NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTAMENTO DE QUÍMICA

# NMR INVESTIGATION OF ION-PAIR MODULATION OF PROTEIN STRUCTURE AND DYNAMICS AND THEIR RELATION TO PROTEIN STABILITY

MICAEL SIMÕES SILVA

Mestre em Química Bioorgânica

DOUTORAMENTO EM BIOCIÊNCIAS MOLECULARES

Universidade NOVA de Lisboa

Julho, 2021

NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTAMENTO DE QUÍMICA

# NMR INVESTIGATION OF ION-PAIR MODULATION OF PROTEIN STRUCTURE AND DYNAMICS AND THEIR RELATION TO PROTEIN STABILITY

MICAEL SIMÕES SILVA

Mestre em Química Bioorgânica

DOUTORAMENTO EM BIOCIÊNCIAS MOLECULARES

Universidade NOVA de Lisboa

Julho, 2021

NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTAMENTO DE QUÍMICA

# NMR INVESTIGATION OF ION-PAIR MODULATION OF PROTEIN STRUCTURE AND DYNAMICS AND THEIR RELATION TO PROTEIN STABILITY

MICAEL SIMÕES SILVA

Mestre em Química Bioorgânica

**Orientador:** Doutor Eurico José da Silva Cabrita, Professor Associado com Agregação, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

**Coorientador:** Doutor Ângelo Miguel Mendes Pinto de Figueiredo, NMR Facility Manager, ISMB, University College London and Birkbeck College

**Júri:**

**Presidente:** Doutora Maria João Lobo de Reis Madeira Crispim Romão, Professora Catedrática, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

**Arguentes:** Doutor Óscar Millet Aguilar-Galindo, Investigador Principal, CIC bioGUNE – Centro de Investigación Cooperativa en Biociencias

Doutora Mara Guadalupe Freire Martins, Investigadora Coordenadora, Universidade de Aveiro

**Vogais:** Doutor Manuel José Estevez Prieto, Professor Catedrático, Instituto Superior Técnico, Universidade de Lisboa

Doutora Maria João Lobo de Reis Madeira Crispim Romão, Professora Catedrática, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Doutor Eurico José da Silva Cabrita, Professor Associado com Agregação, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Doutor Aldino José Martins Viegas, Investigador Auxiliar, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

DOUTORAMENTO EM BIOCIÊNCIAS MOLECULARES

Especialidade em Mecanismos Moleculares de Processos Biológicos

Universidade NOVA de Lisboa
Julho, 2021

**NMR investigation of ion-pair modulation of protein structure and dynamics and their relation to protein stability**

*To my family*

*And to those that are in love with fundamental science*

# Acknowledgements

I would like to thank the Assessment and Examination Committee members Prof. Dr. Oscar Millet, Dr. Mara Freire, Prof. Dr. Manuel Prieto, Prof. Dr. Maria João Romão and Dr. Aldino Viegas, for their time and consideration of this thesis. I would also like to thank Dr. Tiago Cordeiro for his guidance as member of my Thesis Advisory Committee.

To *Fundação para a Ciência e Tecnologia* for my PhD fellowship (PD/BD/128202/2016) and financial support within the scope of the PhD program Molecular Biosciences (PD/00133/2012). To School of Science and Technology of NOVA University of Lisbon for educated me since 2010, through a Bachelor, a Master, and now a PhD degree.

First and foremost, I would like to thank my supervisor, Professor Eurico Cabrita. You taught me not only everything that I know about Nuclear Magnetic Resonance (NMR), but you also helped me to grown as a person and as a researcher. In the last years, you supervised my work and put my problems in perspective (including not only this PhD, but also my Master thesis as well as my final project in Biochemistry a long time ago). Thank you for always having trusted me, believed in me, and supported me (including as landlord!). Other person that I have a lifelong gratitude is my co-supervisor, Angelo Figueiredo. Although you move out to London a few years ago, I always felt your enthusiastic support, especially when I struggled with some problem. Now I realise your effort to visit me (and the lab) in your holidays! More than supervisors, I see in you good friends who wish the best for me, personally and professionally. I cannot thank you enough for all writing corrections, discussions, advises, and encouragements, especially when things got difficult or more challenging.

A big thanks to Aldino Viegas. Thank you for all discussions and for giving me feedback and help whenever needed. I also thank the other post-docs in the group, Filipa Marcelo, Ana Sofia Ferreira, and Jorge Dias, for all the support and good atmosphere. To my lab mates, especially Ana, Helena (recent Doctor!), Wagner (very recent Doctor!), Carmen, Sara, Philip and Carlos, I had a great time with you, helping me to have more discipline and attitude in the lab, and I wish you all the best!

To my colleagues in Molecular Biosciences, especially Filipe and Ricardo, thanks for all the jokes and good humour, passing the doctoral training at ITQB with you was the best! To the Young Biophysicists, which helped me to organise a one-day meeting, creating a fantastic party on biophysics! To the brilliant NMR minds that I had the pleasure of meeting at Germany in the EMBOs of 2016 and 2019 (in particular, at the poster session's sharing a beer!).

