kirschner. "A microtubule-associated protein from *Xenopus* eggs that specifically promotes assembly at the plus-end". *J Cell Biol*, 105(5):2203–15, (1987).

[67] R. J. Vásquez, D. L. Gard, and L. Cassimeris. "XMAP from *Xenopus* eggs promotes rapid plus end assembly of microtubules and rapid microtubule polymer turnover". *J Cell Biol*, 127(4):985–93, (1994).

[68] J. Al-Bassam, N. A. Larsen, A. A. Hyman, and S. C. Harrison. "Crystal structure of a TOG domain: conserved features of XMAP215/Dis1-family TOG domains and implications for tubulin binding". *Structure*, 15(3):355–62, (2007) http://dx.doi.org/10.1016/j.str.2007.01.012.

[69] G. J. Brouhard, J. H. Stear, T. L. Noetzel, J. Al-Bassam, K. Kinoshita, S. C. Harrison, J. Howard, and A. A. Hyman. "XMAP215 is a processive microtubule polymerase". *Cell*, 132(1):79–88, (2008) http://dx.doi.org/10.1016/j.cell.2007.11.043.

[70] F. Gergely, C. Karlsson, I. Still, J. Cowell, J. Kilmartin, and J. W. Raff. "The TACC domain identifies a family of centrosomal proteins that can interact with microtubules". *Proc Natl Acad Sci U S A*, 97(26):14352–7, (2000) http://dx.doi.org/10.1073/pnas.97.26.14352.

[71] I. Peset and I. Vernos. "The TACC proteins: TACC-ling microtubule dynamics and centrosome function". *Trends Cell Biol*, 18(8):379–88, (2008) http://dx.doi.org/10.1016/j.tcb.2008.06.005.

[72] J. W. Raff. "Centrosomes and cancer: lessons from a TACC". *Trends Cell Biol*, 12(5):222–5, (2002).

[73] M. J. Lee, F. Gergely, K. Jeffers, S. Y. Peak-Chew, and J. W. Raff. "Msps/XMAP215 interacts with the centrosomal protein D-TACC to regulate microtubule behaviour". *Nat Cell Biol*, 3(7):643–9, (2001) http://dx.doi.org/10.1038/35083033.

[74] L. L. O'Brien, A. J. Albee, L. L. Liu, W. Tao, P. Dobrzyn, S. B. Lizarraga, and C. Wiese. "The *Xenopus* TACC homologue, maskin, functions in mitotic spindle assembly". *Mol Biol Cell*, 16(6):2836–2847, (2005).

[75] J. M. Bellanger and P. Gonczy. "TAC-1 and ZYG-9 form a complex that promotes microtubule assembly in *C-elegans* embryos". *Curr Biol*, 13(17):1488–1498, (2003).

[76] K. Kinoshita, T. L. Noetzel, L. Pelletier, K. Mechtler, D. N. Drechsel, A. Schwager, M. Lee, J. W. Raff, and A. A. Hyman. "Aurora A phosphorylation of TACC3/maskin is required for centrosome-dependent microtubule assembly in mitosis". *J Cell Biol*, 170(7):1047–55, (2005) http://dx.doi.org/10.1083/jcb.200503023.

[77] G. B. Mortuza, T. Cavazza, M. F. García-Mayoral, D. Hermida, I. Peset, J. G. Pedrero, N. Merino, F. J. Blanco, J. Lyngso, M. Bruix, J. S. Pedersen, I. Vernos, and G. Montoya. "XTACC3-XMAP215 association reveals an asymmetric interaction promoting microtubule elongation". *Nat Commun*, 5:6072, (2014) http://dx.doi.org/10.1038/Ncomms6072.

[78] M. A. Treviño, M. F. García-Mayoral, M. A. Jiménez, U. Bastolla, and M. Bruix. "Emergence of structure through protein-protein interactions and pH changes in dually predicted coiled-coil and disordered regions of centrosomal proteins". *Biochim Biophys Acta*, 1844(10):1808–1819, (2014) http://dx.doi.org/10.1016/j.bbapap.2014.07.019.

[79] M. Pancera, T. Zhou, A. Druz, I. S. Georgiev, C. Soto, J. Gorman, J. Huang, P. Acharya, G. Y. Chuang, G. Ofek, G. B. Stewart-Jones, J. Stuckey, R. T. Bailer, M. G. Joyce, M. K. Louder, N. Tumba, Y. Yang, B. Zhang, M. S. Cohen, B. F. Haynes, J. R. Mascola, L. Morris, J. B. Munro, S. C. Blanchard, W. Mothes, M. Connors, and P. D. Kwong. "Structure and immune recognition of trimeric pre-fusion HIV-1 Env". *Nature*, 514(7523):455–61, (2014) http://dx.doi.org/10.1038/nature13808.

[80] M. Pelay-Gimeno, A. Glas, O. Koch, and T. N. Grossmann. "Structure-based design of inhibitors of protein-protein interactions: mimicking peptide binding epitopes". *Angew Chem Int Ed Engl*, 54(31):8896–927, (2015) http://dx.doi.org/10.1002/anie.201412070.

[81] M. A. Jiménez, F. J. Blanco, M. Rico, J. Santoro, J. Herranz, and J. L. Nieto. "Periodic properties of proton conformational shifts in isolated protein helices". *Eur J Biochem*, 207(1):39–49, (1992) http://dx.doi.org/10.1111/j.1432-1033.1992.tb17017.x.

[82] V. Muñoz, L. Serrano, M. A. Jiménez, and M. Rico. "Structural analysis of peptides encompassing all α-helices of three α/β parallel proteins: Che-Y, flavodoxin and P21-Ras: Implications for α-Helix stability and the folding of α/β parallel proteins". *J Mol Biol*, 247(4):648–669, (1995) http://dx.doi.org/10.1016/S0022-2836(05)80145-7.

[83] A. R. Viguera, M. A. Jiménez, M. Rico, and L. Serrano. "Conformational analysis of peptides corresponding to β-hairpins and a β-sheet that represent the entire sequence of the α-spectrin SH3 domain". *J Mol Biol*, 255(3):507–521, (1996) http://dx.doi.org/10.1006/jmbi.1996.0042.

[84] E. Rujas, J. M. Caaveiro, A. Partida-Hanon, N. Gulzar, K. Morante, B. Apellániz, M. García-Porras, M. Bruix, K. Tsumoto, J. K. Scott, M. A. Jiménez, and J. L. Nieva. "Structural basis for broad neutralization of HIV-1 through the molecular recognition of 10E8 helical epitope at the membrane interface". *Sci Rep*, 6:38177, (2016) http://dx.doi.org/10.1038/srep38177.

[85] O. Spiga, A. Bernini, M. Scarselli, A. Ciutti, L. Bracci, L. Lozzi, B. Lelli, D. Di Maro, D. Calamandrei, and N. Niccolai. "Peptide-protein interactions studied by surface plasmon and nuclear magnetic resonances". *FEBS Letters*, 511(1-3):33–35, (2002) http://dx.doi.org/10.1016/S0014-5793(01)03274-4.

[86] D. Helmer and K. Schmitz. "Peptides and peptide analogs to inhibit protein-protein interactions". *Adv Exp Med Biol*, 917:147–83, (2016) http://dx.doi.org/10.1007/978-3-319-32805-8_8.

[87] D. Roccatano, G. Colombo, M. Fioroni, and A. E. Mark. "Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study". *Proc Natl Acad Sci U S A*, 99(19):12179–84, (2002) http://dx.doi.org/10.1073/pnas.182199699.

[88] M. Buck. "Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins". *Q Rev Biophys*, 31(3):297–355, (1998).

[89] M. A. Jiménez, M. Bruix, C. González, F. J. Blanco, J. L. Nieto, J. Herranz, and M. Rico. "CD and 1H-NMR studies on the conformational properties of peptide fragments from the C-terminal domain of thermolysin". *Eur J Biochem*, 211(3):569–81, (1993).

[90] C. A. Rohl and R. L. Baldwin. "Comparison of NH exchange and circular dichroism as techniques for measuring the parameters of the helix-coil transition in peptides". *Biochemistry*, 36(28):8435–42, (1997) http://dx.doi.org/10.1021/bi9706677.

[91] L. Whitmore and B. A. Wallace. "Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases". *Biopolymers*, 89(5):392–400, (2008) http://dx.doi.org/10.1002/bip.20853.

[92] M. Piotto, V. Saudek, and V. Sklenar. "Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions". *J Biomol NMR*, 2(6):661–665, (1992) http://dx.doi.org/10.1007/Bf02192855.

[93] J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, and K. Wüthrich. "Recommendations for the presentation of NMR structures of proteins and nucleic acids–IUPAC-IUBMB-IUPAB Inter-Union Task Group on the standardization of data bases of protein and nucleic acid structures determined by NMR spectroscopy". *Eur J Biochem*, 256(1):1–15, (1998).

[94] K. Wüthrich. "NMR of proteins and nucleic acids". *Wiley-Interscience, John Wiley Sons*, (1986).

[95] T. D. Goddard and D. G. Kneller. "Sparky 3, University of California, San Francisco". (2005).

[96] P. Guntert. "Automated NMR structure calculation with CYANA". *Methods Mol Biol*, 278:353–78, (2004) http://dx.doi.org/10.1385/1-59259-809-9:353.

[97] Y. Shen and A. Bax. "Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks". *J Biomol NMR*, 56(3):227–41, (2013) http://dx.doi.org/10.1007/s10858-013-9741-y.

[98] R. Koradi, M. Billeter, and K. Wüthrich. "MOLMOL: a program for display and analysis of macromolecular structures". *J Mol Graph*, 14(1):51–5, 29–32, (1996).

[99] L. Schrödinger. "The PyMOL molecular graphics system, version 1.3r1". (2010).

[100] D. S. Wishart and B. D. Sykes. "Chemical shifts as a tool for structure determination". *Methods Enzymol*, 239:363–92, (1994).

[101] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes. "H-1, C-13 and N-15 random coil NMR chemical-shifts of the common amino-acids .Investigations of nearest-neighbor effects". *J Biomol NMR*, 5(3):67, (1995).

[102] C. M. Santiveri, M. Rico, and M. A. Jiménez. "13C(alpha) and 13C(beta) chemical shifts as a tool to delineate beta-hairpin structures in peptides". *J Biomol NMR*, 19(4):331–45, (2001).

[103] M. E. Holtzer, S. Kumar, A. Holtzer, and D. L. Crimmins. "The CD of two-chain coiled coils: experiments on tropomyosin and tropomyosin segments in the tyrosine/disulfide spectral region". *Biopolymers*, 28(9):1597–612, (1989) http://dx.doi.org/10.1002/bip.360280909.

[104] A. Chakrabartty, T. Kortemme, and R. L. Baldwin. "Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions". *Protein Sci*, 3(5):843–52, (1994) http://dx.doi.org/10.1002/pro.5560030514.

[105] D. Stigter and K. A. Dill. "Charge effects on folded and unfolded proteins". *Biochemistry*, 29(5):1262–71, (1990).

[106] M. A. Treviño, M. Rodríguez-Rodríguez, I. Correas, M. Marcilla, J. P. Albar, M. Rico, M. A. Jiménez, and M. Bruix. "NMR characterisation of the minimal interacting regions of centrosomal proteins 4.1R and NuMA1: effect of phosphorylation". *BMC Biochemistry*, 11:7, (2010).

[107] S. Hormeño, B. Ibarra, F. J. Chichón, K. Habermann, B. M. H. Lange, J. M. Valpuesta, J. L. Carrascosa, and J. R. Arias-González. "Single centrosome manipulation reveals its electric charge and associated dynamic structure". *Biophys J*, 97(4):1022–1030, (2009).

[108] M. Mompeán, R. Hervás, Y. Xu, T. H. Tran, C. Guarnaccia, E. Buratti, F. Baralle, L. Tong, M. Carrión-Vazquez, A. E. McDermott, and D. V. Laurents. "Structural evidence of amyloid fibril formation in the putative aggregation domain of TDP-43". *J Phys Chem Lett*, 6(13):2608–15, (2015) http://dx.doi.org/10.1021/acs.jpclett.5b00918.

[109] X. Robert and P. Gouet. "Deciphering key features in protein structures with the new ENDscript server". *Nucleic Acids Res*, 42(W1):W320–W324, (2014) http://dx.doi.org/10.1093/nar/gku316.

[110] D. Singh, J. M. Chan, P. Zoppoli, F. Niola, R. Sullivan, A. Castano, E. M. Liu, J. Reichel, P. Porrati, S. Pellegatta, K. Qiu, Z. Gao, M. Ceccarelli, R. Riccardi, D. J. Brat, A. Guha, K. Aldape, J. G. Golfinos, D. Zagzag, T. Mikkelsen, G. Finocchiaro, A. Lasorella, R. Rabadan, and A. Iavarone. "Transforming fusions of FGFR and TACC genes in human glioblastoma". *Science*, 337(6099):1231–5, (2012) http://dx.doi.org/10.1126/science.1220834.

[111] P. Luciw. "Human immunodeficiency viruses and their replication". *Fields virology*, (1996).

[112] F. Barre-Sinoussi, J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier. "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)". *Science*, 220(4599):868–71, (1983).

[113] M. S. Gottlieb, R. Schroff, H. M. Schanker, J. D. Weisman, P. T. Fan, R. A. Wolf, and A. Saxon. "Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency". *N Engl J Med*, 305(24):1425–31, (1981) http://dx.doi.org/10.1056/NEJM198112103052401.

[114] UNAIDS. "Report on the Global AIDS Epidemic". *Geneva*, (2017).

[115] S. Nyamweya, A. Hegedus, A. Jaye, S. Rowland-Jones, K. L. Flanagan, and D. C. Macallan. "Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis". *Rev Med Virol*, 23(4):221–40, (2013) http://dx.doi.org/10.1002/rmv.1739.

[116] K. Schmitt and R. Akkina. "Ultra-sensitive HIV-1 latency viral outgrowth assays using humanized mice". *Front Immunol*, 9:344, (2018) http://dx.doi.org/10.3389/fimmu.2018.00344.

[117] M. B. Feinberg and W. C. Greene. "Molecular insights into human immunodeficiency virus type 1 pathogenesis". *Curr Opin Immunol*, 4(4):466–74, (1992).

[118] A. S. Perelson and R. M. Ribeiro. "Estimating drug efficacy and viral dynamic parameters: HIV and HCV". *Stat Med*, 27(23):4647–57, (2008) http://dx.doi.org/10.1002/sim.3116.

[119] A. Rambaut, D. Posada, K. A. Crandall, and E. C. Holmes. "The causes and consequences of HIV evolution". *Nat Rev Genet*, 5(1):52–61, (2004) http://dx.doi.org/10.1038/nrg1246.

[120] D. L. Robertson, B. H. Hahn, and P. M. Sharp. "Recombination in AIDS viruses". *J Mol Evol*, 40(3):249–59, (1995).

[121] S. L. McGovern, E. Caselli, N. Grigorieff, and B. K. Shoichet. "A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening". *J Med Chem*, 45(8):1712–22, (2002).

[122] S. Fiorentini, E. Marini, S. Caracciolo, and A. Caruso. "Functions of the HIV-1 matrix protein p17". *New Microbiol*, 29(1):1–10, (2006).

[123] Y. Zeinolabediny, F. Caccuri, L. Colombo, F. Morelli, M. Romeo, A. Rossi, S. Schiarea, C. Ciaramelli, C. Airoldi, R. Weston, L. Donghui, J. Krupinski, R. Corpas, E. García-Lara, S. Sarroca, C. Sanfeliu, M. Slevin, A. Caruso, M. Salmona, and L. Diomede. "HIV-1 matrix protein p17 misfolding forms toxic amyloidogenic assemblies that induce neurocognitive disorders". *Sci Rep*, 7(1):10313, (2017) http://dx.doi.org/10.1038/s41598-017-10875-0.

[124] G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. Zhang. "Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics". *Nature*, 497(7451):643–6, (2013) http://dx.doi.org/10.1038/nature12162.

[125] L. Dawson and X. F. Yu. "The role of nucleocapsid of HIV-1 in virus assembly". *Virology*, 251(1):141–57, (1998) http://dx.doi.org/10.1006/viro.1998.9374.

[126] G. Li and E. De Clercq. "HIV Genome-wide protein associations: a review of 30 years of research". *Microbiol Mol Biol Rev*, 80(3):679–731, (2016) http://dx.doi.org/10.1128/MMBR.00065-15.

[127] W. S. Hu and H. M. Temin. "Retroviral recombination and reverse transcription". *Science*, 250(4985):1227–33, (1990).

[128] D. C. Chan, D. Fass, J. M. Berger, and P. S. Kim. "Core structure of gp41 from the HIV envelope glycoprotein". *Cell*, 89(2):263–73, (1997).

[129] D. C. Chan and P. S. Kim. "HIV entry and its inhibition". *Cell*, 93(5):681–4, (1998).

[130] S. Serrano, N. Huarte, E. Rujas, D. Andreu, J. L. Nieva, and M. A. Jiménez. "Structure-related roles for the conservation of the HIV-1 fusion peptide sequence revealed by Nuclear Magnetic Resonance". *Biochemistry*, 56(41):5503–5511, (2017) http://dx.doi.org/10.1021/acs.biochem.7b00745.

[131] R. Wyatt and J. Sodroski. "The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens". *Science*, 280(5371):1884–8, (1998).

[132] S. Serrano, A. Araujo, B. Apellániz, S. Bryson, P. Carravilla, I. de la Arada, N. Huarte, E. Rujas, E. F. Pai, J. L. Arrondo, C. Domene, M. A. Jiménez, and J. L. Nieva. "Structure and immunogenicity of a peptide vaccine, including the complete HIV-1 gp41 2F5 epitope: implications for antibody recognition mechanism and immunogen design". *J Biol Chem*, 289(10):6565–80, (2014) http://dx.doi.org/10.1074/jbc.M113.527747.

[133] B. Apellániz, E. Rujas, S. Serrano, K. Morante, K. Tsumoto, J. M. Caaveiro, M. A. Jiménez, and J. L. Nieva. "The atomic structure of the HIV-1 gp41 transmembrane domain and its connection to the immunogenic membrane-proximal external region". *J Biol Chem*, 290(21):12999–3015, (2015) http://dx.doi.org/10.1074/jbc.M115.644351.

[134] D. H. Gabuzda, A. Lever, E. Terwilliger, and J. Sodroski. "Effects of deletions in the cytoplasmic domain on biological functions of human immunodeficiency virus type 1 envelope glycoproteins". *J Virol*, 66(6):3306–15, (1992).

[135] M. A. Checkley, B. G. Luttge, and E. O. Freed. "HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation". *J Mol Biol*, 410(4):582–608, (2011) http://dx.doi.org/10.1016/j.jmb.2011.04.042.

[136] J. W. Dubay, S. J. Roberts, B. Brody, and E. Hunter. "Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity". *J Virol*, 66(8):4748–56, (1992).

[137] M. Lu, S. C. Blacklow, and P. S. Kim. "A trimeric structural domain of the HIV-1 transmembrane glycoprotein". *Nat Struct Biol*, 2(12):1075–82, (1995).

[138] W. Weissenhorn, A. Dessen, S. C. Harrison, J. J. Skehel, and D. C. Wiley. "Atomic structure of the ectodomain from HIV-1 gp41". *Nature*, 387(6631):426–30, (1997) http://dx.doi.org/10.1038/387426a0.

[139] R. Brasseur, M. Vandenbranden, B. Cornet, A. Burny, and J. M. Ruysschaert. "Orientation into the lipid bilayer of an asymmetric amphipathic helical peptide located at the N-terminus of viral fusion proteins". *Biochim Biophys Acta*, 1029(2):267–73, (1990).

[140] R. A. Furuta, C. T. Wild, Y. Weng, and C. D. Weiss. "Capture of an early fusion-active conformation of HIV-1 gp41". *Nat Struct Biol*, 5(4):276–9, (1998).

[141] J. M. Kilby, S. Hopkins, T. M. Venetta, B. DiMassimo, G. A. Cloud, J. Y. Lee, L. Alldredge, E. Hunter, D. Lambert, D. Bolognesi, T. Matthews, M. R. Johnson, M. A. Nowak, G. M. Shaw, and M. S. Saag. "Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry". *Nat Med*, 4(11):1302–7, (1998) http://dx.doi.org/10.1038/3293.

[142] J. LaBonte, J. Lebbos, and P. Kirkpatrick. "Enfuvirtide". *Nat Rev Drug Discov*, 2:345, (2003) http://dx.doi.org/10.1038/nrd1091.

[143] Q. Wang, W. Bi, X. Zhu, H. Li, Q. Qi, F. Yu, L. Lu, and S. Jiang. "Nonneutralizing antibodies induced by the HIV-1 gp41 NHR domain gain neutralizing activity in the presence of the HIV fusion inhibitor Enfuvirtide: a potential therapeutic vaccine strategy". *J Virol*, 89(13):6960–4, (2015) http://dx.doi.org/10.1128/JVI.00791-15.

[144] C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, and T. J. Matthews. "Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection". *Proc Natl Acad Sci U S A*, 91(21):9770–4, (1994).

[145] J. Cao, L. Bergeron, E. Helseth, M. Thali, H. Repke, and J. Sodroski. "Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein". *J Virol*, 67(5):2747–55, (1993).

[146] M. B. Zwick, A. F. Labrijn, M. Wang, C. Spenlehauer, E. O. Saphire, J. M. Binley, J. P. Moore, G. Stiegler, H. Katinger, D. R. Burton, and P. W. Parren. "Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41". *J Virol*, 75(22):10892–905, (2001) http://dx.doi.org/10.1128/JVI.75.22.10892-10905.2001.

[147] N. Kondo, K. Miyauchi, F. Meng, A. Iwamoto, and Z. Matsuda. "Conformational changes of the HIV-1 envelope protein during membrane fusion are inhibited by the replacement of its membrane-spanning domain". *J Biol Chem*, 285(19):14681–8, (2010) http://dx.doi.org/10.1074/jbc.M109.067090.

[148] L. Shang and E. Hunter. "Residues in the membrane-spanning domain core modulate conformation and fusogenicity of the HIV-1 envelope glycoprotein". *Virology*, 404(2):158–67, (2010) http://dx.doi.org/10.1016/j.virol.2010.03.016.

[149] L. Shang, L. Yue, and E. Hunter. "Role of the membrane-spanning domain of human immunodeficiency virus type 1 envelope glycoprotein in cell-cell fusion and virus infection". *J Virol*, 82(11):5417–28, (2008) http://dx.doi.org/10.1128/JVI.02666-07.

[150] A. A. Waheed, S. D. Ablan, F. Soheilian, K. Nagashima, A. Ono, C. P. Schaffner, and E. O. Freed. "Inhibition of human immunodeficiency virus type 1 assembly and release by the cholesterol-binding compound amphotericin B methyl ester: evidence for Vpu dependence". *J Virol*, 82(19):9776–81, (2008) http://dx.doi.org/10.1128/JVI.00917-08.

[151] A. A. Waheed, S. D. Ablan, R. C. Sowder, J. D. Roser, C. P. Schaffner, E. Chertova, and E. O. Freed. "Effect of mutations in the human immunodeficiency virus type 1 protease on cleavage of the gp41 cytoplasmic tail". *J Virol*, 84(6):3121–6, (2010) http://dx.doi.org/10.1128/JVI.02002-09.

[152] R. C. Kennedy, R. D. Henkel, D. Pauletti, J. S. Allan, T. H. Lee, M. Essex, and G. R. Dreesman. "Antiserum to a synthetic peptide recognizes the HTLV-III envelope glycoprotein". *Science*, 231(4745):1556–9, (1986).

[153] L. Lu, Y. Zhu, J. Huang, X. Chen, H. Yang, S. Jiang, and Y. H. Chen. "Surface exposure of the HIV-1 env cytoplasmic tail LLP2 domain during the membrane fusion process: interaction with gp41 fusion core". *J Biol Chem*, 283(24):16723–31, (2008) http://dx.doi.org/10.1074/jbc.M801083200.

[154] G. B. Karlsson Hedestam, R. A. M. Fouchier, S. Phogat, D. R. Burton, J. Sodroski, and R. T. Wyatt. "The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus". *Nat Rev Microbiol*, 6:143, (2008) http://dx.doi.org/10.1038/nrmicro1819.

[155] J. Huang, G. Ofek, L. Laub, M. K. Louder, N. A. Doria-Rose, N. S. Longo, H. Imamichi, R. T. Bailer, B. Chakrabarti, S. K. Sharma, S. M. Alam, T. Wang, Y. Yang, B. Zhang, S. A. Migueles, R. Wyatt, B. F. Haynes, P. D. Kwong, J. R. Mascola, and M. Connors. "Broad and potent neutralization of HIV-1 by a gp41-specific human antibody". *Nature*, 491(7424):406–12, (2012) http://dx.doi.org/10.1038/nature11544.

[156] A. S. Kim, D. P. Leaman, and M. B. Zwick. "Antibody to gp41 MPER alters functional properties of HIV-1 Env without complete neutralization". *PLoS Pathog*, 10(7):e1004271, (2014) http://dx.doi.org/10.1371/journal.ppat.1004271.

[157] V. K. Gangupomu and C. F. Abrams. "All-atom models of the membrane-spanning domain of HIV-1 gp41 from metadynamics". *Biophys J*, 99(10):3438–44, (2010) http://dx.doi.org/10.1016/j.bpj.2010.09.054.

[158] G. Manzo, M. Carboni, A. C. Rinaldi, M. Casu, and M. A. Scorciapino. "Characterization of sodium dodecylsulphate and dodecylphosphocholine mixed micelles through NMR and dynamic light scattering". *Magn Reson Chem*, 51(3):176–83, (2013) http://dx.doi.org/10.1002/mrc.3930.

[159] R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein, and J. M. Thornton. "AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR". *J Biomol NMR*, 8(4):477–86, (1996).

[160] L. M. Gordon, P. W. Mobley, R. Pilpa, M. A. Sherman, and A. J. Waring. "Conformational mapping of the N-terminal peptide of HIV-1 gp41 in membrane environments using (13)C-enhanced Fourier transform infrared spectroscopy". *Biochim Biophys Acta*, 1559(2):96–120, (2002).

[161] A. Kato, K. Maki, T. Ebina, K. Kuwajima, K. Soda, and Y. Kuroda. "Mutational analysis of protein solubility enhancement using short peptide tags". *Biopolymers*, 85(1):12–8, (2007) http://dx.doi.org/10.1002/bip.20596.

[162] A. M. Ruschak, J. D. Rose, M. P. Coughlin, and T. L. Religa. "Engineered solubility tag for solution NMR of proteins". *Protein Sci*, 22(11):1646–54, (2013) http://dx.doi.org/10.1002/pro.2337.

[163] J. L. Nieva and A. Agirre. "Are fusion peptides a good model to study viral cell fusion?". *Biochim Biophys Acta*, 1614(1):104–15, (2003).

[164] W. Weissenhorn, A. Dessen, L. J. Calder, S. C. Harrison, J. J. Skehel, and D. C. Wiley. "Structural basis for membrane fusion by enveloped viruses". *Mol Membr Biol*, 16(1):3–9, (1999).

[165] S. K. McDonald and K. G. Fleming. "Aromatic side chain water-to-lipid transfer free energies show a depth-dependence across the membrane normal". *J Am Chem Soc*, (2016) http://dx.doi.org/10.1021/jacs.6b03460.

[166] S. H. White, A. S. Ladokhin, S. Jayasinghe, and K. Hristova. "How membranes shape protein structure". *J Biol Chem*, 276(35):32395–8, (2001) http://dx.doi.org/10.1074/jbc.R100008200.

[167] S. H. White and W. C. Wimley. "Membrane protein folding and stability: physical principles". *Annu Rev Biophys Biomol Struct*, 28:319–65, (1999).

[168] W. C. Wimley and S. H. White. "Experimentally determined hydrophobicity scale for proteins at membrane interfaces". *Nat Struct Biol*, 3(10):842–8, (1996).

[169] B. Apellaniz, S. Nir, and J. L. Nieva. "Distinct mechanisms of lipid bilayer perturbation induced by peptides derived from the membrane-proximal external region of HIV-1 gp41". *Biochemistry*, 48(23):5320–31, (2009) http://dx.doi.org/10.1021/bi900504t.

[170] K. Shin, M. Sarker, S. K. Huang, and J. K. Rainey. "Apelin conformational and binding equilibria upon micelle interaction primarily depend on membrane-mimetic headgroup". *Sci Rep*, 7(1):15433, (2017) http://dx.doi.org/10.1038/s41598-017-14784-0.

[171] L. de Armas-Rillo, M. S. Valera, S. Marrero-Hernandez, and A. Valenzuela-Fernandez. "Membrane dynamics associated with viral infection". *Rev Med Virol*, 26(3):146–60, (2016) http://dx.doi.org/10.1002/rmv.1872.

[172] J. Davies. "Inactivation of antibiotics and the dissemination of resistance genes". *Science*, 264(5157):375–82, (1994).

[173] A. R. Coates, G. Halls, and Y. Hu. "Novel classes of antibiotics or more of the same?". *Br J Pharmacol*, 163(1):184–94, (2011) http://dx.doi.org/10.1111/j.1476-5381.2011.01250.x.

[174] X. Zhao, H. Wu, H. Lu, G. Li, and Q. Huang. "LAMP: A Database Linking Antimicrobial Peptides". *PLoS One*, 8(6):e66557, (2013) http://dx.doi.org/10.1371/journal.pone.0066557.

[175] I. M. Gould and A. M. Bal. "New antibiotic agents in the pipeline and how they can help overcome microbial resistance". *Virulence*, 4(2):185–91, (2013) http://dx.doi.org/10.4161/viru.22507.

[176] S. Sengupta, M. K. Chattopadhyay, and H. P. Grossart. "The multifaceted roles of antibiotics and antibiotic resistance in nature". *Front Microbiol*, 4:47, (2013) http://dx.doi.org/10.3389/fmicb.2013.00047.

[177] G. D. Wright. "Something old, something new: revisiting natural products in antibiotic drug discovery". *Can J Microbiol*, 60(3):147–54, (2014) http://dx.doi.org/10.1139/cjm-2014-0063.

[178] G. L. French. "The continuing crisis in antibiotic resistance". *Int J Antimicrob Agents*, 36 Suppl 3:S3–7, (2010) http://dx.doi.org/10.1016/S0924-8579(10)70003-0.

[179] HPA. "Mandatory MRSA bacteraemia and *Clostridium difficile* infection (up to January-March 2010)". *Health Protection Agency, Quarterly Epidemiological Commentary*, (2010).

[180] CDC. "Emerging infectious diseases. Centers for Disease Control and Prevention, Office of Infectious Disease Antibiotic resistance threats in the United States". 24(3):212, (2018).

[181] CDC. "Antibiotics resistance threats in the United States. U.S department of health and human services centers for disease control and prevention.". *Office of Infectious Disease Antibiotic resistance threats in the United States*, Accessed January 26, (2018).

[182] B. Spellberg and D. N. Gilbert. "The future of antibiotics and resistance: a tribute to a career of leadership by John Bartlett". *Clin Infect Dis*, 59 Suppl 2:S71–5, (2014) http://dx.doi.org/10.1093/cid/ciu392.

[183] W. L. Maloy and U. P. Kari. "Structure-activity studies on magainins and other host defense peptides". *Biopolymers*, 37(2):105–22, (1995) http://dx.doi.org/10.1002/bip.360370206.

[184] M. Vaara. "Agents that increase the permeability of the outer membrane". *Microbiol Rev*, 56(3):395–411, (1992).

[185] Q. Huang, X. Zhao, J. Huang, H. Wu, H. Lu, and G. Li. "LAMP: A database linking antimicrobial peptides.". Accessed January 27, 2018(http://biotechlab.fudan.edu.cn/database/lamp/index.php), (2016).

[186] R. J. Dubos. "Studies on a bactericidal agent extracted from a soil *Bacillus* : I. Preparation of the agent. Its activity *in vitro* ". *J Exp Med*, 70(1):1–10, (1939).

[187] R. D. Hotchkiss and R. J. Dubos. "Fractionation of the bactericidal agent from cultures of a soil *Bacillus*.". *J Biol Chem* , 132:791–792, (1940).

[188] C. H. Rammelkamp and L. Weinstein. "Toxic effects of Tyrothricin, Gramicidin and Tyrocidine". *J Infect Dis*, 71(2):166–173, (1942) http://dx.doi.org/10.1093/infdis/71.2.166.

[189] A. A. Bahar and D. Ren. "Antimicrobial peptides". *Pharmaceuticals (Basel)*, 6(12):1543–75, (2013) http://dx.doi.org/10.3390/ph6121543.

[190] J. P. Silva, S. Dhall, M. García, A. Chan, C. Costa, M. Gama, and M. Martins-Green. "Improved burn wound healing by the antimicrobial peptide LLKKK18 released from conjugates with dextrin embedded in a carbopol gel". *Acta Biomater*, 26:249–62, (2015) http://dx.doi.org/10.1016/j.actbio.2015.07.043.

[191] H. Jenssen, P. Hamill, and R. E. Hancock. "Peptide antimicrobial agents". *Clin Microbiol Rev*, 19(3):491–511, (2006) http://dx.doi.org/10.1128/CMR.00056-05.

[192] K. Sivanesam, B. L. Kier, S. D. Whedon, C. Chatterjee, and N. H. Andersen. "Hairpin structure stability plays a role in the activity of two antimicrobial peptides". *FEBS Lett*, 590(24):4480–4488, (2016) http://dx.doi.org/10.1002/1873-3468.12477.

[193] R. Eckert, F. Qi, D. K. Yarbrough, J. He, M. H. Anderson, and W. Shi. "Adding selectivity to antimicrobial peptides: rational design of a multidomain peptide against Pseudomonas spp". *Antimicrob Agents Chemother*, 50(4):1480–8, (2006) http://dx.doi.org/10.1128/AAC.50.4.1480-1488.2006.

[194] K. Matsuzaki. "Control of cell selectivity of antimicrobial peptides". *Biochim Biophys Acta*, 1788(8):1687–92, (2009) http://dx.doi.org/10.1016/j.bbamem.2008.09.013.

[195] S. Pacor, A. Giangaspero, M. Bacac, G. Sava, and A. Tossi. "Analysis of the cytotoxicity of synthetic antimicrobial peptides on mouse leucocytes: implications for systemic use". *J Antimicrob Chemother*, 50(3):339–48, (2002).

[196] N. Papo, Z. Oren, U. Pag, H. G. Sahl, and Y. Shai. "The consequence of sequence alteration of an amphipathic alpha-helical antimicrobial peptide and its diastereomers". *J Biol Chem*, 277(37):33913–21, (2002) http://dx.doi.org/10.1074/jbc.M204928200.

[197] M. Sieprawska-Lupa, P. Mydel, K. Krawczyk, K. Wojcik, M. Puklo, B. Lupa, P. Suder, J. Silberring, M. Reed, J. Pohl, W. Shafer, F. McAleese, T. Foster, J. Travis, and J. Potempa. "Degradation of human antimicrobial peptide LL-37 by *Staphylococcus aureus*-derived proteinases". *Antimicrob Agents Chemother*, 48(12):4673–9, (2004) http://dx.doi.org/10.1128/AAC.48.12.4673-4679.2004.

[198] J. Svenson, W. Stensen, B. O. Brandsdal, B. E. Haug, J. Monrad, and J. S. Svendsen. "Antimicrobial peptides with stability toward tryptic degradation". *Biochemistry*, 47(12):3777–88, (2008) http://dx.doi.org/10.1021/bi7019904.

[199] B. Cornet, J. M. Bonmatin, C. Hetru, J. A. Hoffmann, M. Ptak, and F. Vovelle. "Refined three-dimensional solution structure of insect defensin A". *Structure*, 3(5):435–48, (1995).

[200] M. R. Yeaman and N. Y. Yount. "Mechanisms of antimicrobial peptide action and resistance". *Pharmacol Rev*, 55(1):27–55, (2003) http://dx.doi.org/10.1124/pr.55.1.2.

[201] G. R. Zimmermann, P. Legault, M. E. Selsted, and A. Pardi. "Solution structure of bovine neutrophil beta-defensin-12: the peptide fold of the beta-defensins is identical to that of the classical defensins". *Biochemistry*, 34(41):13663–71, (1995).

[202] E. F. Haney and H. J. Vogel. "NMR of antimicrobial peptides". *Annual Reports on NMR Spectroscopy*, 65:1–51, (2009) http://dx.doi.org/10.1016/S0066-4103(08)00201-9.

[203] A. M. McManus, N. F. Dawson, J. D. Wade, L. E. Carrington, D. J. Winzor, and D. J. Craik. "Three-dimensional structure of RK-1: a novel alpha-defensin peptide". *Biochemistry*, 39(51):15757–64, (2000).

[204] M. Uteng, H. H. Hauge, P. R. Markwick, G. Fimland, D. Mantzilas, J. Nissen-Meyer, and C. Muhle-Goll. "Three-dimensional structure in lipid micelles of the pediocin-like antimicrobial peptide sakacin P and a sakacin P variant that is structurally stabilized by an inserted C-terminal disulfide bridge". *Biochemistry*, 42(39):11417–26, (2003) http://dx.doi.org/10.1021/bi034572i.

[205] K. A. Brogden. "Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?". *Nat Rev Microbiol*, 3(3):238–50, (2005) http://dx.doi.org/10.1038/nrmicro1098.

[206] A. Bastian and H. Schafer. "Human alpha-defensin 1 (HNP-1) inhibits adenoviral infection in vitro". *Regul Pept*, 101(1-3):157–61, (2001).

[207] A. Belaid, M. Aouni, R. Khelifa, A. Trabelsi, M. Jemmali, and K. Hani. "In vitro antiviral activity of dermaseptins against herpes simplex virus type 1". *J Med Virol*, 66(2):229–34, (2002).

[208] W. S. Horne, C. M. Wiethoff, C. Cui, K. M. Wilcoxen, M. Amorin, M. R. Ghadiri, and G. R. Nemerow. "Antiviral cyclic D,L-alpha-peptides: targeting a general biochemical pathway in virus infections". *Bioorg Med Chem*, 13(17):5145–53, (2005) http://dx.doi.org/10.1016/j.bmc.2005.05.051.

[209] Jr. Robinson, W. E., B. McDougall, D. Tran, and M. E. Selsted. "Anti-HIV-1 activity of indolicidin, an antimicrobial peptide from neutrophils". *J Leukoc Biol*, 63(1):94–100, (1998).

[210] N. Sitaram and R. Nagaraj. "Interaction of antimicrobial peptides with biological and model membranes: structural and charge requirements for activity". *Biochim Biophys Acta*, 1462(1-2):29–54, (1999).

[211] B. Yasin, W. Wang, M. Pang, N. Cheshenko, T. Hong, A. J. Waring, B. C. Herold, E. A. Wagar, and R. I. Lehrer. "Theta defensins protect cells from infection by herpes simplex virus by inhibiting viral adhesion and entry". *J Virol*, 78(10):5147–56, (2004).

[212] N. L. van der Weerden, R. E. Hancock, and M. A. Anderson. "Permeabilization of fungal hyphae by the plant defensin NaD1 occurs through a cell wall-dependent process". *J Biol Chem*, 285(48):37513–20, (2010) http://dx.doi.org/10.1074/jbc.M110.134882.

[213] L. Moerman, S. Bosteels, W. Noppe, J. Willems, E. Clynen, L. Schoofs, K. Thevissen, J. Tytgat, J. Van Eldere, J. Van Der Walt, and F. Verdonck. "Antibacterial and antifungal properties of alpha-helical, cationic peptides in the venom of scorpions from southern Africa". *Eur J Biochem*, 269(19):4799–810, (2002).

[214] Y. Park, S. H. Jang, D. G. Lee, and K. S. Hahm. "Antinematodal effect of antimicrobial peptide, PMAP-23, isolated from porcine myeloid against *Caenorhabditis elegans*". *J Pept Sci*, 10(5):304–11, (2004) http://dx.doi.org/10.1002/psc.518.

[215] Y. Shai. "Mode of action of membrane active antimicrobial peptides". *Biopolymers*, 66(4):236–48, (2002) http://dx.doi.org/10.1002/bip.10260.

[216] L. Zhang, A. Rozek, and R. E. Hancock. "Interaction of cationic antimicrobial peptides with model membranes". *J Biol Chem*, 276(38):35714–22, (2001) http://dx.doi.org/10.1074/jbc.M104925200.

[217] D. G. Lee, H. K. Kim, S. A. Kim, Y. Park, S. C. Park, S. H. Jang, and K. S. Hahm. "Fungicidal effect of indolicidin and its interaction with phospholipid membranes". *Biochem Biophys Res Commun*, 305(2):305–10, (2003).

[218] C. Subbalakshmi and N. Sitaram. "Mechanism of antimicrobial action of indolicidin". *FEMS Microbiol Lett*, 160(1):91–6, (1998).

[219] Y. M. Zhang and C. O. Rock. "Transcriptional regulation in bacterial membrane lipid synthesis". *J Lipid Res*, 50 Suppl:S115–9, (2009) http://dx.doi.org/10.1194/jlr.R800046-JLR200.

[220] F. Madani, S. Lindberg, U. Langel, S. Futaki, and A. Graslund. "Mechanisms of cellular uptake of cell-penetrating peptides". *J Biophys*, 2011:414729, (2011) http://dx.doi.org/10.1155/2011/414729.

[221] A. Pokorny and P. F. Almeida. "Permeabilization of raft-containing lipid vesicles by delta-lysin: a mechanism for cell sensitivity to cytotoxic peptides". *Biochemistry*, 44(27):9538–44, (2005) http://dx.doi.org/10.1021/bi0506371.

[222] M. Miteva, M. Andersson, A. Karshikoff, and G. Otting. "Molecular electroporation: a unifying concept for the description of membrane pore formation by antibacterial peptides, exemplified with NK-lysin". *FEBS Lett*, 462(1-2):155–8, (1999).

[223] A. T. Jones. "Macropinocytosis: searching for an endocytic identity and role in the uptake of cell penetrating peptides". *J Cell Mol Med*, 11(4):670–84, (2007) http://dx.doi.org/10.1111/j.1582-4934.2007.00062.x.