To my best friends that faculty and science joined, Poeira, Açoriano and Paiva. Thank you for your truthful friendship. I am sure that you have a bright future ahead. To my great friends Miguel (aka Ressacas), Maria, Rafaela, Ana (near MDs!), Sónia, Susana and Tiago as well as G2 and awesome friends, thank you for the good time and many conversations around a dining table. Thank you all for kept me sane throughout this thesis!

*Em bom português, também quero agradecer à minha família por todo o suporte que me deram para eu conseguir estudar longe de casa, nunca me faltou nada. Apesar de muitas vezes não o dizerem, sei que têm muito orgulho em mim por eu estar a concluir o doutoramento. Um abraço para o meu irmão que sempre me apoiou e apesar de termos estado nos últimos anos em cidades diferentes, sempre o senti muito próximo. Vais ter muito sucesso, tenho a certeza!*

Finally, very special thanks to my fiancée, Rita, for her love, understanding and unconditional support. Together we can do anything!

# Abstract

Salt ions differ in their ability to salt out or salt in proteins from aqueous solutions. Their effects on protein stability are known to be connected to ion hydration, ion pairing and ion-specific interactions with the protein. In general, cations follow the Hofmeister series for protein stabilisation behaviour, while for anions this is only true for proteins where the backbone effect is stronger than that of the positively charged side chains. Since at low concentrations electrostatic effects are expected to be the most significant, ion-specific effects become dominant at high salt concentrations. However, the molecular details of how ions interact with proteins have not yet been fully understood. In this thesis, using different nuclear magnetic resonance (NMR) methodologies and ionic liquids (ILs) as models for the investigation of salt and/or ion-pair effects on protein stability, I prove that molecular mechanisms which result in protein stabilisation or destabilisation are opposed (in enthalpic and entropic terms). These mechanisms depend not only on physical and chemical nature of ions but also on the protein properties. The variety of combinations to form ILs allowed the investigation of the effects of choline glutamate ([Ch][Glu], stabiliser) or 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca], denaturant) on two model proteins with distinct stability and structural properties: GB1 ($\Delta G_{F\rightarrow U}$ ~ 7 kcal/mol) and drkN SH3 ($\Delta G_{F\rightarrow U}$ ~ 0 kcal/mol). The possibility of ion-specific interactions with the protein was studied and the changes of protein structure, diffusion, dynamics, and solvent exchange in the presence of IL were characterized. The data gathered for GB1 and drkN SH3 revealed that [Ch][Glu] at high concentrations (> 1 M) stabilises proteins, not only via electrostatic contacts, but also through a preferential accumulation mechanism at the protein surface, in agreement with an entropic-driven mechanism due to excluded-volume effects. At low IL concentrations, a preferential hydration of the protein is not entirely excluded, which could lead to an initial protein destabilisation. On the other hand, [Bmim][dca] denatures proteins by preferential and direct but unfavourable hydrophobic interactions. These interactions occur not only with the core of the folded state, but also with the unfolded state (slowing the folding process). This was revealed by structural and thermodynamic studies with drkN SH3 where it was possible to directly evaluate the perturbations on the unfolded state due to the slow exchange between the two states. This interaction leads to a stabilisation of a residual helical structure in the unfolded ensemble, which contradicts the random coil-like structure typically found in the presence of denaturing agents. The data gathered provided a thorough understanding of IL-protein interactions as well as the mechanism by which they can affect protein's equilibrium thermodynamics and kinetics, illustrating the importance of the unfolded state and a possible impact in the rational design of solvents in biotechnological processes, for example, increasing not only the catalytic activity but also the enzyme thermostability in such media.



**Keywords:** NMR; Ionic Liquids; Protein Stability and Structure; IL-protein interactions.

# Resumo

Os sais diferem na sua capacidade de promover a precipitação ou a solubilização de proteínas em soluções aquosas. Os seus efeitos na estabilidade das proteínas são normalmente descritos tomando em consideração a hidratação dos iões, a força do par iónico e as interações especificas do catião/anião com a proteína. Em geral, em termos de efeito sobre a estabilidade de proteínas, os catiões seguem a série de Hofmeister, enquanto que para os aniões a ordem de estabilidade de Hofmeister é apenas seguida se os efeitos sobre a cadeia principal da proteína forem superiores àqueles com as cadeias laterais carregadas positivamente. Uma vez que a baixas concentrações prevalecem os contactos electroestáticos inespecíficos, que podem destabilizar a proteína, os efeitos específicos dos iões são dominantes a altas concentrações. Apesar de toda a literatura existente, o mecanismo de ação dos iões sobre a estabilidade das proteínas não está totalmente elucidado ao nível molecular. Nesta tese, utilizando diferentes metodologias de ressonância magnética nuclear (RMN) e líquidos iónicos (LIs), como modelos para a investigação de efeitos de sais e/ou pares iónicos sobre a estabilidade e estrutura de proteínas, eu demonstro que os mecanismos moleculares dependem não só das propriedades físicas e químicas dos iões, mas também da proteína. A versatilidade de combinações possíveis para formar LIs permitiu investigar os efeitos do glutamato de colina ([Ch][Glu], estabilizador) e do dicianamida de 1-butil-3-metilimidazólio ([Bmim][dca], desnaturante) em proteínas-modelo com estabilidade e estrutura muito distintas: GB1 ($\Delta G_{F\rightarrow U}$ ~ 7 kcal/mol) e drkN SH3 ($\Delta G_{F\rightarrow U}$ ~ 0 kcal/mol). Entre outros aspetos, foram estudadas a ocorrência de interações especificas ião/proteína e foram caracterizadas as alterações estruturais, difusionais, de permuta e de dinâmica das proteínas na presença dos LIs. Os dados recolhidos para a GB1 e a drkN SH3 indicam que a concentrações altas o [Ch][Glu] (> 1 M) estabiliza as proteínas modelo não apenas por contactos electroestáticos, mas também por acumulação preferencial na superfície da proteína, de acordo com um mecanismo geral entrópico por efeito de volume excluído. Para concentrações baixas deste LI, a hidratação preferencial da proteína não é totalmente excluída, o que poderá levar a uma desestabilização da proteína. Em contraste, o [Bmim][dca] desestabiliza fortemente as proteínas através de interações preferenciais, diretas e hidrofóbicas. Estas interações ocorrem não apenas com o núcleo do estado *enrolado* mas também com estados da população *sem enrolamento.* Utilizando a drkN SH3, devido à permuta lenta entre os dois estados na escala de tempo do desvio químico em RMN, foi possível avaliar diretamente as perturbações dos estados *enrolado* e *sem enrolamento.* Para o estado *sem enrolamento,* esta interação leva à estabilização de uma estrutura residual de hélice, o que contradiz a estrutura de *enrolamento aleatório (random coil)* normalmente observada na presença de agentes desnaturantes.

Os resultados desta tese ilustram uma metodologia que permite uma compreensão detalhada, a nível molecular, das interações proteína-líquido iónico, bem como do mecanismo através do qual é afetada a termodinâmica e cinéticas do equilíbrio entre estados *enrolado* e *sem enrolamento.* Trata-se de uma contribuição importante para o desenho e otimização racional de solventes para utilização em processos biotecnológicos, como por exemplo, para o aumento da termoestabilidade e da atividade catalítica de enzimas.