[224] H. G. Boman, B. Agerberth, and A. Boman. "Mechanisms of action on *Escherichia coli* of cecropin P1 and PR-39, two antibacterial peptides from pig intestine". *Infect Immun*, 61(7):2978–84, (1993).

[225] P. Nicolas. "Multifunctional host defense peptides: intracellular-targeting antimicrobial peptides". *FEBS J*, 276(22):6483–96, (2009) http://dx.doi.org/10.1111/j.1742-4658.2009.07359.x.

[226] K. H. Scheit, E. S. Reddy, and P. M. Bhargava. "Seminaplasmin is a potent inhibitor of *E. coli* RNA polymerase in vivo". *Nature*, 279(5715):728–31, (1979).

[227] S. N. Chitnis, K. S. Prasad, and P. M. Bhargava. "Bacteriolytic activity of seminalplasmin". *J Gen Microbiol*, 133(5):1265–71, (1987) http://dx.doi.org/10.1099/00221287-133-5-1265.

[228] S. N. Chitnis, K. S. Prasad, and P. M. Bhargava. "Isolation and characterization of autolysis-defective mutants of *Escherichia coli* that are resistant to the lytic activity of seminalplasmin". *J Gen Microbiol*, 136(3):463–9, (1990) http://dx.doi.org/10.1099/00221287-136-3-463.

[229] M. Ishikawa, T. Kubo, and S. Natori. "Purification and characterization of a diptericin homologue from *Sarcophaga peregrina* (flesh fly)". *Biochem J*, 287 (Pt 2):573–8, (1992).

[230] H. V. Westerhoff, D. Juretic, R. W. Hendler, and M. Zasloff. "Magainins and the disruption of membrane-linked free-energy transduction". *Proc Natl Acad Sci U S A*, 86(17):6597–601, (1989).

[231] C. Subbalakshmi, R. Nagaraj, and N. Sitaram. "Biological activities of C-terminal 15-residue synthetic fragment of melittin: design of an analog with improved antibacterial activity". *FEBS Lett*, 448(1):62–6, (1999).

[232] I. Zelezetsky, S. Pacor, U. Pag, N. Papo, Y. Shai, H. G. Sahl, and A. Tossi. "Controlled alteration of the shape and conformational stability of alpha-helical cell-lytic peptides: effect on mode of action and cell specificity". *Biochem J*, 390(Pt 1):177–88, (2005) http://dx.doi.org/10.1042/BJ20042138.

[233] Y. Huang, J. Huang, and Y. Chen. "Alpha-helical cationic antimicrobial peptides: relationships of structure and function". *Protein Cell*, 1(2):143–52, (2010) http://dx.doi.org/10.1007/s13238-010-0004-3.

[234] A. Tossi, L. Sandri, and A. Giangaspero. "Amphipathic, alpha-helical antimicrobial peptides". *Biopolymers*, 55(1):4–30, (2000) http://dx.doi.org/10.1002/1097-0282(2000)55:1<4::AID-BIP30>3.0.CO;2-M.

[235] M. Fernandez-Vidal, S. Jayasinghe, A. S. Ladokhin, and S. H. White. "Folding amphipathic helices into membranes: amphiphilicity trumps hydrophobicity". *J Mol Biol*, 370(3):459–70, (2007) http://dx.doi.org/10.1016/j.jmb.2007.05.016.

[236] A. Aumelas, M. Mangoni, C. Roumestand, L. Chiche, E. Despaux, G. Grassy, B. Calas, and A. Chavanieu. "Synthesis and solution structure of the antimicrobial peptide protegrin-1". *Eur J Biochem*, 237(3):575–83, (1996).

[237] S. S. Harwig, K. M. Swiderek, T. D. Lee, and R. I. Lehrer. "Determination of disulphide bridges in PG-2, an antimicrobial peptide from porcine leukocytes". *J Pept Sci*, 1(3):207–15, (1995) http://dx.doi.org/10.1002/psc.310010308.

[238] V. N. Kokryakov, S. S. Harwig, E. A. Panyutich, A. A. Shevchenko, G. M. Aleshina, O. V. Shamova, H. A. Korneva, and R. I. Lehrer. "Protegrins: leukocyte antimicrobial peptides that combine features of corticostatic defensins and tachyplesins". *FEBS Lett*, 327(2):231–6, (1993).

[239] O. A. Mirgorodskaya, A. A. Shevchenko, K. O. Abdalla, I. V. Chernushevich, T. A. Egorov, A. X. Musoliamov, V. N. Kokryakov, and O. V. Shamova. "Primary structure of three cationic peptides from porcine neutrophils. Sequence determination by the combined usage of electrospray ionization mass spectrometry and Edman degradation". *FEBS Lett*, 330(3):339–42, (1993).

[240] R. L. Fahrner, T. Dieckmann, S. S. Harwig, R. I. Lehrer, D. Eisenberg, and J. Feigon. "Solution structure of protegrin-1, a broad-spectrum antimicrobial peptide from porcine leukocytes". *Chem Biol*, 3(7):543–50, (1996).

[241] A. A. Langham, A. S. Ahmad, and Y. N. Kaznessis. "On the nature of antimicrobial activity: a model for protegrin-1 pores". *J Am Chem Soc*, 130(13):4338–46, (2008) http://dx.doi.org/10.1021/ja0780380.

[242] R. Mani, S. D. Cady, M. Tang, A. J. Waring, R. I. Lehrer, and M. Hong. "Membrane-dependent oligomeric structure and pore formation of a beta-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR". *Proc Natl Acad Sci U S A*, 103(44):16242–7, (2006) http://dx.doi.org/10.1073/pnas.0605079103.

[243] C. Roumestand, V. Louis, A. Aumelas, G. Grassy, B. Calas, and A. Chavanieu. "Oligomerization of protegrin-1 in the presence of DPC micelles. A proton high-resolution NMR study". *FEBS Lett*, 421(3):263–7, (1998).

[244] K. S. Usachev, S. V. Efimov, O. A. Kolosova, A. V. Filippov, and V. V. Klochkov. "High-resolution NMR structure of the antimicrobial peptide protegrin-2 in the presence of DPC micelles". *J Biomol NMR*, 61(3-4):227–34, (2015) http://dx.doi.org/10.1007/s10858-014-9885-4.

[245] K. S. Usachev, S. V. Efimov, O. A. Kolosova, E. A. Klochkova, A. V. Aganov, and V. V. Klochkov. "Antimicrobial peptide protegrin-3 adopt an antiparallel dimer in the presence of DPC micelles: a high-resolution NMR study". *J Biomol NMR*, 62(1):71–9, (2015) http://dx.doi.org/10.1007/s10858-015-9920-0.

[246] N. Bhattacharjee and P. Biswas. "Position-specific propensities of amino acids in the beta-strand". *BMC Struct Biol*, 10:29, (2010) http://dx.doi.org/10.1186/1472-6807-10-29.

[247] P. Y. Chou and G. D. Fasman. "Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins". *Biochemistry*, 13(2):211–22, (1974).

[248] P. Koehl and M. Levitt. "Structure-based conformational preferences of amino acids". *Proc Natl Acad Sci U S A*, 96(22):12524–9, (1999).

[249] S. S. Harwig, A. Waring, H. J. Yang, Y. Cho, L. Tan, and R. I. Lehrer. "Intramolecular disulfide bonds enhance the antimicrobial and lytic activities of protegrins at physiological sodium chloride concentrations". *Eur J Biochem*, 240(2):352–7, (1996).

[250] M. E. Mangoni, A. Aumelas, P. Charnet, C. Roumestand, L. Chiche, E. Despaux, G. Grassy, B. Calas, and A. Chavanieu. "Change in membrane permeability induced by protegrin 1: implication of disulphide bridges for pore formation". *FEBS Lett*, 383(1-2):93–8, (1996).

[251] H. Mohanram and S. Bhattacharjya. "Cysteine deleted protegrin-1 (CDP-1): antibacterial activity, outer-membrane disruption and selectivity". *Biochim Biophys Acta*, 1840(10):3006–16, (2014) http://dx.doi.org/10.1016/j.bbagen.2014.06.018.

[252] R. Mani, A. J. Waring, R. I. Lehrer, and M. Hong. "Membrane-disruptive abilities of beta-hairpin antimicrobial peptides correlate with conformation and activity: a 31P and 1H NMR study". *Biochim Biophys Acta*, 1716(1):11–8, (2005) http://dx.doi.org/10.1016/j.bbamem.2005.08.008.

[253] S. P. Mielke and V. V. Krishnan. "Characterization of protein secondary structure from NMR chemical shifts". *Prog Nucl Magn Reson Spectrosc*, 54(3-4):141–165, (2009) http://dx.doi.org/10.1016/j.pnmrns.2008.06.002.

[254] M. Tang and M. Hong. "Structure and mechanism of beta-hairpin antimicrobial peptides in lipid bilayers from solid-state NMR spectroscopy". *Mol Biosyst*, 5(4):317–22, (2009) http://dx.doi.org/10.1039/b820398a.

[255] C. M. Santiveri, E. León, M. Rico, and M. A. Jiménez. "Context-dependence of the contribution of disulfide bonds to beta-hairpin stability". *Chemistry*, 14(2):488–99, (2008) http://dx.doi.org/10.1002/chem.200700845.

[256] H. Jang, B. Ma, R. Lal, and R. Nussinov. "Models of toxic beta-sheet channels of protegrin-1 suggest a common subunit organization motif shared with toxic alzheimer beta-amyloid ion channels". *Biophys J*, 95(10):4631–42, (2008) http://dx.doi.org/10.1529/biophysj.108.134551.

[257] C. P. Hill, J. Yee, M. E. Selsted, and D. Eisenberg. "Crystal structure of defensin HNP-3, an amphiphilic dimer: mechanisms of membrane permeabilization". *Science*, 251(5000):1481–5, (1991).

[258] Z. O. Shenkarev, S. V. Balandin, K. I. Trunov, A. S. Paramonov, S. V. Sukhanov, L. I. Barsukov, A. S. Arseniev, and T. V. Ovchinnikova. "Molecular mechanism of action of beta-hairpin antimicrobial peptide arenicin: oligomeric structure in dodecylphosphocholine micelles and pore formation in planar lipid bilayers". *Biochemistry*, 50(28):6255–65, (2011) http://dx.doi.org/10.1021/bi200746t.

[259] T. J. Silhavy, D. Kahne, and S. Walker. "The bacterial cell envelope". *Cold Spring Harb Perspect Biol*, 2(5):a000414, (2010) http://dx.doi.org/10.1101/cshperspect.a000414.

[260] S. Razin. "Cholesterol incorporation into bacterial membranes". *J Bacteriol*, 124(1):570–2, (1975).

[261] J. P. Saenz, D. Grosser, A. S. Bradley, T. J. Lagny, O. Lavrynenko, M. Broda, and K. Simons. "Hopanoids as functional analogues of cholesterol in bacterial membranes". *Proc Natl Acad Sci U S A*, 112(38):11971–6, (2015) http://dx.doi.org/10.1073/pnas.1515607112.

[262] C. L. Schoch, G. H. Sung, F. López-Giráldez, J. P. Townsend, J. Miadlikowska, V. Hofstetter, B. Robbertse, P. B. Matheny, F. Kauff, Z. Wang, C. Gueidan, R. M. Andrie, K. Trippe, L. M. Ciufetti, A. Wynns, E. Fraker, B. P. Hodkinson, G. Bonito, J. Z. Groenewald, M. Arzanlou, G. S. de Hoog, P. W. Crous, D. Hewitt, D. H. Pfister, K. Peterson, M. Gryzenhout, M. J. Wingfield, A. Aptroot, S. O. Suh, M. Blackwell, D. M. Hillis, G. W. Griffith, L. A. Castlebury, A. Y. Rossman, H. T. Lumbsch, R. Lucking, B. Budel, A. Rauhut, P. Diederich, D. Ertz, D. M. Geiser, K. Hosaka, P. Inderbitzin, J. Kohlmeyer, B. Volkmann-Kohlmeyer, L. Mostert, K. O'Donnell, H. Sipman, J. D. Rogers, R. A. Shoemaker, J. Sugiyama, R. C. Summerbell, W. Untereiner, P. R. Johnston, S. Stenroos, A. Zuccaro, P. S. Dyer, P. D. Crittenden, M. S. Cole, K. Hansen, J. M. Trappe, R. Yahr, F. Lutzoni, and J. W. Spatafora. "The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits". *Syst Biol*, 58(2):224–39, (2009) http://dx.doi.org/10.1093/sysbio/syp020.

[263] T. Y. James, F. Kauff, C. L. Schoch, P. B. Matheny, V. Hofstetter, C. J. Cox, G. Celio, C. Gueidan, E. Fraker, J. Miadlikowska, H. T. Lumbsch, A. Rauhut, V. Reeb, A. E. Arnold, A. Amtoft, J. E. Stajich, K. Hosaka, G. H. Sung, D. Johnson, B. O'Rourke, M. Crockett, M. Binder, J. M. Curtis, J. C. Slot, Z. Wang, A. W. Wilson, A. Schussler, J. E. Longcore, K. O'Donnell, S. Mozley-Standridge, D. Porter, P. M. Letcher, M. J. Powell, J. W. Taylor, M. M. White, G. W. Griffith, D. R. Davies, R. A. Humber, J. B. Morton, J. Sugiyama, A. Y. Rossman, J. D. Rogers, D. H. Pfister, D. Hewitt, K. Hansen, S. Hambleton, R. A. Shoemaker, J. Kohlmeyer, B. Volkmann-Kohlmeyer, R. A. Spotts, M. Serdani, P. W. Crous, K. W. Hughes, K. Matsuura, E. Langer, G. Langer, W. A. Untereiner, R. Lucking, B. Budel, D. M. Geiser, A. Aptroot, P. Diederich, I. Schmitt, M. Schultz, R. Yahr, D. S. Hibbett, F. Lutzoni, D. J. McLaughlin, J. W. Spatafora, and R. Vilgalys. "Reconstructing the early evolution of fungi using a six-gene phylogeny". *Nature*, 443(7113):818–22, (2006) http://dx.doi.org/10.1038/nature05110.

[264] M. L. Berbee and J. W. Taylor. "Two ascomycete classes based on fruiting-body characters and ribosomal DNA sequence". *Mol Biol Evol*, 9(2):278–84, (1992) http://dx.doi.org/10.1093/oxfordjournals.molbev.a040719.

[265] A. C. Leeder, J. Palma-Guerrero, and N. L. Glass. "The social network: deciphering fungal language". *Nat Rev Microbiol*, 9(6):440–51, (2011) http://dx.doi.org/10.1038/nrmicro2580.

[266] A. de Bary. "Vergleichende morphologie und biologie der pilze, mycetozoen, und bacterie". *Wilhelm Engelmann*, (1884).

[267] D. Turrà and A. Di Pietro. "Chemotropic sensing in fungus-plant interactions". *Curr Opin Plant Biol*, 26:135–40, (2015) http://dx.doi.org/10.1016/j.pbi.2015.07.004.

[268] G. A. Zentmyer. "Chemotaxis of zoospores for root exudates". *Science*, 133(3464):1595–6, (1961) http://dx.doi.org/10.1126/science.133.3464.1595.

[269] R. A. Arkowitz. "Chemical gradients and chemotropism in yeast". *Cold Spring Harb Perspect Biol*, 1(2):a001958, (2009) http://dx.doi.org/10.1101/cshperspect.a001958.

[270] D. Turrà, M. El Ghalid, F. Rossi, and A. Di Pietro. "Fungal pathogen uses sex pheromone receptor for chemotropic sensing of host plant signals". *Nature*, 527(7579):521–4, (2015) http://dx.doi.org/10.1038/nature15516.

[271] H. F. Link. "Observationes in ordines plantarum naturals". *Mag Ges Naturf Freunde*, 3:3–42, (1809).

[272] H. W. Wollenweber and O. A. Reinking. "Die Fusarien, ihre beschreibung, schadwirkung und bekampfung.". *P Parey, Berlin*, page 365 pp., (1935).

[273] W. C. Snyder and H. N. Hansen. "The species concept in *Fusarium*". *Amer J Bot*, 27:64–67, (1940).

[274] P. E. Nelson, T. A. Toussoun, and W. F. O. Marasas. *"Fusarium species*: an illustrated manual for identification". *Pennsylvania State University Press, University Park*, (1983).

[275] P. E. Nelson, M. C. Dignani, and E. J. Anaissie. "Taxonomy, biology, and clinical aspects of *Fusarium species*". *Clin Microbiol Rev*, 7(4):479–504, (1994).

[276] A. Rodríguez, F. Perestelo, A. Carnicero, V. Regalado, R. Perez, G. De la Fuente, and M. A. Falcon. "Degradation of natural lignins and lignocellulosic substrates by soil-inhabiting fungi imperfecti". *FEMS Microbiology Ecology*, 21(3):213–219, (1996) http://dx.doi.org/10.1016/S0168-6496(96)00057-8.

[277] J. B. Sutherland, A. L. Pometto III, and D. L. Crawford. "Lignocellulose degradation by *Fusarium* species". *Can J Bot*, 61(4):1194–1198, (1983) http://dx.doi.org/10.1139/b83-126.

[278] P. Lemanceau, P. A. Bakker, W. J. De Kogel, C. Alabouvette, and B. Schippers. "Antagonistic effect of nonpathogenic *Fusarium oxysporum* Fo47 and pseudobactin 358 upon pathogenic *Fusarium oxysporum F. sp. dianthi*". *Appl Environ Microbiol*, 59(1):74–82, (1993).

[279] B. Nel, C. Steinberg, N. Labuschagne, and A. Viljoen. "The potential of nonpathogenic *Fusarium oxysporum* and other biological control organisms for suppressing *Fusarium* wilt of banana". *Plant Pathology*, 55(2):217–223, (2006) http://dx.doi.org/10.1111/j.1365-3059.2006.01344.x.

[280] V. Hornillos, E. Carrillo, L. Rivas, F. Amat-Guerri, and A. U. Acuña. "Synthesis of BODIPY-labeled alkylphosphocholines with leishmanicidal activity, as fluorescent analogues of miltefosine". *Bioorg Med Chem Lett*, 18(24):6336–6339, (2008) http://dx.doi.org/10.1016/j.bmcl.2008.10.089.

[281] G. P. Munkvold. *"Fusarium* species and their associated mycotoxins". *Methods Mol Biol*, 1542:51–106, (2017) http://dx.doi.org/10.1007/978-1-4939-6707-0_4.

[282] F. Wu and G. P. Munkvold. "Mycotoxins in ethanol co-products: modeling economic impacts on the livestock industry and management strategies". *J Agric Food Chem*, 56(11):3900–11, (2008) http://dx.doi.org/10.1021/jf072697e.

[283] J. M. Awika. "Major cereal grains production and use around the world". *Advances in Cereal Science: Implications to Food Processing and Health Promotion*, 1089(1089):1–13, (2011) http://dx.doi.org/doi:10.1021/bk-2011-1089.ch00110.1021/bk-2011-1089.ch001.

[284] A. Waśkiewicz, M. Beszterda, and P. Goliński. "Occurrence of fumonisins in food - An interdisciplinary approach to the problem". *Food Control*, 26(2):491–499, (2012) http://dx.doi.org/10.1016/j.foodcont.2012.02.007.

[285] D. M. Geiser, T. Aoki, C. W. Bacon, S. E. Baker, M. K. Bhattacharyya, M. E. Brandt, D. W. Brown, L. W. Burgess, S. Chulze, J. J. Coleman, J. C. Correll, S. F. Covert, P. W. Crous, C. A. Cuomo, G. S. De Hoog, A. Di Pietro, W. H. Elmer, L. Epstein, R. J. Frandsen, S. Freeman, T. Gagkaeva, A. E. Glenn, T. R. Gordon, N. F. Gregory, K. E. Hammond-Kosack, L. E. Hanson, M. Jiménez-Gasco Mdel, S. Kang, H. C. Kistler, G. A. Kuldau, J. F. Leslie, A. Logrieco, G. Lu, E. Lysoe, L. J. Ma, S. P. McCormick, Q. Migheli, A. Moretti, F. Munaut, K. O'Donnell, L. Pfenning, R. C. Ploetz, R. H. Proctor, S. A. Rehner, V. A. Robert, A. P. Rooney, B. Bin Salleh, M. M. Scandiani, J. Scauflaire, D. P. Short, E. Steenkamp, H. Suga, B. A. Summerell, D. A. Sutton, U. Thrane, F. Trail, A. Van Diepeningen, H. D. Vanetten, A. Viljoen, C. Waalwijk, T. J. Ward, M. J. Wingfield, J. R. Xu, X. B. Yang, T. Yli-Mattila, and N. Zhang. "One fungus, one name: defining the genus *Fusarium* in a scientifically robust way that preserves longstanding use". *Phytopathology*, 103(5):400–8, (2013) http://dx.doi.org/10.1094/PHYTO-07-12-0150-LE.

[286] K. O'Donnell, A. P. Rooney, R. H. Proctor, D. W. Brown, S. P. McCormick, T. J. Ward, R. J. Frandsen, E. Lysoe, S. A. Rehner, T. Aoki, V. A. Robert, P. W. Crous, J. Z. Groenewald, S. Kang, and D. M. Geiser. "Phylogenetic analyses of RPB1 and RPB2 support a middle Cretaceous origin for a clade comprising all agriculturally and medically important fusaria". *Fungal Genet Biol*, 52:20–31, (2013) http://dx.doi.org/10.1016/j.fgb.2012.12.004.