**Palavras-chave:** RMN; Líquidos Iónicos; Estabilidade e estrutura de proteínas; Interações proteína-líquido iónico.

# Table of Contents

# List of Figures

# List of Tables

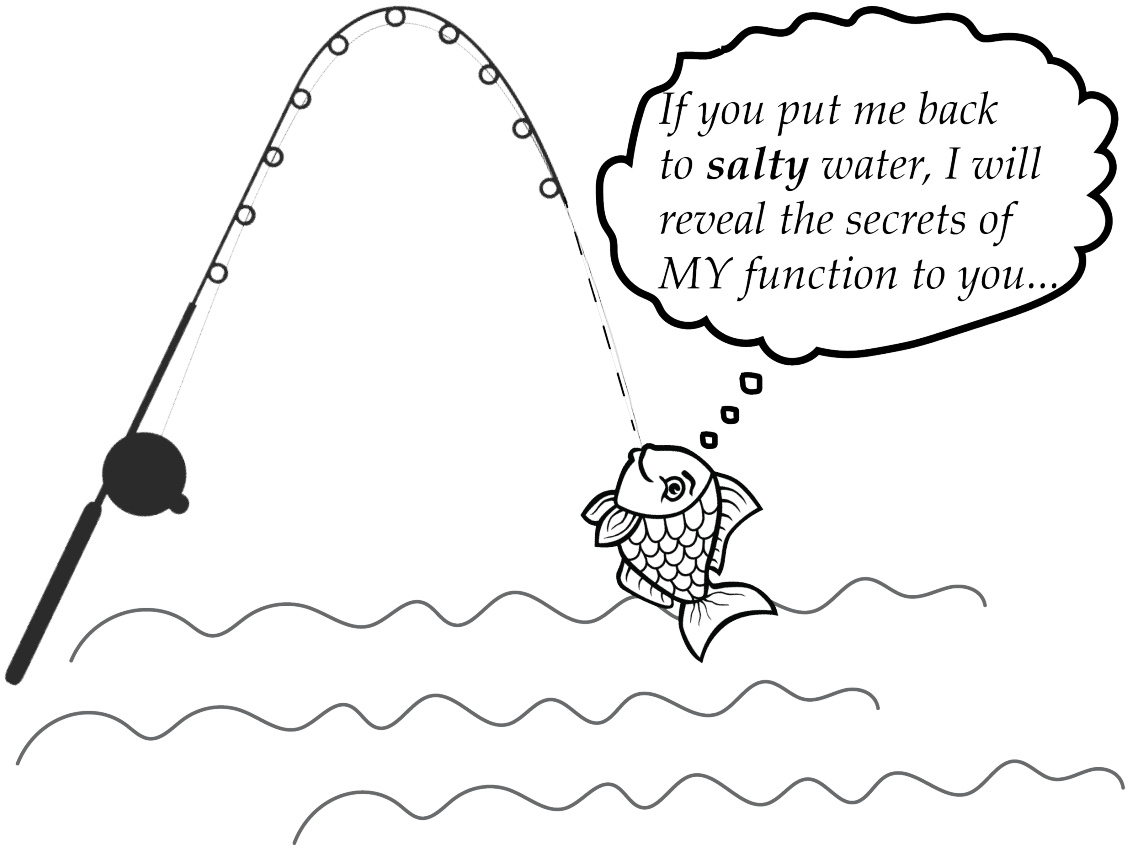
# List of Abbreviations

| | |
|---|---|
| A/ Ala | Alanine |
| [Bmim][dca] | 1-Butyl-3-methylimidazolium dicyanamide |
| [Bmim]Cl | 1-Butyl-3-methylimidazolium chloride |
| C/ Cys | Cysteine |
| CLEANEX-PM | Phase-modulated clean chemical exchange |
| [Ch]Cl | Choline chloride |
| [Ch][Glu] | Choline glutamate |
| CSP | Chemical Shift Perturbations |
| D/ Asp | Aspartic acid |
| drkN SH3 | N-terminal SH3 domain of *Drosophila* adapter protein Drk |
| E/ Glu | Glutamic acid |
| F | Folded state |
| F/ Phe | Phenylalanine |
| G/ Gly | Glycine |
| GB1 | B1 domain of protein G *streptococcus* |
| GdmCl | Guanidinium Chloride |
| H/ His | Histidine |
| HetNOE | Heteronuclear steady-state NOE |
| HSQC | Heteronuclear Single Quantum Coherence |
| I/ Ile | Isoleucine |
| IDP | Intrinsically Disordered protein |
| IL | Ionic Liquid |
| K/ Lys | Lysine |
| $k_{ex}$ | Rate of exchange |
| $k_{fu}$ | Unfolding rate |
| $k_{uf}$ | Folding rate |
| L/ leu | Leucine |
| M/ Met | Methionine |
| N/ Asn | Asparagine |
| Na[DCA] | Sodium dicyanamide |
| NaCl | Sodium chloride |
| [$Na_2SO_4$] | Sodium sulphate |
| NMR | Nuclear Magnetic Resonance |
| NOE | Laboratory-frame nuclear Overhauser effect |

| | |
|---|---|
| P/ Pro | Proline |
| $p_f$ | Folded population |
| pI | Isoelectric point |
| $p_u$ | Unfolded population |
| Q/ Gln | Glutamine |
| R/ Arg | Arginine |
| $R_1$ | Longitudinal relaxation |
| $R_2$ | Transverse relaxation |
| $R_{ex}$ | Chemical exchange |
| ROE | Rotating-frame NOE |
| S/ Ser | Serine |
| $S^2$ | Generalized squared order parameter |
| SSP | Secondary Structure Propensity |
| STD-NMR | Saturation transfer difference |
| T/ Thr | Threonine |
| $T_m$ | Melting temperature |
| TS | Transition state |
| $T_s$ | Temperature of maximum stability |
| U | Unfolded state |
| V/ Val | Valine |
| W/ Trp | Tryptophan |
| Y/ Tyr | Tyrosine |
| Zzex | ZZ exchange |
| $\Delta C_{p,\mathrm{u}}$ | Equilibrium heat capacity of unfolding standard |
| $\Delta G_{\mathrm{HX}}$ | Free energy of opening |
| $\Delta G_u$ | Equilibrium free energy of unfolding |
| $\Delta G_{\mathrm{U,F\text{-}TS}}$ | Activation free energy for folding or unfolding standard |
| $\Delta H_u$ | Equilibrium enthalpy of unfolding |
| $\Delta S_{\mathrm{conf}}$ | Change in conformational entropy |
| $\Delta S_u$ | Equilibrium entropy of unfolding |
| $\Delta\delta_{\mathrm{comb}}$ | Combined chemical shift |
| $\Delta\omega$ | Frequency difference |
| $\Delta\sigma_{\mathrm{HN}}/\Delta T$ | Amide proton temperature coefficient |
| $\tau_m$ | Correlation time |

1

# Chapter 1: General Introduction




**Cartoon representing that "salty" water (with ions and osmolytes) is essential for proper functioning of biological molecules (fish in the drawing), but the functionality dominantly belongs to the biomolecule itself. Adapted from Jungwirth[1].**

## Introduction

It has been known for more than 130 years that salt ions modulate the physical properties of proteins in aqueous solutions[2]. This concept, and the ability of ions to salt out proteins, has been illustrated in the well-known Hofmeister series. While it is clear that this series is intimately connected to ion hydration, ion pairing and ion-specific effects, its molecular origin has not yet been fully understood.

Additionally, ions can associate, and form molten salts known as ionic liquids (ILs)[3,4]. These are a class of non-molecular compounds composed of a combination of organic cations and organic/inorganic anions that exist as liquid at temperatures below 100°C[5]. Consequently, because small and charged metabolites are abundantly found in the cytoplasmatic milieu, the question rises whether the possible formation of naturally occurring ILs is possible and, if so, it may influence the stability of proteins beyond Hofmeister chemistry.

The following introductory sections provide some fundamental aspects of protein stability, role of cosolute interactions, and biomolecular Nuclear Magnetic Resonance (NMR) spectroscopy. I start by introducing concepts of protein stability and kinetics of protein folding, and then I focus on folding pathways and energy landscape. Next, the role of cosolute interactions, such as ion specific effects, is discussed. The main emphasis is on ionic liquids and their effects on proteins. The implications of these interactions on enthalpy-entropy compensation are considered. Finally, the background for the choice of ILs and model proteins is provided. In the last section, protein NMR techniques are presented, from a simple HSQC to the backbone and side chain assignment experiments. The methods for exploring protein dynamics, from ps-ns to seconds time scales, are also reviewed.

## Protein Stability

Most of our understanding of the molecular basis of numerous biological processes has come from the elucidation of the conformation of proteins and on the investigation of the relationship between their three-dimensional structures and their biological function. The structure-function paradigm states that the amino acid sequence of a globular protein encodes its singular, minimum-energy and biologically active conformation[6].

In the early 1960s, Anfinsen discovered that the small protein ribonuclease A and staphylococcal nuclease could be reversibly denatured[7]. On removal of a chemical denaturant, such as urea, they spontaneously refold to their native structures. Similarly, they spontaneously refold on cooling after thermal denaturation. Not only do the amino acid sequences of these proteins encode their final folded structures, but they also encode the information on how to get to the structures (i.e., the folding pathway). This seminal work was a major milestone in protein science as it was a formative demonstration that protein folding is a straightforward biophysical process. These ideas, coined in the Anfinsen's dogma or the thermodynamic hypothesis, state that "the three-dimensional structure of the native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other) is the one in which the Gibbs-free energy of the whole system is the lowest; that is, that the native conformation is

determined by the totality of the interatomic interactions and hence by the amino acid sequence, in a given environment"[7]. However, most of proteins will not renature after being denatured and larger or multimeric proteins require assistance to fold (which is provided by molecular chaperones). In the light of their existence, the Anfinsen thermodynamic hypothesis was reconfigured reflecting a funnel-shaped energy landscape diagram, which states that proteins must fold energetically downhill, shrinking in conformational extent as they go[8,9].

**Kinetics of Protein Folding**

To understand protein stability, we must first consider the physical nature of protein folding, that follows very simple rate laws, governed by a few basic principles. A denatured protein makes many interactions with solvent water. As the protein folds, it exchanges those intermolecular noncovalent interactions (e.g., hydrophobic effects, van der Waals forces, coulombic attractions) by intramolecular ones: its hydrophobic side chains tend to pack with one another, and many of its hydrogen bond donors and acceptors pair with each other, especially those in the polypeptide backbone, forming the hydrogen-bonded networks in $\alpha$-helices and β-sheets. While each interaction's energy is small, due to their large numbers, the total interaction energies in the native and denatured states are very large, adding up to some thousands of kilocalories per mole, depending on the size of the protein. Yet, proteins are only marginally stable, their free energies of unfolding ranging from 5 to 15 kcal/mol (20 - 60 kJ/mol)[10], energetically equivalent to the free energy of one or two typical hydrogen bonds in water. This tiny amount of energy (or the low stability of a typical globular protein) is expressed by the low difference in standard state Gibbs free energy of unfolding ($\Delta G_u$) between the native/structurally defined folded state (F) and the thermodynamic ensemble of denatured/unfolded states (U). Considering only a two-state kinetic model, these states are almost balanced in energy and their difference is crucial and has to be positive for a stable conformation,

$$F \rightleftharpoons U,$$

$$\Delta G_u^{0\prime} = G_U^{0\prime} - G_F^{0\prime}, \qquad [1.1]$$

in which the reversible two-state process is characterised by the equilibrium constant,

$$K_{eq} = \frac{[U]}{[F]} = exp\left(\frac{-\Delta G_u^{0\prime}}{RT}\right), \qquad [1.2]$$

where R is the gas constant. The low $\Delta G_u$ is founded in the mutual compensation of large enthalpic and entropic contributions, $\Delta H_u$ and $\Delta S_u$, respectively, known as enthalpy-entropy compensation[11],

$$\Delta G_u^{0\prime} = \Delta H_u^{0\prime} - T\Delta S_u^{0\prime}. \qquad [1.3]$$