[287] K. O'Donnell, D. A. Sutton, M. G. Rinaldi, C. Gueidan, P. W. Crous, and D. M. Geiser. "Novel multilocus sequence typing scheme reveals high genetic diversity of human pathogenic members of the *Fusarium incarnatum F. equiseti* and *F. chlamydosporum* species complexes within the United States". *J Clin Microbiol*, 47(12):3851–61, (2009) http://dx.doi.org/10.1128/JCM.01616-09.

[288] R. C. Ploetz. " *Fusarium* wilt of banana is caused by several pathogens referred to as *Fusarium oxysporum F. sp. cubense* ". *Phytopathology*, 96(6):653–656, (2006) http://dx.doi.org/10.1094/PHYTO-96-0653.

[289] F. Wu. "Measuring the economic impacts of *Fusarium* toxins in animal feeds". *Animal Feed Science and Technology*, 137(3):363–374, (2007) http://dx.doi.org/10.1016/j.anifeedsci.2007.06.010.

[290] C. Lass-Florl and M. Cuenca-Estrella. "Changes in the epidemiological landscape of invasive mould infections and disease". *J Antimicrob Chemother*, 72(suppl 1i):5–11, (2017) http://dx.doi.org/10.1093/jac/dkx028.

[291] A. M. Tortorano, M. Richardson, E. Roilides, A. van Diepeningen, M. Caira, P. Munoz, E. Johnson, J. Meletiadis, Z. D. Pana, M. Lackner, P. Verweij, T. Freiberger, O. A.

Cornely, S. Arikan-Akdagli, E. Dannaoui, A. H. Groll, K. Lagrou, A. Chakrabarti, F. Lanternier, L. Pagano, A. Skiada, M. Akova, M. C. Arendrup, T. Boekhout, A. Chowdhary, M. Cuenca-Estrella, J. Guinea, J. Guarro, S. de Hoog, W. Hope, S. Kathuria, O. Lortholary, J. F. Meis, A. J. Ullmann, G. Petrikkos, C. Lass-Florl, M. European Society of Clinical, G. Infectious Diseases Fungal Infection Study, and M. European Confederation of Medical. "ESCMID and ECMM joint guidelines on diagnosis and management of hyalohyphomycosis: *Fusarium* spp., *Scedosporium* spp. and others". *Clin Microbiol Infect*, 20 Suppl 3:27–46, (2014) http://dx.doi.org/10.1111/1469-0691.12465.

[292] M. Cuenca-Estrella, M. Bassetti, C. Lass-Florl, Z. Racil, M. Richardson, and T. R. Rogers. "Detection and investigation of invasive mould disease". *J Antimicrob Chemother*, 66 Suppl 1:15–24, (2011) http://dx.doi.org/10.1093/jac/dkq438.

[293] M. Garnica, M. O. da Cunha, R. Portugal, A. Maiolino, A. L. Colombo, and M. Nucci. "Risk factors for invasive fusariosis in patients with acute myeloid leukemia and in hematopoietic cell transplant recipients". *Clin Infect Dis*, 60(6):875–80, (2015) http://dx.doi.org/10.1093/cid/ciu947.

[294] M. Nucci and E. Anaissie. *"Fusarium* infections in immunocompromised patients". *Clin Microbiol Rev*, 20(4):695–704, (2007) http://dx.doi.org/10.1128/CMR.00014-07.

[295] S. Shoham and K. A. Marr. "Invasive fungal infections in solid organ transplant recipients". *Future Microbiol*, 7(5):639–55, (2012) http://dx.doi.org/10.2217/fmb.12.28.

[296] A. Alastruey-Izquierdo, M. Cuenca-Estrella, A. Monzon, E. Mellado, and J. L. Rodriguez-Tudela. "Antifungal susceptibility profile of clinical *Fusarium* spp. isolates identified by molecular methods". *J Antimicrob Chemother*, 61(4):805–9, (2008) http://dx.doi.org/10.1093/jac/dkn022.

[297] A. Alastruey-Izquierdo, E. Mellado, T. Pelaez, J. Peman, S. Zapico, M. Alvarez, J. L. Rodriguez-Tudela, M. Cuenca-Estrella, and F. S. Group. "Population-based survey of filamentous fungi and antifungal resistance in Spain (FILPOP Study)". *Antimicrob Agents Chemother*, 57(7):3380–7, (2013) http://dx.doi.org/10.1128/AAC.00383-13.

[298] J. Guarro, A. S. Kantarcioglu, R. Horre, J. L. Rodriguez-Tudela, M. Cuenca Estrella, J. Berenguer, and G. S. de Hoog. *"Scedosporium apiospermum*: changing clinical spectrum of a therapy-refractory opportunist". *Med Mycol*, 44(4):295–327, (2006) http://dx.doi.org/10.1080/13693780600752507.

[299] E. Bucking-Throm, W. Duntze, L. H. Hartwell, and T. R. Manney. "Reversible arrest of haploid yeast cells in the initiation of DNA synthesis by a diffusible sex factor". *Exp Cell Res*, 76(1):99–110, (1973).

[300] D. Turrà, D. Nordzieke, S. Vitale, M. El Ghalid, and A. Di Pietro. "Hyphal chemotropism in fungal pathogenicity". *Semin Cell Dev Biol*, 57:69–75, (2016) http://dx.doi.org/10.1016/j.semcdb.2016.04.020.

[301] J. Kurjan and I. Herskowitz. "Structure of a yeast pheromone gene (MF alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor". *Cell*, 30(3):933–43, (1982).

[302] K. Schrick, B. Garvik, and L. H. Hartwell. "Mating in *Saccharomyces cerevisiae*: the role of the pheromone signal transduction pathway in the chemotropic response to pheromone". *Genetics*, 147(1):19–32, (1997).

[303] J. E. Segall. "Polarization of yeast cells in spatial gradients of alpha mating factor". *Proc Natl Acad Sci U S A*, 90(18):8332–6, (1993).

[304] A. Dmochowska, D. Dignard, D. Henning, D. Y. Thomas, and H. Bussey. "Yeast KEX1 gene encodes a putative protease with a carboxypeptidase B-like function involved in killer toxin and alpha-factor precursor processing". *Cell*, 50(4):573–84, (1987).

[305] Jr. Jones, S. K. and R. J. Bennett. "Fungal mating pheromones: choreographing the dating game". *Fungal Genet Biol*, 48(7):668–76, (2011) http://dx.doi.org/10.1016/j.fgb.2011.04.001.

[306] D. Julius, A. Brake, L. Blair, R. Kunisawa, and J. Thorner. "Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-alpha-factor". *Cell*, 37(3):1075–89, (1984).

[307] V. L. Boyartchuk and J. Rine. "Roles of prenyl protein proteases in maturation of *Saccharomyces cerevisiae* a-factor". *Genetics*, 150(1):95–101, (1998).

[308] B. He, P. Chen, S. Y. Chen, K. L. Vancura, S. Michaelis, and S. Powers. "RAM2, an essential gene of yeast, and RAM1 encode the two polypeptide components of the farnesyltransferase that prenylates a-factor and RAS proteins". *Proc Natl Acad Sci U S A*, 88(24):11373–7, (1991).

[309] S. Powers, S. Michaelis, D. Broek, S. Santa Anna, J. Field, I. Herskowitz, and M. Wigler. "RAM, a gene of yeast required for a functional modification of RAS proteins and for production of mating pheromone a-factor". *Cell*, 47(3):413–22, (1986).

[310] A. Tam, F. J. Nouvet, K. Fujimura-Kamada, H. Slunt, S. S. Sisodia, and S. Michaelis. "Dual roles for Ste24p in yeast a-factor maturation: NH2-terminal proteolysis and COOH-terminal CAAX processing". *J Cell Biol*, 142(3):635–49, (1998).

[311] M. G. Abel, Y. L. Zhang, H. F. Lu, F. Naider, and J. M. Becker. "Structure-function analysis of the *Saccharomyces cerevisiae* tridecapeptide pheromone using alanine-scanned analogs". *J Pept Res*, 52(2):95–106, (1998).

[312] F. Naider and J. M. Becker. "The alpha-factor mating pheromone of *Saccharomyces cerevisiae*: a model for studying the interaction of peptide hormones and G protein-coupled receptors". *Peptides*, 25(9):1441–63, (2004) http://dx.doi.org/10.1016/j.peptides.2003.11.028.

[313] P. Shenbagamurthi, B. Kundu, S. Raths, J. M. Becker, and F. Naider. "Biological activity and conformational isomerism in position 9 analogues of the des-1-tryptophan,3-beta-cyclohexylalanine-alpha-factor from *Saccharomyces cerevisiae*". *Biochemistry*, 24(25):7070–6, (1985).

[314] J. L. Patton and R. L. Lester. "The phosphoinositol sphingolipids of *Saccharomyces cerevisiae* are highly localized in the plasma membrane". *J Bacteriol*, 173(10):3101–8, (1991).

[315] M. E. van der Rest, A. H. Kamminga, A. Nakano, Y. Anraku, B. Poolman, and W. N. Konings. "The plasma membrane of *Saccharomyces cerevisiae*: structure, function, and biogenesis". *Microbiol Rev*, 59(2):304–22, (1995).

[316] A. A. Moosavi-Movahedi, M. Gharanfoli, K. Nazari, M. Shamsipur, J. Chamani, B. Hemmateenejad, M. Alavi, A. Shokrollahi, M. Habibi-Rezaei, C. Sorenson, and N. Sheibani. "A distinct intermediate of RNase A is induced by sodium dodecyl sulfate at its pK(a)". *Colloids Surf B Biointerfaces*, 43(3-4):150–7, (2005) http://dx.doi.org/10.1016/j.colsurfb.2005.04.008.

[317] R. Amiri, A. K. Bordbar, D. V. Laurents, A. R. Khosropour, and I. Mohammadpoor-Baltork. "Thermal stability and enzymatic activity of RNase A in the presence of cationic gemini surfactants". *Int J Biol Macromol*, 50(4):1151–7, (2012) http://dx.doi.org/10.1016/j.ijbiomac.2012.01.027.

[318] M. J. Rosen and D. J. Tracy. "Gemini surfactants". *J Surfactants Deterg*, 1(4):547–554, (1998).

[319] E. Pambou, J. Crewe, M. Yaseen, F. N. Padia, S. Rogers, D. Wang, H. Xu, and J. R. Lu. "Structural features of micelles of zwitterionic dodecyl-phosphocholine (C(1)(2)PC) surfactants studied by small-angle neutron scattering". *Langmuir*, 31(36):9781–9, (2015) http://dx.doi.org/10.1021/acs.langmuir.5b02077.

[320] I. Castrillo, N. A. Araujo, J. Alegre-Cebollada, J. G. Gavilanes, A. Martínez-del Pozo, and M. Bruix. "Specific interactions of sticholysin I with model membranes: an NMR study". *Proteins*, 78(8):1959–70, (2010) http://dx.doi.org/10.1002/prot.22712.

[321] M. F. García-Mayoral, M. Moussaoui, B. G. de la Torre, D. Andreu, E. Boix, M. V. Nogues, M. Rico, D. V. Laurents, and M. Bruix. "NMR structural determinants of eosinophil cationic protein binding to membrane and heparin mimetics". *Biophys J*, 98(11):2702–11, (2010) http://dx.doi.org/10.1016/j.bpj.2010.02.039.

[322] D. Pulido, M. F. García-Mayoral, M. Moussaoui, D. Velázquez, M. Torrent, M. Bruix, and E. Boix. "Structural basis for endotoxin neutralization by the eosinophil cationic protein". *FEBS J*, 283(22):4176–4191, (2016) http://dx.doi.org/10.1111/febs.13915.

[323] S. J. Perkins and K. Wüthrich. "Ring current effects in the conformation dependent NMR chemical shifts of aliphatic protons in the basic pancreatic trypsin inhibitor". *Biochim Biophys Acta*, 576(2):409–23, (1979).

[324] C. Grathwohl and K. Wüthrich. "NMR studies of the molecular conformations in the linear oligopeptides H-(L-Ala)n-L-Pro-OH". *Biopolymers*, 15(10):2043–57, (1976) http://dx.doi.org/10.1002/bip.1976.360151013.

[325] N. J. Zondlo. "Aromatic-proline interactions: electronically tunable CH/pi interactions". *Acc Chem Res*, 46(4):1039–49, (2013) http://dx.doi.org/10.1021/ar300087y.

[326] S. Rothemund, M. Beyermann, E. Krause, G. Krause, M. Bienert, R. S. Hodges, B. D. Sykes, and F. D. Sonnichsen. "Structure effects of double D-amino acid replacements: a nuclear magnetic resonance and circular dichroism study using amphipathic model helices". *Biochemistry*, 34(40):12954–62, (1995).

[327] E. Strandberg and J. A. Killian. "Snorkeling of lysine side chains in transmembrane helices: how easy can it get?". *FEBS Lett*, 544(1-3):69–73, (2003).

[328] B. C. Cunningham and J. A. Wells. "High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis". *Science*, 244(4908):1081–5, (1989).

[329] P. Grieco, V. Luca, L. Auriemma, A. Carotenuto, M. R. Saviello, P. Campiglia, D. Barra, E. Novellino, and M. L. Mangoni. "Alanine scanning analysis and structure-function relationships of the frog-skin antimicrobial peptide temporin-1Ta". *J Pept Sci*, 17(5):358–65, (2011) http://dx.doi.org/10.1002/psc.1350.

[330] S. H. Martin, B. D. Wingfield, M. J. Wingfield, and E. T. Steenkamp. "Causes and consequences of variability in peptide mating pheromones of ascomycete fungi". *Mol Biol Evol*, 28(7):1987–2003, (2011) http://dx.doi.org/10.1093/molbev/msr022.

[331] A. Domsalla and M. F. Melzig. "Occurrence and properties of proteases in plant latices". *Planta Med*, 74(7):699–711, (2008) http://dx.doi.org/10.1055/s-2008-1074530.

[332] S. B. Rodriguez, B. L. Stitt, and D. E. Ash. "Cysteine 351 is an essential nucleophile in catalysis by *Porphyromonas gingivalis* peptidylarginine deiminase". *Arch Biochem Biophys*, 504(2):190–6, (2010) http://dx.doi.org/10.1016/j.abb.2010.09.008.

[333] A. L. Levonen, A. Landar, A. Ramachandran, E. K. Ceaser, D. A. Dickinson, G. Zanoni, J. D. Morrow, and V. M. Darley-Usmar. "Cellular mechanisms of redox cell signalling: role of cysteine modification in controlling antioxidant defences in response to electrophilic lipid oxidation products". *Biochem J*, 378(Pt 2):373–82, (2004) http://dx.doi.org/10.1042/BJ20031049.

[334] L. J. McDonald and J. Moss. "Enzymatic and nonenzymatic ADP-ribosylation of cysteine". *Mol Cell Biochem*, 138(1-2):221–6, (1994).

[335] S. Vitale, A. Partida-Hanon, S. Serrano, A. Martínez-Del-Pozo, A. Di Pietro, D. Turrà, and M. Bruix. "Structure-activity relationship of alpha mating pheromone from the fungal pathogen *Fusarium oxysporum*". *J Biol Chem*, 292(9):3591–3602, (2017) http://dx.doi.org/10.1074/jbc.M116.766311.

[336] L. Marina Robles, C. Millán-Pacheco, N. Pastor, and G. Del Río. "Structure-function studies of the alpha pheromone receptor from yeast.". *TIP*, 20(1):16–26, (2017) http://dx.doi.org/10.1016/j.recqb.2016.11.002.

[337] D. J. Schibli, R. C. Montelaro, and H. J. Vogel. "The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined helix in dodecylphosphocholine micelles". *Biochemistry*, 40(32):9570–8, (2001).

Appendices

A

Structural study of the centrosomal TACC3-XMAP protein-protein interaction from a minimized system of peptides and cosolvents

A.1 Chemical shift tables for peptide TACC-α2

Table A.1. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide TACC-α2 in H_2O/D_2O (9:1 v/v), 5 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Val	**823**	8.31	4.03	2.04	0.93	-	-	-	62.63
Glu	**824**	8.74	4.28	2.02, 2.03	2.27	-	-	-	56.33
Met	**825**	8.59	4.45	2.10	2.55, 2.64	-	2.08	-	55.60
Gly	**826**	8.61	3.91	-	-	-	-	-	45.24
Lys	**827**	8.14	4.30	1.73, 1.79	1.34, 1.42	1.65	2.96	-	56.12
Ile	**828**	8.35	4.11	1.84	1.19, 1.50, 0.93	0.85	-	-	60.90
Ile	**829**	8.47	4.11	1.84	1.17, 1.47, 0.89	0.84	-	-	60.78
Ala	**830**	8.52	4.27	1.35	-	-	-	-	52.43
Glu	**831**	8.44	4.17	1.88, 1.90	2.21	-	-	-	56.69
Phe	**832**	8.42	4.61	3.08, 3.11	-	7.25	7.32	7.28	57.75
Glu	**833**	8.50	4.22	1.89, 2.02	2.21	-	-	-	56.85
Gly	**834**	8.03	3.93	-	-	-	-	-	45.21
Thr	**835**	8.19	4.34	4.22	1.21	-	-	-	62.06
Ile	**836**	8.45	4.18	1.89	1.20, 1.50, 0.90	0.85	-	-	61.63
Thr	**837**	8.30	4.21	4.13	1.19	-	-	-	62.78
Gln	**838**	8.43	4.25	2.00, 2.06	2.33, 2.40	6.97, 7.72		-	55.72
Ile	**839**	8.38	4.03	1.85	1.52, 1.16, 0.89	0.84	-	-	62.35
Leu	**840**	8.42	4.33	1.68	1.56	0.85, 0.91	-	-	55.78
Glu	**841**	8.54	4.19	1.98, 2.03	2.27, 2.34	-	-	-	56.90
Asp	**842**	8.60	4.54	2.70	-	-	-	-	55.49
Ser	**843**	8.44	4.31	3.95, 3.99	-	-	-	-	60.01
Gln	**844**	8.33	4.22	2.12, 2.15	2.39, 2.42	6.97, 7.73	-	-	56.42
Arg	**845**	8.13	4.20	1.86, 1.89	1.70, 1.63	3.23	7.37	6.70, 6.57	56.76
Gln	**846**	8.39	4.20	2.13	2.41, 2.43	6.96, 7.65	-	-	56.76
Lys	**847**	8.28	4.18	1.79, 1.87	1.41, 1.48	1.67	2.96	-	57.18
Glu	**848**	8.36	4.25	1.96, 2.08	2.25, 2.31	-	-	-	57.33
Thr	**849**	8.36	4.22	4.22	1.24	-	-	-	63.36
Ala	**850**	8.31	4.24	1.42	-	-	-	-	53.27
Lys	**851**	8.19	4.26	1.73, 1.81	-	1.68	2.98	-	57.31
Leu	**852**	8.18	4.20	1.57	1.50	0.87, 0.93	-	-	55.30
Glu	**853**	8.37	4.23	1.97, 2.05	2.26, 2.32	-	-	-	57.07
Leu	**854**	8.26	4.26	1.67	1.58	0.87, 0.93	-	-	55.47
Asn	**855**	8.49	4.64	2.78, 2.87	7.03, 7.75		-	-	53.37
Lys	**856**	8.29	4.23	1.78, 1.89	1.49	1.68	2.98	-	57.40