As a protein unfolds, $\Delta H_u$ has a positive and a negative contribution, the first from unfavourable changes arising from the breakage of intra-protein bonds and water-water bonds, and the latter from favourable changes arising from the solvation of newly exposed surfaces[12]. The entropic contribution to $\Delta G_u$, $-T\Delta S_u$, is a balance of a negative change as the

protein becomes more disordered with unfolding (favourable), and a positive change, as water molecules order themselves around the newly exposed protein surfaces (unfavourable). Because the folded state possesses the lowest free energy and the unfolded ensemble is entropically favoured (less structurally ordered), increasing $\Delta H_u$ and/or decreasing $\Delta S_u$ increases protein stability by increasing $\Delta G_u$. This equilibrium can be manipulated by adding cosolutes[13]. The cosolute-induced changes of $\Delta H_u$ and $\Delta S_u$ relative to the cosolute-free can provide a thermodynamic fingerprint of the protein in the presence of such cosolute, i.e., enthalpic (de)stabilisation offset with entropic (de)stabilisation.

Additional insights on the understanding of protein stability came from denaturation studies[14]. Denaturation, a phenomenon known for a long time, can occur by changing the protein's physical or chemical environments, such as variations in temperature or pressure, or the addition of cosolutes. Since the denatured state has considerably more conformational freedom than a native state, it has an inherently high configurational entropy [$S = k_B \ln(W)$, where $k_B$ is the Boltzmann constant and W is the number of accessible microstates][15]. Conversely, the native state is very conformationally restricted and has a low entropy. Thus, as a protein folds, it loses considerable entropy, which must be balanced by a gain of enthalpy for the free energy to favour folding. The enthalpy of packing side chains in the native state is favourable and compensates, just barely, for its low entropy[16]. The classic thermodynamic description of protein folding[17] implies large negative values of $\Delta H_u$ and $\Delta S_u$ (absolute values of $\Delta H_u$ are typically ~ 100 kcal/mol[15]). Furthermore, these observed thermodynamic values must include the contribution of solvent (entropy and enthalpy of water) through the gain in entropy of release of water on folding and the specific heat of unfolding ($C_p$). The heat-capacity increase upon unfolding ($\Delta C_{p,u}^{0'} = C_{p,U}^{0'} - C_{p,F}^{0'}$), dominated by the hydration of hydrophobic groups[18], leads to a generic effect of temperature dependence of free energy of unfolding (i.e., a protein stability curve)[19].

One of the hallmarks of a two-state system is that kinetic and thermodynamic parameters extracted through different methods yield similar rates and stabilities, and the process is completely reversible[20]. A two-state folder is separated by a single activation barrier, the so-called transition state (TS or ‡) or transition-state ensemble[21]. This state consists in an ensemble of conformations that differ only slightly from one another in their energy around the saddle point. This type of energy surface should lead to the transition states in protein folding being very wide and with long saddle points, with many small dips in the profile (Fig. 1.1).

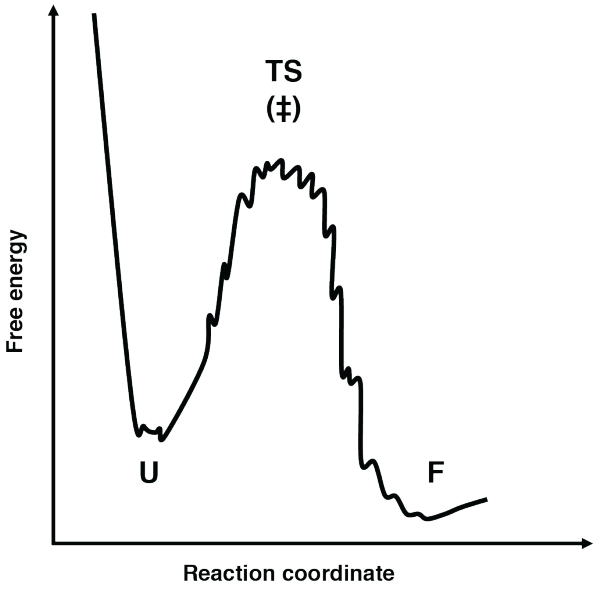



**Figure 1.1. Free energy versus reaction coordinate diagram for protein folding.**

The reaction coordinate is the path taken as the reaction proceeds. The unfolded ensemble (U) and the folded state (F) sit at the bottom of potential energy wells, whereas the transition state (TS or ‡) is at a maximum. In practice, there are many minor maxima and minima because there are many small conformational changes of low energy. At each point along the pathway, the structure lies at a minimum energy for structural changes caused by movements in structure perpendicular to the direction of motion. There are thus minor slightly higher energy paths that parallel the major lowest energy one. Adapted from Fersht[21].