Table A.2. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide TACC-α2 in 70 % H_2O/D_2O (9:1 v/v), 30 % TFE, 25 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Val	**823**	8.01	4.00	2.07	0.98, 1.01	-	-	-	62.18
Glu	**824**	8.63	4.33	2.052.15	2.36	-	-	-	56.14
Met	**825**	8.26	4.35	2.10	2.37, 2.60	-	-	-	56.23
Gly	**826**	8.58	3.89	-	-	-	-	-	45.87
Lys	**827**	7.66	4.17	1.73, 1.99	1.47	1.58	2.99	-	57.78
Ile	**828**	7.57	3.80	2.12	1.21, 1.71, 0.91	0.90	-	-	64.05
Ile	**829**	8.25	3.78	1.93	1.22, 1.75, 0.99	0.86	-	-	63.60
Ala	**830**	7.96	4.24	1.57	-	-	-	-	54.02
Glu	**831**	8.19	4.17	2.18, 2.28	2.36, 2.45	-	-	-	56.59
Phe	**832**	8.66	4.32	3.19, 3.22	-	7.09	7.17	7.13	60.01
Glu	**833**	9.03	3.84	2.11, 2.25	2.29, 2.68	-	-	-	58.46
Gly	**834**	8.57	3.84	-	-	-	-	-	45.36
Thr	**835**	8.02	4.04	4.36	1.24	-	-	-	62.18
Ile	**836**	8.28	3.72	1.76	1.04, 1.76, 0.80	0.64	-	-	62.75
Thr	**837**	8.04	3.84	4.30	1.26	-	-	-	66.08
Gln	**838**	7.59	4.07	2.29	2.41, 2.53	6.73, 7.49		-	57.43
Ile	**839**	8.09	3.80	1.89	1.25, 0.83, 1.10	-	-	-	63.46
Leu	**840**	8.62	4.08	1.89, 2.01	1.52	0.87, 0.97	-	-	59.88
Glu	**841**	8.57	4.07	2.01, 2.24	2.31, 2.59	-	-	-	57.72
Asp	**842**	8.59	4.52	2.75, 2.93	-	-	-	-	56.18
Ser	**843**	8.43	4.26	4.07, 4.10	-	-	-	-	60.58
Gln	**844**	8.09	4.17	2.27, 2.27	2.48	6.73, 7.45	-	-	57.78
Arg	**845**	8.24	4.15	1.85, 2.05	1.68	3.23	7.41	-	58.03
Gln	**846**	8.34	4.12	2.24, 2.24	2.43, 2.58	6.68, 7.40	-	-	57.82
Lys	**847**	7.99	4.11	2.00, 2.00	1.68	-	-	-	57.99
Glu	**848**	8.34	4.17	2.18, 2.18	2.45	-	-	-	56.76
Thr	**849**	8.16	4.05	4.34, 4.34	1.30	-	-	-	65.63
Ala	**850**	8.02	4.19	1.52, 1.52	-	-	-	-	53.49
Lys	**851**	7.89	4.06	1.74, 1.96	1.44	1.62	-	-	57.53
Leu	**852**	7.85	4.20	1.80, 1.92	1.67	0.93, 0.97	-	-	58.39
Glu	**853**	8.13	4.18	1.92, 2.13	2.35, 2.39	-	-	-	57.58
Leu	**854**	8.07	4.23	1.79, 2.03	1.63	0.92	-	-	57.77
Asn	**855**	8.03	4.68	2.84, 2.89	6.85, 7.62	-	-	-	-
Lys	**856**	7.92	4.26	1.88, 1.95	1.49, 1.57	1.73	-	-	55.30

Table A.3. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide TACC-α2 in H_2O/D_2O (9:1 v/v), 20 mM DPC, 25 ºC, pH 5.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Val	**823**	8.37	4.00	2.06	0.92, 0.98	-	-	-	62.78
Glu	**824**	8.58	4.45	2.04, 2.15	2.30, 2.34	-	-	-	55.96
Met	**825**	8.92	4.22	2.04, 2.16	2.58	-	1.96	-	54.98
Gly	**826**	8.99	3.79	-	-	-	-	-	-
Lys	**827**	7.61	4.20	1.72, 1.97	1.42	1.55	2.97	-	57.51
Ile	**828**	7.72	3.83	1.55	1.24, 1.42, 0.84	0.90	-	-	64.39
Ile	**829**	8.35	3.68	1.97	1.13, 1.42, 0.95	0.85	-	-	64.14
Ala	**830**	8.04	4.21	1.43	-	-	-	-	53.21
Glu	**831**	8.18	4.21	2.09	2.34, 2.45	-	-	-	57.80
Phe	**832**	8.53	4.42	3.21, 3.27	-	7.22	7.17	7.09	60.51
Glu	**833**	8.86	3.81	2.14	2.29, 2.67	-	-	-	-
Gly	**834**	8.56	3.79	-	-	-	-	-	-
Thr	**835**	8.00	3.88	4.06	1.23	-	-	-	64.54
Ile	**836**	8.10	3.68	1.89	0.84, 1.09, 1.42	0.68	-	-	64.11
Thr	**837**	8.22	3.81	4.29	1.25	-	-	-	67.07
Gln	**838**	7.70	4.07	2.24	2.40, 2.47	6.84, 7.63	-	-	58.68
Ile	**839**	7.95	3.87	1.98	1.19, 1.87, 0.93	0.84	-	-	64.43
Leu	**840**	8.22	4.08	1.86, 1.90	1.55	0.87, 0.92	-	-	60.01
Glu	**841**	8.35	4.10	2.10, 2.16	2.34, 2.49	-	-	-	58.49
Asp	**842**	8.47	4.53	2.76, 2.86	-	-	-	-	55.97
Ser	**843**	8.18	4.29	3.82, 4.01	-	-	-	-	60.67
Gln	**844**	8.00	4.21	2.15, 2.22	2.44	6.87, 7.54	-	-	58.54
Arg	**845**	8.03	4.21	1.76, 1.93	1.68	3.24	7.41	-	56.64
Gln	**846**	8.24	4.22	2.08, 2.17	2.44	6.92, 7.59	-	-	58.51
Lys	**847**	8.05	4.21	-	-	-	2.97	-	56.83
Glu	**848**	8.29	4.25	2.08	2.36	-	-	-	57.25
Thr	**849**	8.17	4.16	4.24	1.24	-	-	-	63.31
Ala	**850**	8.17	4.25	1.43	-	-	-	-	53.18
Lys	**851**	8.13	4.21	-	-	-	2.98	-	57.06
Leu	**852**	7.94	4.21	1.78	1.55	0.94	-	-	54.53
Glu	**853**	8.32	4.23	1.89	-	-	-	-	57.14
Leu	**854**	8.04	4.26	1.68, 1.73	1.58	0.87, 0.92	-	-	58.02
Asn	**855**	8.36	4.65	2.78, 2.86	6.93, 7.64	-	-	-	53.18
Lys	**856**	8.15	4.25	1.77, 1.97	1.58	1.67	2.98	-	56.09

A.2 Chemical shift tables for peptide Ct-TACC-α4 wt

Table A.4. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Ct-TACC-α4 wt in H_2O/D_2O (9:1 v/v), 5 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Met	**897**	8.45	4.40	1.97	-	-	-	55.30
Lys	**898**	8.57	4.37	1.66, 1.75	-	-	-	55.63
Ile	**899**	8.37	4.15	1.88	-	0.93	-	61.85
Gln	**900**	8.68	4.38	2.02, 2.12	2.40	6.99, 7.69		55.46
Ser	**901**	8.53	4.46	3.89, 3.94	-	-	-	58.00
Leu	**902**	8.52	4.37	1.68	-	0.90, 0.95	-	55.98
Glu	**903**	8.51	4.23	1.99, 2.06	2.31	-	-	57.01
Arg	**904**	8.42	4.37	1.81, 1.89	1.68	3.23	7.34	58.16
Ser	**905**	8.52	4.43	3.94, 4.02	-	-	-	58.98
Leu	**906**	8.49	4.31	1.64, 1.73	-	-	-	56.45
Glu	**907**	8.42	4.23	1.94, 2.00	-	-	-	57.07
Gln	**908**	8.38	4.27	2.05	2.41	6.99, 7.72		56.03
Lys	**909**	8.45	4.31	1.84, 1.78	-	-	-	56.94
Ser	**910**	8.45	4.42	3.94, 3.97	-	-	-	58.77
Lys	**911**	8.40	4.33	1.79, 1.86	-	-	-	56.76
Glu	**912**	8.46	4.24	2.32, 2.26	-	-	-	56.84
Asn	**913**	8.50	4.70	2.90, 2.83	7.06, 7.78	-	-	53.75
Asp	**914**	8.43	4.59	2.74	-	-	-	55.08
Glu	**915**	8.38	4.22	1.98, 2.02	-	-	-	57.31
Leu	**916**	8.25	4.35	2.72	-	-	-	56.20
Thr	**917**	8.13	4.02	4.24	1.24	-	-	62.84
Lys	**918**	8.21	4.34	1.66, 1.75	-	-	-	56.40
Ile	**919**	8.26	4.15	1.86	-	-	-	61.59
Ser	**920**	8.36	4.43	3.90, 3.98	-	-	-	58.86
Asp	**921**	8.61	4.57	2.72	-	-	-	55.37
Asp	**922**	8.39	4.53	2.69	-	-	-	55.18
Leu	**923**	8.09	4.25	1.78	1.66	0.90, 0.97	-	56.62
Ile	**924**	7.99	4.01	1.95	1.23	0.92	-	62.13
Leu	**925**	8.16	4.29	1.56, 1.69	-	-	-	55.91
Lys	**926**	8.20	4.26	1.83, 1.89	-	1.71	-	56.19
Met	**927**	8.30	4.45	2.11	2.59, 2.69	-	-	55.57
Glu	**928**	8.43	4.27	2.24, 2.28	-	-	-	56.37
Lys	**929**	8.33	4.34	1.76, 1.83	-	-	-	56.32
Ile	**930**	8.23	4.13	1.88	0.96, 1.24	0.90	-	60.91

Table A.5. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Ct-TACC-α4 wt in 70 % H_2O/D_2O (9:1 v/v), 30 % TFE, 25 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Met	**897**	8.19	4.44	2.09, 2.07	2.62	-	-	55.02
Lys	**898**	8.34	4.38	1.93, 1.87	1.56	1.77	3.03	55.96
Ile	**899**	7.97	4.10	1.94	0.99, 1.29, 1.56	-	-	61.17
Gln	**900**	8.36	4.31	2.17	2.47	6.76, 7.48	-	56.01
Ser	**901**	8.10	4.40	4.09, 3.99	-	-	-	59.02
Leu	**902**	8.11	4.26	1.85, 1.76	1.69	0.94, 0.98	-	56.71
Glu	**903**	8.61	4.00	2.35, 2.15	2.45	-	-	58.50
Arg	**904**	8.08	4.26	2.00, 1.84	1.79	3.28	7.38	59.32
Ser	**905**	8.12	4.38	4.17, 4.00	-	-	-	60.41
Leu	**906**	8.41	4.19	1.93, 1.83	1.59	0.90, 0.91	-	57.57
Glu	**907**	8.25	4.09	2.22, 2.17	2.47	-	-	58.10
Gln	**908**	8.09	4.15	2.26	2.51	6.75, 7.67	-	57.49
Lys	**909**	8.25	4.24	2.00	1.61	1.73	3.02	57.25
Ser	**910**	8.29	4.34	4.18, 4.07	-	-	-	59.77
Lys	**911**	7.96	4.30	2.02	1.54	1.66	3.03	56.99
Glu	**912**	8.23	4.18	2.18, 1.92	2.34, 2.51	-	-	57.10
Asn	**913**	8.30	4.66	2.96, 2.93	6.80, 7.66	-	-	53.92
Asp	**914**	8.30	4.57	2.91, 2.82	-	-	-	55.28
Glu	**915**	8.34	4.18	2.16, 2.10	2.39	-	-	56.60
Leu	**916**	8.24	4.25	1.89, 1.81	1.72	0.94, 0.98	-	56.87
Thr	**917**	8.05	4.10	4.38	1.32	-	-	64.17
Lys	**918**	7.78	4.27	2.00, 1.75	1.51	1.62	3.03	56.99
Ile	**919**	7.98	4.07	2.00	0.99, 1.27, 1.45	0.93	-	62.57
Ser	**920**	8.02	4.33	4.11, 3.99	-	-	-	59.58
Asp	**921**	8.44	4.51	2.84, 2.74	-	-	-	55.39
Asp	**922**	8.05	4.50	2.80, 2.77	-	-	-	55.63
Leu	**923**	8.02	4.15	1.87, 1.79	-	0.94, 0.99	-	57.63
Ile	**924**	8.06	3.80	2.02	0.98, 1.28, 1.76	0.92	-	63.41
Leu	**925**	7.82	4.18	1.93, 1.81	1.68	0.95, 0.98	-	56.84
Lys	**926**	7.98	4.08	2.06	1.43	1.67	2.99	58.19
Met	**927**	8.42	4.27	2.31, 2.14	2.62, 2.79	-	-	57.40
Glu	**928**	8.28	4.17	2.37, 2.20	2.50, 2.51	-	-	57.74
Lys	**929**	7.90	4.28	2.00, 1.75	1.59	1.74	3.02	56.46
Ile	**930**	7.89	4.08	1.98	1.00, 1.27, 1.45	0.91	-	61.80

Table A.6. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Ct-TACC-α4 wt in H_2O/D_2O (9:1 v/v), 20 mM DPC, 25 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Met	**897**	8.50	4.37	2.04	2.55, 2.60	-	-	55.79
Lys	**898**	8.54	4.39	1.84, 1.70	1.43	1.50	2.99	56.01
Ile	**899**	8.46	3.98	1.93	1.25, 1.55, 0.93	0.89	-	62.68
Gln	**900**	8.54	4.20	2.08, 2.06	2.41	6.91, 7.65	-	58.11
Ser	**901**	8.12	4.34	3.98, 3.92	-	-	-	59.94
Leu	**902**	8.11	4.23	1.76	1.69	0.90, 0.93	-	56.93
Glu	**903**	8.40	3.98	2.10, 1.91	2.27, 2.43	-	-	57.08
Arg	**904**	8.18	4.22	1.91, 1.75	1.70	3.22	7.35	57.05
Ser	**905**	8.20	4.35	4.01, 3.91	-	-	-	59.74
Leu	**906**	8.24	4.17	-	1.59	0.900.91	-	57.31
Glu	**907**	8.15	4.22	-	-	-	-	57.15
Gln	**908**	8.26	4.07	2.11	2.41	6.89, 7.63	-	57.95
Lys	**909**	8.15	4.18	2.12	1.51	1.86	2.98	57.04
Ser	**910**	8.26	4.34	3.96, 3.92	-	-	-	59.84
Lys	**911**	8.22	4.13	1.87	1.51	1.62	2.97	58.09
Glu	**912**	8.17	4.25	2.02	2.42	-	-	56.79
Asn	**913**	8.20	4.72	2.88, 2.79	7.00, 7.75	-	-	-
Asp	**914**	8.47	4.57	2.76	-	-	-	55.29
Glu	**915**	8.45	4.16	2.11, 2.08	2.39	-	-	56.88
Leu	**916**	8.32	4.17	1.81, 1.74	1.66	0.89, 0.95	-	57.51
Thr	**917**	8.05	4.02	4.24	1.24	-	-	62.77
Lys	**918**	7.86	4.20	1.95, 1.90	1.45, 1.52	1.69	2.98	57.65
Ile	**919**	8.06	4.07	-	1.23	-	-	-
Ser	**920**	8.24	4.10	3.97, 3.89	-	-	-	57.83
Asp	**921**	8.43	4.43	2.82, 2.68	-	-	-	56.39
Asp	**922**	8.05	4.47	2.82, 2.69	-	-	-	56.56
Leu	**923**	8.24	4.25	1.87	1.51	0.90, 0.95	-	56.88
Ile	**924**	8.25	3.70	2.00	1.15, 1.78, 0.92	0.89	-	-
Leu	**925**	7.81	4.13	1.87, 1.74	1.65	0.91, 0.94	-	57.67
Lys	**926**	7.93	4.07	2.02, 1.87	-	1.49, 1.65	2.95	58.06
Met	**927**	8.29	4.19	2.11, 2.01	2.54, 2.70	-	2.25	56.67
Glu	**928**	8.21	4.04	2.22, 2.20	2.37, 2.62	-	-	-
Lys	**929**	7.73	4.21	1.95, 1.69	1.51	1.61	2.98	57.37
Ile	**930**	7.63	3.94	1.96	1.27, 1.69, 0.98	0.91	-	62.55

A.3 Chemical shift tables for peptide Ct-TACC-α4 mut

Table A.7. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Ct-TACC-α4 mut in H_2O/D_2O (9:1 v/v), 5 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε
Met	**897**	8.14	4.38	1.71	2.28	-	-
Lys	**898**	8.25	4.34	1.50, 1.45	1.38, 1.46	2.93	-
Ile	**899**	8.06	4.12	-	1.21, 1.76, 0.90	0.61	-
Gln	**900**	8.36	4.36	1.79, 1.69	2.39	6.99, 7.68	-
Ser	**901**	8.21	4.42	3.62, 3.55	-	-	-
Leu	**902**	8.21	4.34	-	-	0.88, 0.95	-
Glu	**903**	8.19	4.20	2.06, 1.99	-	-	-
Arg	**904**	8.11	4.34	1.56, 1.48	1.64, 1.68	3.20	7.34
Ser	**905**	8.21	4.40	3.71, 3.62	-	-	-
Leu	**906**	8.18	4.27	1.42	1.61	0.88, 0.93	-
Glu	**907**	8.07	4.18	2.23, 2.33	-	-	-
Gln	**908**	8.04	4.21	2.08	2.40	6.98, 7.73	-
Lys	**909**	8.12	4.27	1.78, 1.51	-	-	-
Ser	**910**	8.11	4.39	3.64, 3.61	-	-	-
Lys	**911**	8.04	4.29	-	-	-	-
Glu	**912**	8.10	4.20	2.06, 1.99	-	-	-
Asn	**913**	8.18	4.66	2.58, 2.52	7.05, 7.79	-	-
Asp	**914**	8.13	4.54	2.46, 2.43	-	-	-
Glu	**915**	8.06	4.15	1.78	-	-	-
Leu	**916**	7.90	4.27	1.51	1.64, 1.73	0.96, 0.89	-
Thr	**917**	7.85	4.17	3.97	1.24	-	-
Lys	**918**	7.77	4.23	-	1.42	1.52	3.00
Ile	**919**	7.85	4.03	-	0.95, 1.15, 1.92	0.88	-
Ser	**920**	8.03	4.34	3.68, 3.61	-	-	-
Ala	**921**	8.12	4.22	1.15	-	-	-
Ala	**922**	7.74	4.22	1.14	-	-	-
Leu	**923**	7.66	4.25	1.44	1.67	0.90	-
Ile	**924**	7.66	3.96	1.59	0.90, 1.15, 1.91	0.86	-
Leu	**925**	7.69	4.26	1.39, 1.34	1.61	0.89, 0.91	-
Lys	**926**	7.79	4.22	-	-	-	-
Met	**927**	7.99	4.42	1.79	2.59, 2.69	-	-
Glu	**928**	8.05	4.23	2.08, 1.99	-	-	-
Lys	**929**	8.00	4.31	1.57, 1.52	1.43	1.49	3.00
Ile	**930**	7.88	4.10	1.56	0.96, 1.33, 1.87	0.95	-

Table A.8. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Ct-TACC-α4 mut in 70 % H_2O/D_2O (9:1 v/v), 30 % TFE, 5 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Met	**897**	8.25	4.46	2.12, 2.10	2.64	-	-	55.16
Lys	**898**	8.39	4.38	1.91, 1.89	1.51, 1.59	1.78	3.06	56.17
Ile	**899**	7.98	4.10	1.99	1.00, 1.60, 1.33	0.96	-	61.47
Gln	**900**	8.37	4.29	2.19	2.50	6.79, 7.48	-	57.31
Ser	**901**	8.10	4.40	4.11, 4.03	-		-	59.31
Leu	**902**	8.09	4.27	1.89, 1.79	1.71	0.93, 0.98	-	57.21
Glu	**903**	8.60	4.01	2.39, 2.19	2.39, 2.52	-	-	58.66
Arg	**904**	8.13	4.23	2.03, 1.88	1.82	3.31	7.40	60.91
Ser	**905**	8.13	4.39	4.20, 4.03	-	-	-	60.68
Leu	**906**	8.46	4.20	1.97, 1.86	1.62	0.93	-	56.65
Glu	**907**	8.36	4.10	2.30, 2.22	2.36, 2.57	-	-	58.58
Gln	**908**	8.13	4.16	2.27, 1.97	2.56	6.77, 7.67	-	57.82
Lys	**909**	8.30	4.24	2.02	1.60	1.63	3.04	57.47
Ser	**910**	8.34	4.33	4.10, 4.11	-	-	-	60.08
Lys	**911**	8.02	4.31	2.07	-	1.70	-	57.54
Glu	**912**	8.30	4.17	2.28, 2.20	2.37, 2.53	-	-	57.39
Asn	**913**	8.39	4.63	2.98	6.78, 7.70	-	-	54.44
Asp	**914**	8.35	4.56	2.99, 2.86	-	-	-	55.92
Glu	**915**	8.33	4.17	2.28, 2.21	2.49	-	-	58.09
Leu	**916**	8.42	4.24	1.93, 1.84	1.84	0.96, 1.00	-	56.79
Thr	**917**	8.26	3.98	4.47	1.34	-	-	65.51
Lys	**918**	7.83	4.10	2.08, 1.98	1.48	1.74, 1.76	3.01	58.20
Ile	**919**	8.31	3.89	1.99	1.03, 1.23, 1.84	0.94	-	63.36
Ser	**920**	8.19	4.21	4.14, 3.97	-	-	-	61.44
Ala	**921**	8.10	4.13	1.55	-	-	-	53.73
Ala	**922**	7.88	4.13	1.60	-	-	-	53.87
Leu	**923**	8.22	4.16	1.99, 1.90	1.71	0.75	-	56.67
Ile	**924**	8.20	3.78	2.04	0.97, 1.23, 1.83	0.89	-	63.85
Leu	**925**	8.12	4.15	1.95	1.65	0.96	-	57.55
Lys	**926**	8.01	4.13	2.08	1.51	1.66	3.02	57.93
Met	**927**	8.42	4.30	2.35, 2.20	2.67, 2.81	-	-	55.98
Glu	**928**	8.40	4.21	2.43, 2.23	2.57	-	-	57.24
Lys	**929**	7.95	4.31	2.04, 1.78	1.58	1.62	-	56.73
Ile	**930**	7.90	4.13	2.01	1.02, 1.31, 1.66	0.97	-	60.93

Table A.9. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Ct-TACC-α4 mut in H_2O/D_2O (9:1 v/v), 20 mM DPC, 40 ºC, pH 5.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Met	**897**	8.53	4.44	2.13, 2.10	2.61, 2.66	-	2.08	55.02
Lys	**898**	8.56	4.48	1.94, 1.78	1.52	1.60	3.06	55.49
Ile	**899**	8.53	3.98	2.04	1.00, 1.35, 1.64	0.97	-	62.15
Gln	**900**	8.53	4.19	2.17	2.50	6.92, 7.65	-	57.57
Ser	**901**	7.96	4.36	4.10, 4.06	-	-	-	59.73
Leu	**902**	7.99	4.24	1.88, 1.78	-	0.97, 1.00	-	56.55
Glu	**903**	8.44	3.90	2.23, 2.02	2.33, 2.59	-	-	58.47
Arg	**904**	8.08	4.19	2.02, 1.87	1.78	3.30	7.37	57.55
Ser	**905**	8.09	4.38	4.10, 3.99	-	-	-	60.52
Leu	**906**	8.22	4.17	1.90	1.64	0.93, 0.99	-	56.83
Glu	**907**	8.27	4.07	2.24	2.38, 2.58	-	-	56.71
Gln	**908**	8.07	4.18	2.24	2.50, 2.56	6.88, 7.58	-	57.83
Lys	**909**	8.19	4.30	1.96, 1.77	1.50	1.62	3.02	55.93
Ser	**910**	8.18	4.35	4.10, 4.01	-	-	-	59.73
Lys	**911**	7.82	4.35	2.02, 1.94	1.62	1.77	3.06	55.78
Glu	**912**	7.98	4.34	2.20, 2.14	2.42, 2.50	-	-	55.80
Asn	**913**	8.11	4.81	2.97, 2.86	6.97, 7.79	-	-	52.51
Asp	**914**	8.57	4.64	2.89, 2.83	-	-	-	54.41
Glu	**915**	8.47	4.20	2.26, 2.13	2.51	-	-	56.67
Leu	**916**	8.43	4.17	1.89, 1.78	-	0.95, 1.00	-	56.74
Thr	**917**	8.17	4.03	4.37	1.35	-	-	61.44
Lys	**918**	7.89	4.18	2.04, 1.95	1.50	1.69	3.02	56.39
Ile	**919**	8.31	3.90	2.03	1.00, 1.26, 1.90	0.95	-	63.27
Ser	**920**	8.10	4.35	4.09, 4.02	-	-	-	58.94
Ala	**921**	8.10	4.13	1.59	-	-	-	53.91
Ala	**922**	7.72	4.23	1.62	-	-	-	53.50
Leu	**923**	8.28	4.09	1.92	1.74	0.94, 0.99	-	56.79
Ile	**924**	8.30	3.72	2.01	0.98, 1.23, 1.83	0.92	-	63.95
Leu	**925**	7.77	4.17	1.90, 1.81	1.72	0.92, 0.98	-	56.68
Lys	**926**	7.77	4.17	2.11, 2.00	1.62	1.78	2.98	56.54
Met	**927**	8.31	4.24	2.31, 2.18	2.60, 2.75	-	2.05	56.39
Glu	**928**	8.19	4.11	2.30, 2.20	2.51, 2.74	-	-	56.76
Lys	**929**	7.68	4.29	2.04, 1.77	1.60	1.70	3.06	55.91
Ile	**930**	7.62	4.02	2.03	1.05, 1.35, 1.76	0.98	-	61.34

A.4 Chemical shift tables for peptide XMAP-pCt

Table A.10. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide XMAP-pCt in H_2O/D_2O (9:1 v/v), 5 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Ser	**2047**	8.54	4.50	3.90, 3.86	-	-	-	58.12
Thr	**2048**	8.44	4.40	4.28	1.21	-	-	61.39
Asn	**2049**	8.59	4.73	2.84, 2.77	7.05, 7.76	-	-	53.05
Ile	**2050**	8.32	4.10	1.89	0.91, 1.20, 1.56	0.87	-	61.39
Asp	**2051**	8.44	4.49	2.72, 2.70	-	-	-	55.34
Asp	**2052**	8.45	4.58	2.73, 2.69	-	-	-	54.42
Lys	**2053**	8.31	4.09	1.88, 1.70	1.45	1.53	2.99	58.04
Leu	**2054**	8.07	4.22	1.74, 1.68	-	0.90, 0.96	-	55.97
Lys	**2055**	8.10	4.08	1.86, 1.70	1.40	1.51	-	57.90
Lys	**2056**	8.18	4.18	1.84, 1.68	1.45	1.54	-	57.14
Arg	**2057**	8.38	4.19	1.90, 1.72	1.60	3.22	7.33	56.10
Leu	**2058**	8.17	4.22	1.72	1.59	-	-	56.07
Glu	**2059**	8.27	4.18	2.05, 2.03	2.29, 2.38	-	-	57.33
Arg	**2060**	8.22	4.25	1.92, 1.86	1.64, 1.73	3.22	-	55.87
Ile	**2061**	8.18	4.01	1.90	0.92, 1.21, 1.59	0.87	-	62.06
Lys	**2062**	8.38	4.23	1.82	1.43	1.54	2.94	57.11
Ser	**2063**	8.41	4.43	3.96, 3.90	-	-	-	58.48
Ser	**2064**	8.36	4.47	3.96, 3.91	-	-	-	58.34
Arg	**2065**	8.28	4.34	1.89, 1.80	1.64, 1.70	3.22	-	55.95
Lys	**2066**	8.41	4.26	1.88, 1.69	1.45	1.51	-	56.96

Table A.11. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide XMAP-pCt in 70 % H_2O/D_2O (9:1 v/v), 30 % TFE, 25 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Ser	**2047**	8.18	4.54	3.97, 3.87	-	-	-	56.79
Thr	**2048**	8.19	4.42	4.35	1.27	-	-	60.45
Asn	**2049**	8.42	4.81	2.92, 2.85	7.01, 7.75	-	-	51.96
Ile	**2050**	8.07	4.13	1.95	0.98, 1.27, 1.55	0.93	-	60.81
Asp	**2051**	8.13	4.49	2.80, 2.73	-	-	-	54.26
Asp	**2052**	8.25	4.45	2.79, 2.73	-	-	-	55.53
Lys	**2053**	8.10	4.05	2.02, 1.91	1.51	1.68	3.04	57.98
Leu	**2054**	8.10	4.16	1.79, 1.73	1.54	0.92, 0.98	-	56.28
Lys	**2055**	8.18	3.92	1.97, 1.74	1.42	1.63	2.98	58.89
Lys	**2056**	7.91	4.12	1.98, 1.73	1.55	1.69	3.01	56.59
Arg	**2057**	7.83	4.11	2.08, 1.85	1.71	3.21	7.29	57.63
Leu	**2058**	8.48	4.13	1.97, 1.72	1.55	0.90, 0.97	-	56.50
Glu	**2059**	8.32	4.02	2.24, 2.15	2.36, 2.53	-	-	57.91
Arg	**2060**	7.89	4.15	2.11, 1.85	1.73	3.25	7.34	57.63
Ile	**2061**	8.24	3.77	2.04	0.94, 1.11, 1.87	0.89	-	63.91
Lys	**2062**	8.54	4.04	1.97, 1.72	1.47	1.65	2.96	58.10
Ser	**2063**	8.10	4.39	4.10, 4.05	-	-	-	59.04
Ser	**2064**	7.93	4.46	4.10, 4.06	-	-	-	58.46
Arg	**2065**	7.90	4.38	2.02, 1.83	1.75	3.22	7.26	54.99
Lys	**2066**	8.00	4.32	1.95, 1.88	1.52	1.58	3.21	54.77

Table A.12. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide XMAP-pCt in H_2O/D_2O (9:1 v/v), 20 mM DPC, 25 ºC, pH 5.5.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα
Ser	**2047**	8.40	4.47	3.89, 3.84	-	-	-	58.75
Thr	**2048**	8.29	4.36	4.26	1.20	-	-	62.11
Asn	**2049**	8.51	4.74	2.88, 2.79	6.98, 7.69	-	-	53.63
Ile	**2050**	8.32	4.05	1.93	0.92, 1.21, 1.50	0.88	-	
Asp	**2051**	8.43	4.44	2.71, 2.67	-	-	-	55.64
Asp	**2052**	8.22	4.42	2.71, 2.67	-	-	-	56.55
Lys	**2053**	8.16	4.04	1.85	1.45	1.57	2.98	59.22
Leu	**2054**	8.16	4.15	1.90, 1.81	1.69	0.89, 0.95	-	
Lys	**2055**	8.18	3.93	1.91	1.39	1.70	2.96	59.22
Lys	**2056**	8.05	4.13	1.86	1.47	1.64	2.97	58.16
Arg	**2057**	7.86	4.15	1.95, 1.76	1.70	3.18	-	57.72
Leu	**2058**	8.27	4.14	1.81, 1.78	1.63	0.88, 0.91	-	57.34
Glu	**2059**	8.21	4.05	2.12, 2.09	2.28, 2.40	-	-	58.80
Arg	**2060**	7.90	4.21	1.95, 1.77	1.68	3.21	-	58.13
Ile	**2061**	7.95	3.92	1.97	0.91, 1.20, 1.54	0.88	-	63.64
Lys	**2062**	8.26	4.10	1.92, 1.83	1.43	1.54	2.94	58.67
Ser	**2063**	8.22	4.38	3.97, 3.93	-	-	-	59.72
Ser	**2064**	8.12	4.42	3.97, 3.93	-	-	-	59.41
Arg	**2065**	8.00	4.33	1.94, 1.79	1.70	3.18	-	56.35
Lys	**2066**	8.16	4.24	1.86, 1.79	1.44, 1.49	1.68	2.99	56.43

A.5 Chemical shift table for peptide mixtures

Table A.13. Backbone ^{1}H chemical shifts (ppm from DSS) for peptides TACC-α2, Ct-TACC-α4 wt and XMAP-pCt in mixture, in 70 % H_2O/D_2O (9:1 v/v), 30 % TFE, 25 ºC, pH 5.5.

TACC-α2				Ct-TACC-α4 wt				XMAP-pCt			
Residue		HN	Hα	Residue		HN	Hα	Residue		HN	Hα
Val	**823**	8.00	3.98	**Met**	**897**	8.00	3.98	**Ser**	**2047**	8.21	4.54
Glu	**824**	8.62	4.31	**Lys**	**898**	8.62	4.31	**Thr**	**2048**	8.21	4.42
Met	**825**	8.25	4.34	**Ile**	**899**	8.25	4.34	**Asn**	**2049**	8.43	4.81
Gly	**826**	8.55	3.91	**Gln**	**900**	8.55	3.91	**Ile**	**2050**	8.06	4.13
Lys	**827**	7.65	4.16	**Ser**	**901**	7.65	4.16	**Asp**	**2051**	8.14	4.50
Ile	**828**	7.56	3.79	**Leu**	**902**	7.56	3.79	**Asp**	**2052**	8.26	4.45
Ile	**829**	8.23	3.76	**Glu**	**903**	8.23	3.76	**Lys**	**2053**	8.11	4.04
Ala	**830**	7.95	4.22	**Arg**	**904**	7.95	4.22	**Leu**	**2054**	8.10	4.18
Glu	**831**	8.18	4.16	**Ser**	**905**	8.18	4.16	**Lys**	**2055**	8.19	3.91
Phe	**832**	8.65	4.30	**Leu**	**906**	8.65	4.30	**Lys**	**2056**	7.92	4.12
Glu	**833**	9.06	3.88	**Glu**	**907**	9.06	3.88	**Arg**	**2057**	7.83	4.12
Gly	**834**	8.57	3.98	**Gln**	**908**	8.57	3.98	**Leu**	**2058**	8.49	4.12
Thr	**835**	8.00	4.02	**Lys**	**909**	8.00	4.02	**Glu**	**2059**	8.33	4.02
Ile	**836**	8.27	3.71	**Ser**	**910**	8.27	3.71	**Arg**	**2060**	7.90	4.15
Thr	**837**	8.03	3.83	**Lys**	**911**	8.03	3.83	**Ile**	**2061**	8.25	3.78
Gln	**838**	7.59	4.06	**Glu**	**912**	7.59	4.06	**Lys**	**2062**	8.55	4.04
Ile	**839**	8.07	3.79	**Asn**	**913**	8.07	3.79	**Ser**	**2063**	8.11	4.39
Leu	**840**	8.61	4.08	**Asp**	**914**	8.61	4.08	**Ser**	**2064**	7.94	4.46
Glu	**841**	8.57	4.07	**Glu**	**915**	8.57	4.07	**Arg**	**2065**	7.90	4.38
Asp	**842**	8.58	4.50	**Leu**	**916**	8.58	4.50	**Lys**	**2066**	8.01	4.32
Ser	**843**	8.42	4.24	**Thr**	**917**	8.42	4.24				
Gln	**844**	8.09	4.16	**Lys**	**918**	8.09	4.16				
Arg	**845**	8.23	4.13	**Ile**	**919**	8.23	4.13				
Gln	**846**	8.34	4.10	**Ser**	**920**	8.34	4.10				
Lys	**847**	7.98	4.09	**Asp**	**921**	7.98	4.09				
Glu	**848**	8.33	4.16	**Asp**	**922**	8.33	4.16				
Thr	**849**	8.15	4.03	**Leu**	**923**	8.15	4.03				
Ala	**850**	8.01	4.18	**Ile**	**924**	8.01	4.18				
Lys	**851**	7.88	4.05	**Leu**	**925**	7.88	4.05				
Leu	**852**	7.85	4.19	**Lys**	**926**	7.85	4.19				
Glu	**853**	8.12	4.16	**Met**	**927**	8.12	4.16				
Leu	**854**	-	4.23	**Glu**	**928**	-	4.23				
Asn	**855**	8.02	4.67	**Lys**	**929**	8.02	4.67				
Lys	**856**	7.92	4.24	**Ile**	**930**	7.92	4.24				

B

Structural analysis of protein-membrane interactions employing peptides derived from the HIV-1 ectodomain gp41 in cosolvents

B.1 Chemical shift tables for peptide MPERb

Table B.1. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide MPERb in 25 % HFIP, 35 ºC, pH 7.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα	Cβ
Lys	**n-4**	-	4.15	-	-	-	-	-	55.74	-
Lys	**n-3**	-	4.50	1.86	-	-	-	-	56.47	33.24
Lys	**n-2**	8.46	4.40	1.86, 1.92	1.58	1.81	-	-	56.60	33.54
Lys	**n-1**	8.31	4.37	1.84, 1.90	1.50	1.56	-	-	56.59	33.04
Asp	**664**	8.16	4.70	2.70, 2.73	-	-	-	-	54.19	41.6
Lys	**665**	8.33	4.20	1.78, 2.00	1.31	1.61	-	-	57.89	32.54
Trp	**666**	7.88	4.63	3.35, 3.41	-	7.30	9.72, 7.67	7.48	58.30	29.02
Ala	**667**	7.82	4.23	1.47	-	-	-	-	53.81	18.45
Ser	**668**	7.93	4.43	3.92, 3.99	-	-	-	-	59.82	63.5
Lys	**669**	7.76	4.29	1.58, 1.65	-	0.83, 0.90	-	-	57.13	42.16
Trp	**670**	7.78	4.61	3.19, 3.32	-	7.10	9.12, 7.61	7.42	58.55	28.96
Asn	**671**	8.04	4.63	2.75, 2.84	6.57, 7.29	-	-	-	55.79	38.7
Trp	**672**	7.87	4.46	3.35, 3.42	-	7.07	9.21, 7.52	7.43	59.80	28.82
Phe	**673**	8.12	4.14	3.03, 3.11	-	7.19	-	-	60.11	38.68
Asp	**674**	8.10	4.52	2.78, 2.86	-	-	-	-	55.69	39.05
Ile	**675**	8.02	3.98	1.92	1.68, 1.30, 1.03	0.90	-	-	64.41	38.47
Thr	**676**	7.99	3.98	4.07	1.06	-	-	-	65.59	68.99
Asn	**677**	7.86	4.62	2.68, 2.80	5.92, 7.05	-	-	-	54.89	38.69
Trp	**678**	8.04	4.72	3.41, 3.54	-	7.22	9.18, 7.62	7.39	59.81	29.16
Lys	**679**	8.20	4.05	1.86	1.50	0.79, 0.90	-	-	57.70	41.77
Trp	**680**	7.89	4.28	3.39	-	7.03	9.25, 7.39	7.44	60.36	28.82
Tyr	**681**	7.57	4.08	2.91, 3.02	-	7.13	6.85	-	61.23	37.91
Ile	**682**	8.02	3.83	1.85	1.16, 1.50, 0.92	0.79	-	-	64.20	37.59
Lys	**683**	7.83	3.94	1.84, 1.90	1.39	1.63	-	-	59.97	31.92
Lys	**684**	7.64	4.08	1.53, 1.61	1.76	0.90	-	-	58.05	41.65
Phe	**685**	8.18	4.19	3.13, 3.29	-	7.18	7.26	7.22	61.63	38.78
Ile	**686**	8.50	3.59	2.01	1.32, 1.93, 0.96	0.91	-	-	65.10	37.91
Met	**687**	8.05	4.17	2.26, 2.42	2.56, 2.81	-	1.86	-	58.86	32.35
Ile	**688**	8.47	3.76	2.01	1.16, 1.83, 0.90	-	-	-	65.24	37.92
Val	**689**	8.58	3.55	1.86	0.70, 0.82	-	-	-	66.73	31.39
Gly	**690**	8.28	3.81, 3.91	-	-	-	-	-	46.85	-
Lys	**n+1**	7.88	4.24	1.99, 2.09	1.57	1.65	-	-	57.89	32.4
Lys	**n+2**	8.24	4.21	1.99, 2.03	1.55	1.64	-	-	58.10	33.32
Lys	**n+3**	8.34	4.27	1.97	1.55	1.76	-	-	58.35	32.6
Lys	**n+4**	7.96	4.33	1.96, 2.00	1.57	1.61	-	-	56.92	32.45
Lys	**n+5**	8.09	4.34	1.92, 1.98	1.59	1.78	-	-	56.31	32.86

Table B.2. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide MPERb in 20 mM DPC, 35 ºC, pH 7.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ
Lys	**n-4**	-	4.01	-	-	-	-	-	-	55.58	-
Lys	**n-3**	8.38	4.35	1.76, 1.82	1.46	1.70	-	-	-	56.42	-
Lys	**n-2**	8.49	4.28	1.76, 1.81	1.40	1.45	-	-	-	56.39	-
Lys	**n-1**	8.42	4.27	1.74, 1.77	1.37	1.42	-	-	-	56.79	-
Asp	**664**	8.27	4.63	2.71	-	-	-	-	-	54.80	41.86
Lys	**665**	8.53	4.05	1.57, 1.64	1.06	1.50	-	-	-	57.55	32.62
Trp	**666**	8.16	4.57	3.31, 3.37	-	7.39	10.58, 7.51	7.49, 6.95	7.09	58.08	29.12
Ala	**667**	7.93	4.15	1.50	-	-	-	-	-	54.70	18.75
Ser	**668**	8.09	4.33	3.87, 3.94	-	-	-	-	-	60.56	63.02
Lys	**669**	7.88	4.05	1.52	1.61	0.69, 0.80	-	-	-	58.38	42.08
Trp	**670**	7.74	4.62	3.22, 3.41	-	7.29	10.29, 7.59	7.42, 6.90	7.03	58.76	29.52
Asn	**671**	7.98	4.63	2.82, 3.04	-	6.96, 7.36	-	-	-	53.97	38.86
Trp	**672**	7.90	4.43	3.26,	-	7.23	10.41, 7.36	7.42, 6.86	6.96	59.02	29.62
Phe	**673**	7.74	4.38	2.83, 3.04	-	7.11	-	-	-	57.95	-
Asp	**674**	8.13	4.71	2.79	-	-	-	-	-	58.15	-
Ile	**675**	8.16	3.74	1.77	0.87, 1.20, 1.38	0.78	-	-	-	62.83	38.22
Thr	**676**	8.13	3.95	4.10	1.17	-	-	-	-	66.04	68.62
Asn	**677**	8.07	4.74	2.79	-	-	-	-	-	54.61	38.19
Trp	**678**	8.06	4.70	3.35, 3.41	-	7.33	10.47, 7.40	7.43, 6.86	6.95	59.87	29.18
Lys	**679**	7.99	4.06	-	-	-	-	-	-	58.14	41.63
Trp	**680**	8.03	4.10	3.21, 3.31	-	7.06	10.24, 6.65	7.45, 6.79	7.10	58.88	29.45
Tyr	**681**	7.64	4.00	3.01, 3.18	-	7.22	6.88	-	-	61.91	-
Ile	**682**	8.09	3.77	2.13	0.98, 1.02, 1.25	-	-	-	-	65.29	37.78
Lys	**683**	8.30	3.85	1.59, 1.81	1.18	1.34	-	-	-	60.41	31.94
Lys	**684**	7.68	3.89	1.37	1.57	0.62, 0.74	-	-	-	57.97	41.97
Phe	**685**	8.26	4.15	3.23, 3.37	-	7.11	-	-	-	61.51	-
Ile	**686**	8.39	3.47	2.00	0.93, 1.02, 1.19	0.89	-	-	-	65.35	37.9
Met	**687**	7.99	4.09	2.21, 2.27	2.50, 2.73	-	2.05	-	-	58.12	32.44
Ile	**688**	8.22	3.73	1.95	0.98, 1.22	0.88	-	-	-	64.79	37.99
Val	**689**	8.18	3.57	1.92	0.65, 0.77	0.65	-	-	-	65.91	31.41
Gly	**690**	8.12	3.85	-	-	-	-	-	-	-	-
Lys	**n+1**	7.72	4.18	1.70, 1.93	1.50	1.57	2.96	-	-	57.73	31.46
Lys	**n+2**	7.97	4.23	1.69, 1.87	-	1.50	2.97	-	-	56.79	-
Lys	**n+3**	8.09	4.24	1.81, 1.86	1.49	1.68	-	-	-	56.51	-
Lys	**n+4**	8.19	4.25	1.78, 1.85	1.48	1.69	-	7.12	-	56.18	-
Lys	**n+5**	-	-	-	-	-	-	-	-	-	-