The energy surfaces in protein folding also have very high entropy components because of the large changes in configurational entropy, as a protein folds, and the changes in entropies of hydration. In general, due to its transient nature, it is indirectly measured by changes in the free energy of activation ($\Delta G^{\ddagger}$)[15,22]. Further details are given in Chapter 3, which is dedicated to the changes in stability and folding of the marginally stable drkN SH3, where it was possible to record kinetic data from magnetisation transfer rates.

**Folding pathways and energy landscape**

Notwithstanding the tremendous advances in the understanding of protein folding, since Anfinsen's work[7], there is still not a general agreement on the overarching questions of: how do proteins fold, and why do they fold in that way?[23,24].

In the late 1960s, Levinthal argued that because there is an astronomical number of conformations open to the unfolded state, an unbiased search through these conformational space would take too long for a protein to fold[25]. It was a short logical step to argue that there must be defined pathways to simplify the choices in folding. Typical protein folding rates are on the microsecond to second timescale[20]. In order to fold on such fast timescales, protein folding cannot be governed by a random trial and error, but, instead, involve energetically guided pathways (downhill toward the native structure) that dramatically reduce the sampling space and thus the time required to fold, a principle that is captured in funnel-shaped energy landscape[8,9], as represented in Fig. 1.2a. Developed in the 1990s mainly by José Onuchic and Peter Wolynes' groups[26–29], the funnel is a conceptual mechanism for understanding the self-organization of a protein and how it avoids the Levinthal paradox. At the top of the funnel, the protein exists in a large number of random states that have relatively high enthalpy and entropy. There is a competition between the maximization of entropy, keeping the protein as random as possible at the top of the funnel, and the minimization of enthalpy, trying to drag the protein down the funnel as it folds. Progress down the funnel is

accompanied by an increase in native-like structure as folding proceeds (Fig. 1.2b). Proteins fold from the random state by collapsing and reconfiguring. The reconfiguring occurs by a Brownian type of motion between conformations that are geometrically similar and follows a general drift from higher energy to lower energy formations. The funnel is thus a progressive collection of geometrically similar collapsed structures, one of which is more thermodynamically favourable than the rest. Folding is driven by an increase in the acquisition of native-like interactions as the collection of states rapidly interconverts[30].

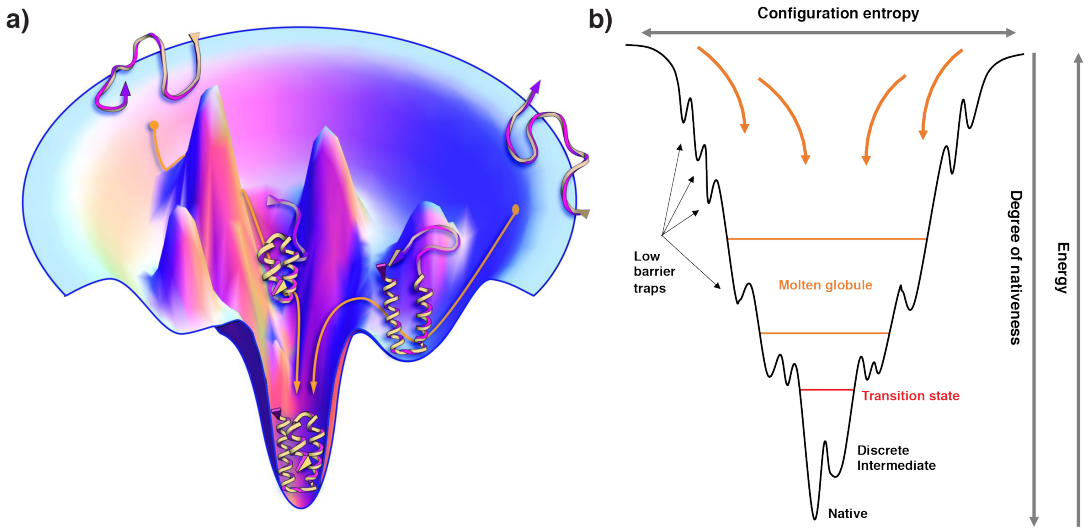



**Figure 1.2. Energy landscape of protein folding.**

Three-dimensional **(a)** and two-dimensional **(b)** representations of a funnelled energy landscape for protein folding, illustrating the principle of minimal frustration. The ensemble of unfolded configurations occupies the top of the energy funnel. As the protein becomes more native-like and folded its accessible configurational entropy, i.e., the width of the funnel, decreases which is offset by the gain of energy from increasingly formed native interactions. There is not, however, a unique and single pathway for folding but the whole process functions via an ensemble of converging pathways. The molten globule state is identified as a loosely collapsed state with fluctuating tertiary interactions and very weak secondary structure, then follows the transition state. Evolution smoothened the surface of the energy landscape sufficiently to prevent entrapment in local minima and ensure the robustness of this molecular self-assembly in a crowded cellular environment. **a)** Adapted from Dill and MacCallum[23]. **b)** Adapted from Onuchic et al[30].