B.2 Chemical shift tables for peptide MPER-H2

Table B.3. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide MPER-H2 in 25 % HFIP, 35 ºC, pH 7.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ
Asn	**671**	-	3.71	2.59, 2.66	-	6.53, 7.24	-	-	-	53.72	41.51
Trp	**672**	-	4.46	3.20, 3.29	-	7.09	9.28, 7.45	7.45, 7.15	7.24	58.91	29.01
Phe	**673**	7.38	4.25	2.85	-	7.05	7.29	7.25	-	60.18	38.54
Asp	**674**	7.76	4.59	2.71, 2.76	-	-	-	-	-	55.50	40.82
Ile	**675**	7.80	3.98	1.92	1.30, 1.60, 1.01	0.92	-	-	-	64.07	38.52
Thr	**676**	7.98	4.04	4.21	1.21	-	-	-	-	65.26	69.02
Asn	**677**	7.83	-	2.72, 2.80	-	6.16, 7.19	-	-	-	55.34	38.60
Trp	**678**	7.96	-	3.38, 3.52	-	7.22	9.23, 7.58	7.39, 7.03	7.18	59.47	29.36
Lys	**679**	8.00	4.04	1.49, 1.77	1.76	0.87, 0.93	-	-	-	57.65	41.72
Trp	**680**	7.81	4.25	3.32, 3.39	-	7.04	9.21, 7.35	7.40, 7.08	7.22	58.94	28.68
Tyr	**681**	7.38	4.05	2.82, 2.97	-	7.07	6.82	-	-	61.04	37.88
Ile	**682**	7.90	3.82	1.88	1.17, 1.54, 0.93	0.82	-	-	-	64.24	37.70
Lys	**683**	7.84	3.92	1.84, 1.88	1.38, 1.60	1.59	2.85	-	-	60.05	32.05
Lys	**684**	7.64	4.06	1.57, 1.73	1.62	0.89	-	-	-	58.12	41.71
Phe	**685**	8.08	4.16	3.13, 3.29	-	7.15	7.23	7.18	-	61.63	38.77
Ile	**686**	8.46	3.56	2.01	1.30, 1.91, 0.93	0.89	-	-	-	65.08	37.78
Met	**687**	8.08	4.15	2.22, 2.41	2.57, 2.77	-	2.07	-	-	58.71	32.36
Ile	**688**	8.45	3.75	2.00	1.14, 1.81, 0.88	0.86	-	-	-	65.22	38.01
Val	**689**	8.52	3.52	1.84	0.68, 0.79	-	-	-	-	66.79	31.36
Gly	**690**	8.28	3.75, 3.87	-	1.50, 1.64	-	-	-	-	46.97	-
Lys	**n+1**	7.88	4.14	1.94, 2.07	1.49, 1.61	1.71	2.98	-	-	58.82	32.36
Lys	**n+2**	8.23	4.11	1.94, 2.02	1.51, 1.58	1.67	2.96	-	-	58.46	32.37
Lys	**n+3**	8.36	4.19	1.95	1.53, 1.59	1.70	-	-	-	57.59	32.52
Lys	**n+4**	7.90	4.25	1.97	-	1.75	3.02	-	-	57.17	32.63
Lys	**n+5**	-	4.29	-	-	1.75	-	-	-	56.42	32.83

Table B.4. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide MPER-H2 in 20 mM DPC, 35 ºC, pH 7.0.
* Ambiguous assignment, Lys residue not determined.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ
Asn	**671**	-	3.90	2.60	-	-	-	-	-	53.66	41.13
Trp	**672**	-	4.42	3.18	-	7.38	7.49	7.47, 6.86	7.05	58.86	29.48
Phe	**673**	7.63	4.49	2.77, 2.84	-	6.98	7.17	7.10	-	-	38.9
Asp	**674**	-	4.48	-	-	-	-	-	-	57.21	
Ile	**675**	8.25	3.82	1.90	1.24, 1.47, 0.85	0.86	-	-	-	63.09	38.34
Thr	**676**	8.20	3.93	4.17	1.23	-	-	-	-	66.11	68.48
Asn	**677**	8.12	4.49	2.73, 2.84	-	6.90, 7.75	-	-	-	57.10	-
Trp	**678**	8.00	-	3.39, 3.45	-	7.33	7.43	7.43, 6.87	7.02	-	-
Lys	**679**	8.17	4.08	1.66	1.84	0.93, 0.98	-	-	-	58.29	41.69
Trp	**680**	-	4.41	3.19	-	7.38	7.50	7.48, 6.96	7.06	58.88	-
Tyr	**681**	7.66	4.01	3.02, 3.19	-	6.84, 7.22	6.58, 6.74	-	-	61.91	-
Ile	**682**	8.14	3.75	2.11	1.20, 1.98, 0.99	0.97	-	-	-	65.38	37.77
Lys	**683**	8.32	3.83	1.66, 1.81	1.33, 1.59	1.58	2.84	-	-	-	31.9
Lys	**684**	7.68	3.88	1.16, 1.59	1.37	0.59, 0.73	-	-	-	-	41.33
Phe	**685**	8.27	4.09	3.20, 3.34	-	7.08	7.08	7.04	-	61.76	-
Ile	**686**	8.41	3.45	2.00	1.16, 1.97, 0.91	0.87	-	-	-	65.37	37.86
Met	**687**	8.00	4.07	2.19, 2.27	2.48, 2.73	-	2.03	-	-	58.92	32.48
Ile	**688**	8.18	3.68	1.92	1.13, 1.78, 0.85	0.82	-	-	-	64.97	37.94
Val	**689**	8.17	3.54	1.91	0.75, 0.64	-	-	-	-	66.09	31.42
Gly	**690**	8.15	3.69, 3.84	-	-	-	-	-	-	-	-
Lys	**n+1**	7.73	4.17	1.89, 1.94	1.53, 1.56	1.68	2.95	-	-	57.15	32.58
Lys	**n+2**	7.95	4.19	1.83, 1.92	1.49, 1.57	1.67	2.94	-	-	57.55	32.6
Lys	**n+3**	-	-	-	-	-	-	-	-	-	-
Lys	**n+4**	-	-	-	-	-	-	-	-	-	-
Lys	**n+5**	-	-	-	-	-	-	-	-	-	32.99
Lys	*	-	4.23	1.80, 1.86	1.47, 1.68	1.68	2.98	-	-	56.70	-

C

Structure and molecular association in sequences derived from the antimicrobial peptide protegrin-1 in the presence of membrane mimetics

C.1 Chemical shift tables for peptide PG-T4

Table C.1. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide PG-T4 in H_2O/D_2O (9:1 v/v), 5 ºC, pH 3.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα	Cβ
Arg	**1**	-	4.05	1.93	1.69	3.23	7.27	-	55.29	30.37
Gly	**2**	8.94	4.05	-	-	-	-	-	44.63	-
Gly	**3**	8.53	3.98	-	-	-	-	-	44.54	-
Arg	**4**	8.44	4.32	1.75, 1.82	1.62	3.19	7.23	-	55.86	30.60
Leu	**5**	8.57	4.40	1.47, 1.63	1.61	0.85, 0.92	-	-	54.82	42.00
Thr	**6**	8.24	4.30	4.13	1.14	-	-	-	61.37	69.75
Tyr	**7**	8.41	4.66	2.93, 3.03	-	7.09	6.78	-	57.47	38.95
Thr	**8**	8.25	4.25	4.08	1.16	-	-	-	61.41	69.66
Arg	**9**	8.47	4.27	1.75, 1.83	1.63	3.20	7.23	-	55.74	30.64
Arg	**10**	8.51	4.24	1.72	1.52, 1.59	3.15	7.17	-	55.86	30.77
Arg	**11**	8.52	4.27	1.68	1.48, 1.56	3.13	7.18	-	55.82	30.72
Phe	**12**	8.48	4.71	3.01, 3.13	-	7.24	7.33	7.28	57.34	39.66
Thr	**13**	8.26	4.33	4.09	1.16	-	-	-	61.38	69.82
Val	**14**	8.41	4.16	2.07	0.96	-	-	-	62.03	32.60
Thr	**15**	8.51	4.36	4.12	1.18	-	-	-	61.58	69.58
Val	**16**	8.44	4.12	2.07	0.95	-	-	-	62.24	32.67
Gly	**17**	8.66	3.95	-	-	-	-	-	44.78	-
Arg	**18**	8.01	4.20	1.71, 1.86	1.58	3.18	7.24	-	56.65	31.03

Table C.2. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide PG-T4 in 30 mM DPC H_2O/D_2O (9:1 v/v), 35 ºC, pH 3.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα	Cβ
Arg	**1**	-	4.09	1.94, 1.96	1.69	3.22	7.25	-	55.58	30.61
Gly	**2**	8.79	4.04	-	-	-	-	-	45.08	-
Gly	**3**	8.34	3.99	-	-	-	-	-	45.09	-
Arg	**4**	8.32	4.37	1.75, 1.80	1.61	3.17	7.25	-	56.31	31.29
Leu	**5**	8.42	4.43	1.53, 1.65	1.61	0.87, 0.92	-	-	55.10	42.62
Thr	**6**	8.03	4.38	4.07	1.11	-	-	-	61.78	70.12
Tyr	**7**	8.22	4.65	2.93, 3.00	-	7.05	6.77	-	-	39.67
Thr	**8**	8.00	4.27	4.09	1.14	-	-	-	-	70.12
Arg	**9**	8.31	4.27	1.74, 1.86	1.62	3.20	7.33	-	56.28	31.29
Arg	**10**	8.27	4.19	1.75, 1.80	1.59	3.16	7.24	-	56.31	30.55
Arg	**11**	8.16	4.23	1.69, 1.73	1.47	3.11	7.20	-	56.31	30.61
Phe	**12**	8.16	4.76	3.02, 3.15	-	7.23	7.28	7.26	57.61	40.30
Thr	**13**	8.15	4.38	4.09	1.13	-	-	-	-	70.05
Val	**14**	8.23	4.26	2.08	0.94	-	-	-	62.90	33.29
Thr	**15**	8.28	4.50	4.11	1.15	-	-	-	-	70.05
Val	**16**	8.15	4.21	2.08	0.93	-	-	-	61.59	33.31
Gly	**17**	8.40	3.96	-	-	-	-	-	45.13	-
Arg	**18**	7.85	4.23	1.70, 1.85	1.58	3.17	7.22	-	56.67	31.42

C.2 Chemical shift tables for peptide PG-C6C15

Table C.3. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide PG-C6C15 in H_2O/D_2O (9:1 v/v) 5 ºC, pH 3.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα	Cβ
Arg	**1**	-	4.10	1.95	1.70	3.23	7.29	-	55.47	30.57
Gly	**2**	8.96	4.06	-	-	-	-	-	44.88	-
Gly	**3**	8.50	3.99	-	-	-	-	-	44.77	-
Arg	**4**	8.37	4.49	1.69, 1.78	1.58	3.14	7.22	-	55.92	31.59
Leu	**5**	8.65	4.62	1.56, 1.63	1.57	0.84, 0.90	-	-	54.59	43.43
Cys	**6**	8.78	5.41	2.78, 2.95	-	-	-	-	56.01	-
Tyr	**7**	8.96	4.83	2.84, 2.97	-	6.80	6.64	-	56.74	-
Thr	**8**	8.74	4.74	3.95	1.07	-	-	-	61.44	70.09
Arg	**9**	8.97	4.30	1.78, 1.87	1.27, 1.59	3.20	7.27	-	56.05	31.55
Arg	**10**	8.96	3.83	1.96	1.60	3.23	7.29	-	57.46	28.22
Arg	**11**	8.13	4.05	1.81	1.46	3.14	7.17	-	56.51	31.63
Phe	**12**	8.28	4.93	3.00, 3.13	-	7.27	7.35	7.28	57.53	40.69
Thr	**13**	8.73	4.73	3.96	1.08	-	-	-	61.66	70.11
Val	**14**	8.68	4.19	1.74	0.91	-	-	-	61.76	31.53
Cys	**15**	8.85	5.28	2.86, 2.92	-	-	-	-	55.21	-
Val	**16**	8.83	4.34	2.18	1.01	-	-	-	61.44	34.10
Gly	**17**	8.68	3.98, 4.08	-	-	-	-	-	44.80	-
Arg	**18**	8.01	4.19	1.71, 1.84	1.58	3.18	7.23	-	57.42	31.63

Table C.4. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide PG-C6C15 in 30 mM DPC H_2O/D_2O (9:1 v/v), 35 ºC, pH 3.5.
* Ambiguous assignment, values can be exchanged

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα	Cβ
Arg	**1**	-	4.09	1.95	1.70	3.22	7.30	54.95	30.00
Gly	**2**	8.82	4.04	-	-	-	-	44.51	-
Gly	**3**	8.31	3.98	-	-	-	-	44.45	-
Arg	**4**	8.05	4.68	1.64, 1.72	1.53	3.07	7.13	57.51	-
Leu	**5**	8.67	4.62	1.63	1.59	0.90, 0.93	-	53.64	43.32
Cys	**6**	8.33	5.78	2.69, 2.91	-	-	-	-	-
Tyr	**7**	9.22	4.75	2.78, 2.85	-	7.02	6.73	56.85	-
Thr	**8**	7.91	5.32	3.66	0.98	-	-	60.26	70.84
Arg	**9**	8.72	4.37	1.70, 1.83	1.57, 1.64	3.35	7.72	54.37	-
Arg	**10**	9.19	3.81	2.06	1.70	3.26	7.60	57.28	27.57
Arg	**11**	7.95	4.03	1.86	1.33, 1.41	3.12	7.26	54.93	-
Phe	**12**	8.13	4.86	3.11	-	7.20	7.27	-	-
Thr	**13**	8.72	5.24	3.70	1.08	-	-	61.07	72.09
Val	**14**	8.87	5.86 *	-	1.04	-	-	-	34.40
Cys	**15**	8.34	5.86	2.70	-	-	-	-	-
Val	**16**	9.00	6.04 *	2.14	0.98	-	-	-	34.42
Gly	**17**	8.57	4.05, 4.18	-	-	-	-	44.56	-
Arg	**18**	7.96	4.22	1.69, 1.83	1.60	3.16	7.36	56.27	31.28

C.3 Chemical shift tables for peptide PG-C8C13

Table C.5. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide PG-C8C13 in H_2O/D_2O (9:1 v/v), 5 ºC, pH 3.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα	Cβ
Arg	**1**	-	4.10, 4.10	1.95	1.70	3.23	7.28	-	55.49	30.62
Gly	**2**	8.95	4.05	-	-	-	-	-	44.88	-
Gly	**3**	8.52	3.99	-	-	-	-	-	44.76	-
Arg	**4**	8.43	4.36	1.74, 1.81	1.61	3.19	7.22	-	56.01	30.99
Cys	**5**	8.58	4.44	1.49, 1.63	1.60	0.85, 0.92	-	-	54.98	42.61
Thr	**6**	8.22	4.49	4.11	1.14	-	-	-	61.43	70.26
Tyr	**7**	8.48	4.63	2.97	-	7.06	6.78	-	57.60	39.41
Thr	**8**	8.57	4.84	3.05	-	-	-	-	55.70	-
Arg	**9**	8.74	4.36	1.79, 1.98	1.64	3.21	7.25	-	56.06	30.49
Arg	**10**	8.11	4.28	1.82, 1.99	1.63	3.26	-	-	56.08	-
Arg	**11**	8.37	4.00	1.70, 1.75	1.30, 1.43	3.10	7.16	-	57.74	29.73
Phe	**12**	8.06	4.72	3.04, 3.14	-	7.34	7.23	7.33	57.35	39.67
Cys	**13**	8.53	4.88	2.96, 3.07	-	-	-	-	55.67	-
Val	**14**	8.41	4.20	2.07	0.95	-	-	-	61.76	33.56
Thr	**15**	8.53	4.51	4.11	1.18	-	-	-	61.76	69.87
Val	**16**	8.46	4.18	2.08	0.95	-	-	-	62.07	33.17
Gly	**17**	8.64	3.97	-	-	-	-	-	44.93	-
Arg	**18**	8.05	4.20	1.72, 1.86	1.59	3.18	7.24	-	57.27	31.44

Table C.6. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide PG-C8C13 in 30 mM DPC, H_2O/D_2O (9:1 v/v), 35 ºC, pH 3.5

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Cα	Cβ
Arg	**1**	-	4.09	1.70, 1.96	1.71	3.24	7.28	55.63	30.70
Gly	**2**	8.82	4.06	-	-		-	45.09	-
Gly	**3**	8.39	4.00	-	-		-	45.06	-
Arg	**4**	8.10	4.71	1.68, 1.72	1.52, 1.58	3.10	7.17	55.83	-
Cys	**5**	8.58	4.62	1.57, 1.69	1.61	0.91, 0.95	-	54.49	44.22
Thr	**6**	8.21	5.26	3.84	1.14		-	61.55	71.50
Tyr	**7**	9.01	4.78	2.86	-	7.03	6.74	57.51	42.16
Thr	**8**	8.36	5.63	2.84, 2.92	-		-	55.53	-
Arg	**9**	8.86	4.38	1.69, 1.91	1.60, 1.64	3.34	7.68	55.05	32.89
Arg	**10**	8.00	3.88	1.87	1.42	3.14	7.28	57.94	29.37
Arg	**11**	8.00	4.13	2.07	1.71	3.28	7.57	55.74	28.21
Phe	**12**	8.00	5.06	3.13	-	7.23	7.26	-	43.45
Cys	**13**	-	5.67	-	-		-	55.48	-
Val	**14**	8.47	5.02, 4.14	2.10	1.06		-	-	35.13
Thr	**15**	8.59	5.42	3.86	1.09		-	-	71.39
Val	**16**	8.90	4.76	2.13	0.99		-	-	34.43
Gly	**17**	8.46	4.07	-	-		-	45.06	-
Arg	**18**	7.91	4.22	1.68, 1.82	1.58	3.15	7.28	57.33	31.96

D

Structure and molecular interactions of the α-mating pheromone from *Fusarium oxysporum* in the presence of membrane mimetics and other cosolvents

D.1 Chemical shift tables for peptide α-pheromone

Table D.1. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *trans* α-pheromone in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	**1**	-	4.33	3.33, 3.39	-	7.24	10.27, 7.56	7.47	56.04
Cys	**2**	8.67	4.52	2.75	-	-	-	-	57.96
Thr	**3**	8.4	4.32	4.14	1.15	-	-	-	62.05
Trp	**4**	8.35	4.66	3.24	-	7.24	10.24, 7.56	7.44	57.38
Arg	**5**	8.21	4.05	1.45, 1.68	1.17, 1.23	2.95	7.04	-	56.05
Gly	**6**	7.67	3.62, 3.80	-	-	-	-	-	44.9
Gln	**7**	8.02	4.48	1.83, 2.02	2.27	6.92, 7.59	-	-	53.41
Pro	**8**	-	4.23	1.55, 1.95	1.94, 1.87	3.57, 3.73	-	-	63.22
Cys	**9**	8.16	4.39	2.75	-	-	-	-	58.05
Trp	**10**	7.77	4.51	3.13, 3.33	-	7.16	10.07, 7.62	7.42	58.49

Table D.2. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *cis* α-pheromone in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε
Trp	**1**	-	-	-	-	-	-
Cys	**2**	8.65	4.47	2.73	-	-	-
Thr	**3**	8.35	4.21	-	1.16	-	-
Trp	**4**	8.2	4.64	3.24	-	-	-
Arg	**5**	8.02	4.09	1.43, 1.67	1.16, 1.27	2.95	7.05
Gly	**6**	7.36	3.46, 3.53	-	-	-	-
Gln	**7**	7.93	4.48	1.80, 1.95	2.25	6.86, 7.55	-
Pro	**8**	-	4.57	1.77, 1.84	-	3.41, 3.57	-
Cys	**9**	8.32	4.36	2.77	-	-	-
Trp	**10**	7.8	4.46	-	-	-	-

Table D.3. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *trans* α-pheromone in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	**1**	-	4.38	3.37, 3.43	-	7.26	10.20, 7.62	7.50	56.13
Cys	**2**	8.68	4.62	2.81	-	-	-	-	57.98
Thr	**3**	8.32	4.53	4.19	1.19	-	-	-	61.63
Trp	**4**	8.34	4.75	3.28	-	7.26	10.18, 7.62	7.47	57.53
Arg	**5**	8.20	4.04	1.48, 1.74	1.08, 1.20	2.94	7.04	-	56.15
Gly	**6**	7.66	3.60, 3.94	-	-	-	-	-	44.97
Gln	**7**	7.83	4.58	1.91, 2.09	2.30	-	6.81, 7.55	-	53.25
Pro	**8**	-	4.32	1.63, 1.94	1.90, 1.99	3.62, 3.76	-	-	63.33
Cys	**9**	8.04	4.50	2.81	-	-	-	-	57.75
Trp	**10**	7.66	4.54	3.20, 3.37	-	7.18	9.94, 7.62	7.43	58.63

Table D.4. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *cis* α-pheromone in 30 % TFE, 5 ºC., pH 5.0

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε
Trp	**1**	-	4,35	3.28, 3.34	-	7.23	10.19
Cys	**2**	8.60	4.63	2.83	-	-	-
Thr	**3**	8.18	4.48	4.20	1.22	-	-
Trp	**4**	8.19	4.71	3.34	-	-	-
Arg	**5**	7.93	4.19	-	1.31	3.00	-
Gly	**6**	7.40	3.50, 3.71	-	-	-	-
Gln	**7**	7.82	-	-	-	-	-
Pro	**8**	-	4.31	1.77, 1.84	-	3.41, 3.57	-
Cys	**9**	8.00	4.50	-	-	-	-
Trp	**10**	-	-	-	-	-	-

Table D.5. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide reduced Gln-Pro *trans* α-pheromone in 20 mM DPC, 25 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ
Trp	**1**	-	4.32	3.42	-	7.39	10.64, 7.62	7.45, 7.10	7.03	53.97	27.05
Cys	**2**	8.60	4.42	2.70	-	-	-	-	-	55.98	25.64
Thr	**3**	8.25	4.38	4.10	1.12	-	-	-	-	59.22	67.28
Trp	**4**	8.18	4.69	3.23, 3.26	-	7.26	10.56, 7.57	7.43, 7.08	7.02	54.75	27.31
Arg	**5**	8.25	4.11	1.56, 1.76	1.24 1.31	3.02	7.30	-	-	53.50	28.83
Gly	**6**	7.79	3.88, 3.67	-	-	-	-	-	-	42.37	-
Gln	**7**	7.88	4.54	1.89, 2.06	2.29	6.82, 7.51	-	-	-	55.63	26.49
Pro	**8**	-	4.32	1.68, 1.99	1.87 1.95	3.60, 3.73	-	-	-	60.54	29.16
Cys	**9**	8.14	4.42	2.81	-	-	-	-	-	55.80	25.89
Trp	**10**	7.60	4.52	3.19, 3.33	-	7.22	10.46, 7.56	7.42, 7.05	6.98	55.65	27.28

Table D.6. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide oxidized Gln-Pro *trans* α-pheromone in 20 mM DPC, 25 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα
Trp	**1**	-	4.20	3.46, 3.50	-	7.47	10.87	7.53, 7.14	7.33	54.76
Cys	**2**	-	4.14	2.70, 2.99	-	-	-	-	-	-
Thr	**3**	7.79	4.69	3.76	0.96	-	-	-	-	68.96
Trp	**4**	9.00	-	3.01, 3.15	-	7.14	10.51	7.43, 7.04	7.08	-
Arg	**5**	8.94	3.60	1.53, 1.70	0.53, 0.88	2.80	7.23	-	-	54.08
Gly	**6**	8.55	3.53, 4.07	-	-	-	-	-	-	-
Gln	**7**	7.63	4.88	1.91, 2.16	2.30, 2.38	6.87, 7.53	-	-	-	53.97
Pro	**8**	-	4.54	1.87, 2.31	1.98, 2.07	3.62, 3.79	-	-	-	60.39
Cys	**9**	7.75	3.77	2.07, 2.98	-	-	-	-	-	65.80
Trp	**10**	7.24	4.38	3.15, 3.39	-	7.15	10.44	7.32, 6.92	7.00	55.54

Table D.7. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide reduced Gln-Pro *trans* α-pheromone in 2 mM Gemini, 25 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα
Trp	**1**		4.26	3.30, 3.36	-	7.28	10.29, 7.58	7.46, 7.16	7.09	56.18
Cys	**2**	8.38	4.49	2.70	-	-	-	-	-	-
Thr	**3**	8.20	4.31	4.13	1.12	-	-	-	-	-
Trp	**4**	8.21	4.67	3.21, 3.24	-	7.22	10.23, 7.56	7.44, 7.19	7.08	57.06
Arg	**5**	8.15	4.12	1.52, 1.73	1.29	3.01	-	-	-	56.22
Gly	**6**	7.74	3.66, 3.86		-	-	-	-	-	45.22
Gln	**7**	7.92	4.50	1.84, 2.01	2.27	6.80, 7.45	-	-	-	-
Pro	**8**		4.24	1.60, 1.86	1.92, 1.95	3.57, 3.71	-	-	-	63.14
Cys	**9**	8.01	4.39	2.74	-	-	-	-	-	56.11
Trp	**10**	7.60	4.48	3.14, 3.31	-	7.19	10.11, 7.59	7.42, 7.12	7.04	58.94

Table D.8. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide oxidized α-pheromone in 2 mM Gemini, 35 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ
Trp	**1**	-	4.16	3.31	-	7.32	10.35 7.58	7.51, 7.20	7.11	56.37	-
Cys	**2**	-	4.52	2.76	-	-	-	-	-	56.79	-
Thr	**3**	8.26	4.56	3.92	1.01	-	-	-	-	61.60	70.97
Trp	**4**	8.69	4.78	3.16, 3.18	-	7.21	10.21 7.64	7.47, 7.18	7.10	57.01	30.07
Arg	**5**	8.69	3.83	1.50, 1.76	0.92, 1.14	2.92	6.95	-	-	56.59	28.68
Gly	**6**	8.06	3.58, 4.07	-	-	-	-	-	-	45.37	-
Gln	**7**	7.59	4.78	1.93, 2.16	2.32	6.77, 7.43	-	-	-	52.87	29.69
Pro	**8**	-	4.43	1.75, 2.21	1.97, 2.06	3.62, 3.81	-	-	-	63.14	32.19
Cys	**9**	7.88	4.15	2.55, 3.03	-	-	-	-	-	56.65	-
Trp	**10**	7.41	4.46	3.17, 3.33	-	7.19	10.09 7.56	7.40, 7.07	7.03	58.38	30.14

D.2 Chemical shift tables for peptide scrambled

Table D.9. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Trp-Pro *trans* scrambled in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα
Trp	**1**	-	4.19	3.17, 3.24	-	6.91	9.82, 7.42	7.37, 6.85	7.05	55.98
Arg	**2**	8.17	4.21	1.47, 1.54	1.30	3.00	7.07	-	-	55.33
Trp	**3**	8.16	4.77	3.06, 3.34	-	7.25	10.22, 7.68	7.46, 7.18	-	55.06
Pro	**4**	-	4.35	1.93, 2.27	2.00, 2.03	3.64, 3.87	-	-	-	63.45
Cys	**5**	8.46	4.39	2.80	-	-	-	-	-	58.56
Cys	**6**	8.45	4.48	2.87	-	-	-	-	-	58.57
Trp	**7**	8.27	4.63	3.28	-	7.19	10.14, 7.57	7.45	-	57.68
Gly	**8**	8.24	3.68, 3.84	-	-	-	-	-	-	45.25
Gln	**9**	8.13	4.35	1.94, 2.10	2.28	6.90, 7.69	-	-	-	55.66
Thr	**10**	8.00	4.15	4.24	1.15	-	-	-	-	63.21

Table D.10. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Trp-Pro *cis* scrambled in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	**1**	-	4.32	3.31, 3.37	-	7.22	10.22, 7.54	7.48	55.64
Arg	**2**	8.56	4.39	1.70, 1.77	1.55	3.15	7.19	-	55.60
Trp	**3**	8.58	4.50	3.17, 3.22	-	7.30	10.34, 7.59	7.52	-
Pro	**4**	-	3.39	1.05, 1.66	1.45, 1.51	3.21	-	-	-
Cys	**5**	8.41	4.23	2.73	-	-	-	-	56.90
Cys	**6**	8.26	4.32	2.71	-	-	-	-	57.92
Trp	**7**	8.12	4.46	3.13, 3.20	-	7.16	10.15, 7.49	-	57.58
Gly	**8**	8.17	3.63, 3.80	-	-	-	-	-	-
Gln	**9**	8.12	-	1.92, 2.07	2.26	-	-	-	-
Thr	**10**	-	-	-	-	-	-	-	-

Table D.11. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Trp-Pro *trans* scrambled in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα
Trp	**1**	-	4.20	3.21, 3.27	-	6.98	9.53, 7.46	7.35, 6.88	7.07	56.26
Arg	**2**	8.07	4.28	1.53, 1.62	1.37	3.04	7.16	-	-	55.54
Trp	**3**	8.00	4.81	3.11, 3.34	-	7.26	10.06, 7.65	7.49, 7.19	-	55.05
Pro	**4**	-	4.32	1.99, 2.22	1.98	3.57, 3.84	-	-	-	63.63
Cys	**5**	8.16	4.44	2.84	-	-	-	-	-	58.41
Cys	**6**	8.29	4.52	2.92	-	-	-	-	-	58.58
Trp	**7**	8.07	4.66	3.29	-	7.16,	9.97, 7.54	7.44, 7.11	-	57.72
Gly	**8**	8.11	3.74, 3.93	-	-	-	-	-	-	45.19
Gln	**9**	8.08	4.40	2.00, 2.15	2.32	6.80, 7.68	-	-	-	55.83
Thr	**10**	7.90	4.22	4.28	1.19	-	-	-	-	63.08

Table D.12. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Trp-Pro *cis* scrambled in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	**1**	-	4.38	3.33, 3.38	-	7.23	10.14, 7.56	7.49	56.13
Arg	**2**	8.53	4.45	1.73, 1.83	1.61	3.19	7.25	-	55.56
Trp	**3**	8.46	4.51	3.18, 3.24	-	7.25	10.15, 7.57	7.51	56.21
Pro	**4**	-	3.38	0.84, 1.66	1.43, 1.51	3.25	-	-	63.01
Cys	**5**	8.19	4.2	2.77	-	-	-	-	58.55
Cys	**6**	8.04	4.29	2.66, 2.71	-	-	-	-	57.94
Trp	**7**	7.92	4.55	3.17, 3.26	-	7.14	10.01, 7.53	7.44	57.68
Gly	**8**	8.05	3.69, 3.91	-	-	-	-	-	-
Gln	**9**	-	-	-	-	-	-	-	-
Thr	**10**	-	-	-	-	-	-	-	-

Table D.13. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Trp-Pro *trans* scrambled in 20 mM DPC, 25 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ
Trp	**1**	-	4.30	3.26, 3.35	-	7.21	10.17, 7.32	7.24, 6.73	6.47	55.52	29.53
Arg	**2**	8.35	4.42	1.67, 1.79	1.52, 1.55	3.14	7.28	-	-	55.76	32.63
Trp	**3**	8.36	4.41	3.14, 3.36	-	7.42	10.74, 7.65	7.50, 7.12	7.08	55.77	29.60
Pro	**4**	-	4.40	1.90, 2.18	1.90	3.62, 3.97	-	-	-	63.11	31.99
Cys	**5**	8.40	4.42	2.90	-	-	-	-	-	58.52	27.85
Cys	**6**	8.28	4.47	2.89	-	-	-	-	-	58.53	28.03
Trp	**7**	8.14	4.71	3.21, 3.33	-	7.20	10.44, 7.47	7.38, 7.02	6.94	57.45	29.64
Gly	**8**	8.46	3.91, 3.98	-	-	-	-	-	-	45.31	-
Gln	**9**	8.19	4.41	2.01, 2.17	2.36	6.85, 7.64	-	-	-	54.44	29.89
Thr	**10**	7.88	4.16	4.24	1.17	-	-	-	-	63.15	70.51

Table D.14. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Trp-Pro *cis* scrambled in 2 mM Gemini, 25 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	**1**	-	4.40	-	-	7.25	10.21	7.46	-
Arg	**2**	-	4.52	1.70, 1.79	1.54	3.09	-	-	55.70
Trp	**3**	8.36	4.39	3.15, 3.20	-	7.30	10.32, 7.58	7.50	-
Pro	**4**	-	4.23	-	1.66	3.44	-	-	-
Cys	**5**	8.19	4.34	2.76	-	-	-	-	56.21
Cys	**6**	8.08	4.56	2.72	-	-	-	-	55.86
Trp	**7**	8.02	4.40	3.15, 3.25	-	7.21	10.16, 7.55	7.44	-
Gly	**8**	8.12	3.77, 3.86	-	-	-	-	-	-
Gln	**9**	8.08	4.37	1.96, 2.12	2.32	-	-	-	55.85
Thr	**10**	7.84	4.15	4.22	1.14	-	-	-	63.13

D.3 Chemical shift tables for peptide D-Ala[1,2]

Table D.15. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *trans* D-Ala[1,2] in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα
D-Ala	**1**	-	4.09	1.53	-	-	-	-	-	51.68
D-Ala	**2**	8.78	4.43	1.44	-	-	-	-	-	52.86
Thr	**3**	8.40	4.41	4.16	1.13	-	-	-	-	61.43
Trp	**4**	8.50	4.65	3.25	-	7.24	10.23, 7.61	7.47	7.19	57.61
Arg	**5**	8.37	4.09	1.50, 1.71	1.30	3.01	7.10	-	-	55.92
Gly	**6**	7.44	3.59, 3.75	-	-	-	-	-	-	44.89
Gln	**7**	8.11	4.51	1.86, 2.03	2.31	6.95, 7.62	-	-	-	53.55
Pro	**8**	-	4.25	1.64, 2.02	1.91, 1.98	3.60, 3.77	-	-	-	63.35
Cys	**9**	8.18	4.40	2.75, 2.81	-	-	-	-	-	58.25
Trp	**10**	7.76	4.52	3.15, 3.35	-	7.18	10.11, 7.66	7.46	-	58.52

Table D.16. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *cis* D-Ala[1,2] in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
D-Ala	**1**	-	-	-	-	-	-	-	-
D-Ala	**2**	-	-	-	-	-	-	-	-
Thr	**3**	-	4.34	4.21	-	-	-	-	61.63
Trp	**4**	8.43	4.63	3.27	-	-	-	-	-
Arg	**5**	8.21	4.12	1.46, 1.67	1.34	3.04	7.12	-	-
Gly	**6**	7.14	3.43, 3.48	-	-	-	-	-	-
Gln	**7**	7.95		1.82, 1.97	2.29	6.89, 7.56	-	-	-
Pro	**8**	-	4.60	1.81, 2.23	1.88	3.45, 3.60	-	-	-
Cys	**9**	8.37	4.39	2.83, 3.02	-	-	-	-	-
Trp	**10**	7.82	4.46	3.13	-	7.16	10.07	7.44	-

Table D.17. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *trans* D-Ala[1,2] in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα
D-Ala	**1**	-	4.11	1.57	-	-	-	-	-	48.74
D-Ala	**2**	8.71	4.46	1.47	-	-	-	-	-	49.90
Thr	**3**	8.23	4.51	4.19	1.16	-	-	-	-	58.19
Trp	**4**	8.44	4.71	3.28	-	7.26	10.20, 7.64	7.47, 7.14	7.21	54.67
Arg	**5**	8.35	4.09	1.52, 1.76	1.25, 1.31	3.02	7.10	-	-	52.97
Gly	**6**	7.30	3.55, 3.83	-	-	-	-	-	-	41.91
Gln	**7**	7.96	4.56	1.92, 2.07	2.31, 2.35	6.83, 7.57	-	-	-	50.43
Pro	**8**	-	4.28	1.66, 2.01	1.93, 2.01	3.64, 3.78	-	-	-	60.40
Cys	**9**	7.95	4.49	2.82	-	-	-	-	-	54.93
Trp	**10**	7.60	4.56	3.21, 3.38	-	7.19	10.00, 7.66	7.46	7.21	55.54

Table D.18. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide Gln-Pro *cis* D-Ala[1,2] in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
D-Ala	**1**	-	-	-	-	-	-	-	-
D-Ala	**2**	-	4.37	-	-	-	-	-	50.51
Thr	**3**	8.25	4.41	4.28	1.18	-	-	-	58.31
Trp	**4**	8.29	4.68	3.32	-	7.25	10.17	7.47	-
Arg	**5**	8.16	4.18	1.51, 1.74	1.37	3.04	7.16	-	52.70
Gly	**6**	7.06	3.45, 3.60	-	-	-	-	-	-
Gln	**7**	7.83	4.35	1.83, 2.00	2.31	6.79, 7.49	-	-	-
Pro	**8**	-	4.64	1.99, 2.12	1.84, 1.91	3.48, 3.61	-	-	59.79
Cys	**9**	8.22	4.46	2.86	-	-	-	-	-
Trp	**10**	7.67	-	3.21, 3.39	-	7.18	9.96	7.44	-

D.4 Chemical shift tables for peptide D-Ala^{6,7}

Table D.19. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide dAla-Pro *trans* D-Ala^{6,7} in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	**1**	-	4.31	3.31, 3.36	-	7.23	10.27, 7.56	7.47	56.19
Cys	**2**	8.65	4.44	2.67	-	-	-	-	58.09
Thr	**3**	8.38	4.21	4.21	1.17	-	-	-	62.42
Trp	**4**	8.08	4.63	3.25, 3.31	-	7.24	10.19, 7.54	7.45	57.47
Arg	**5**	7.85	4.12	1.44, 1.62	1.19	2.95	7.08	-	56.13
D-Ala	**6**	7.82	4.17	1.28	-	-	-	-	52.02
D-Ala	**7**	8.35	4.57	1.31	-	-	-	-	50.60
Pro	**8**	-	4.30	1.68, 2.03	1.91	3.56, 3.71	-	-	63.22
Cys	**9**	8.15	4.27	2.68	-	-	-	-	58.45
Trp	**10**	7.61	4.45	3.14, 3.29	-	7.14	10.08	7.43	58.47

Table D.20. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide dAla-Pro *cis* D-Ala^{6,7} in H_2O, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	1	-	4.30	3.31, 3.35	-	7.22	10.24	7.45	-
Cys	2	8.61	4.40	2.57, 2.64	-	-	-	-	58.06
Thr	3	8.33	4.18	-	1.12	-	-	-	62.42
Trp	4	7.94	4.63	3.22, 3.26	-	7.20	10.20, 7.41	7.43	-
Arg	5	8.00	4.24	1.53, 1.69	1.30	3.02	7.11	-	55.97
D-Ala	6	8.12	4.27	1.30	-	-	-	-	51.66
D-Ala	7	8.44	3.77	0.88	-	-	-	-	50.51
Pro	8	-	4.70	1.75, 2.09	1.75, 1.82	3.37, 3.46	-	-	-
Cys	9	8.09	4.43	"2.55 2.84"	-	-	-	-	58.88
Trp	10	7.88	4.43	"3.04 3.34"	-	7.17	10.00, 7.63	7.38	-

Table D.21. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide dAla-Pro *trans* D-Ala[6,7] in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	1	-	4.31	3.29, 3.38	-	7.24	10.20, 7.60	7.48	56.26
Cys	2	8.57	4.47	2.64, 2.70	-	-	-	-	58.04
Thr	3	8.34	4.27	4.27	1.19	-	-	-	62.56
Trp	4	7.83	4.67	3.29, 3.38	-	7.25	10.12, 7.57	7.47	57.54
Arg	5	7.66	4.19	1.44, 1.69	1.20	2.96	7.09	-	55.99
D-Ala	6	7.69	4.26	1.32	-	-	-	-	51.97
D-Ala	7	8.16	4.55	1.28	-	-	-	-	50.46
Pro	8	-	4.27	1.70, 1.96	1.88, 1.91	3.48, 3.71	-	-	63.26
Cys	9	7.93	4.36	2.74	-	-	-	-	58.24
Trp	10	7.51	4.52	3.20, 3.33	-	7.15	9.95, 7.61	7.42	58.46

Table D.22. ^{1}H and ^{13}C chemical shifts (ppm from DSS) for peptide dAla-Pro *cis* D-Ala[6,7] in 30 % TFE, 5 ºC, pH 5.0.

Residue		HN	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Cα
Trp	1	-	-	-	-	-	-	-	-
Cys	2	-	4.46	2.61	-	-	-	-	-
Thr	3	8.3	4.25	-	1.16	-	-	-	62.58
Trp	4	7.74	4.66	3.28, 3.37	-	-	7.55	-	-
Arg	5	7.76	4.27	1.46, 1.72	1.24	2.99	7.10	-	55.88
D-Ala	6	7.84	4.31	1.36	-	-	-	-	51.82
D-Ala	7	8.2	4.09	1.02	-	-	-	-	50.19
Pro	8	-	4.73	1.74, 2.02	1.76	3.40, 3.48	-	-	-
Cys	9	7.97	4.47	2.59, 2.85	-	-	-	-	-
Trp	10	7.68	4.48	3.14, 3.36	-	7.16	9.91, 7.60	7.40	58.76