The first clue that mechanistic complexity still remains on funnelled landscapes is precisely the ensemble nature of the transition state described above. We can convert the energy landscape funnel (Fig. 1.2) to the free energy-reaction coordinate diagram (Fig. 1.1) by lumping together the various states as ensembles. Because biomolecular stability is dictated by the sum of a large number of attractive and repulsive interactions of similar strength, the resulting free-energy landscape is rugged, with the global minimum in the surface thought to correspond to the native state of the biomolecule. In addition, there are often local minima that are separated from the global minimum by free-energy barriers that can be overcome by thermal fluctuations. These low-lying conformationally excited states have remained elusive to quantitative investigation because their sparse population and transient nature

complicates their study by conventional structural biology methods which are used to characterize highly populated / ground-state conformers[31,32].

Notably, the roughness of the folding funnel (Fig. 1.2) correlates very well with frustration[33–35]. Energy landscape theory[28] postulates that sites of minimal frustration are associated with stable folding cores of proteins and that these minimal frustrations result when inter-residue interactions in a polypeptide chain are not in conflict with each other and cooperatively lead to a low-energy conformation. Therefore, presence of frustration is frequently correlated with local instability of proteins since unfavourable contacts between individual pairs of amino acid side chains at frustrated sites induce dynamic rearrangements and thus instability. In such cases, the landscape is more rugged, and it explains the observation that kinetically stable states (folding intermediates) may frequently be populated during folding[36] and may form part of the ensemble of states populated under a given set of cellular conditions (e.g., pH, temperature, or concentrations of partner ligands).

Despite all the knowledge gathered so far, the nature of protein folding pathways is still an extremely active field of debate[24,37,38]. The new view of multiple microscopic pathways through a funnelled landscape is often compared to the classical view of a defined folding pathway, with predetermined sequential steps. In summary, energy landscape theory could include the existence of foldon-dependent intermediates (that leads to a faster route for folding)[37] but is the combination of multiple microscopic paths and a bias toward native interactions intrinsic to a funnelled landscape - not intermediates - that resolve the Levinthal paradox[38]. Multiple paths even for two-state folding have been experimentally confirmed using recent single-molecule experimental methods[39]. Additionally, and thanks to the large Protein Data Bank of known structures[40], predicting protein structures is now far more successful than was thought possible in the early days (for example, see the latest DeepMind's AlphaFold artificial intelligence software relied on deep neural networks training[41]).

**Intrinsically unstructured proteins**

After half a century of structural studies on well-folded globular proteins, it is perhaps a shock to discover that up to some 40% of the proteins in the human proteome are estimated to be intrinsically disordered and become fully or partially structured upon binding to partners in the cell[42]. Intrinsically disordered regions (IDRs) or entire intrinsically disordered proteins (IDPs)[43] are perfectly functional, playing vital roles in cellular regulation and signalling, and in numerous neurodegenerative diseases[44].

In contrast to the backbone atoms of well-folded proteins, stabilised by hydrogen-bond networks and organized around hydrophobic cores, fluctuate within a narrow well of the energy landscape. IDPs have shallow energy landscapes with an absence of a global minimum, leading to a system without a dominant well-defined conformation, explaining their higher conformational flexibility and structural plasticity. Such unstructured proteins are characterized by biased amino acid composition and low sequence complexity, as well as low proportions of bulky hydrophobic amino acids and high proportions of charged and hydrophilic amino acids (i.e., absence of both hydrophobic core and stable secondary structure). Despite being perfectly functional, such protein sequences are unable to fold

spontaneously into well-defined structures; instead, they are dynamically disordered (which in turn underlies their high conformational entropy) and fluctuate rapidly through a range of conformations, which cover a continuum of conformational space ranging from extended statistical coils to collapsed globules[44–46]. Modelling of these ensembles is based primarily on NMR data, with small-angle X-ray scattering (SAXS) and single-molecule fluorescence energy transfer (smFRET) also playing an important part in their characterisation.

## Role of cosolute-protein interactions

### Beyond the Hofmeister series: Ion-specific effects

Specific ion effects are ubiquitous in chemistry and biology. Such effects have been recognized since as early as 1888 by Hofmeister in a seminal paper entitled "About regularities in the protein precipitating effects of salts and the relation of these effects with the physiological behaviour of salts"[2]. This was the second and, arguably, most important article in a series of seven articles with the running title "About the science of the effect of salts" published between 1887 and 1898 (ref[47]). This particular article presented a set of experiments ordering various salts according to their ability to precipitate egg-white proteins from aqueous solutions. By using salts with a common cation or anion, Hofmeister ingeniously separated cationic and anionic effects, thereby establishing what later became known as the lyotropic or Hofmeister series for ions[48] (Fig. 1.3). Although this ranking is widely used, outliers and exceptions exist and are discussed controversially[49], and a molecular level understanding is still lacking.

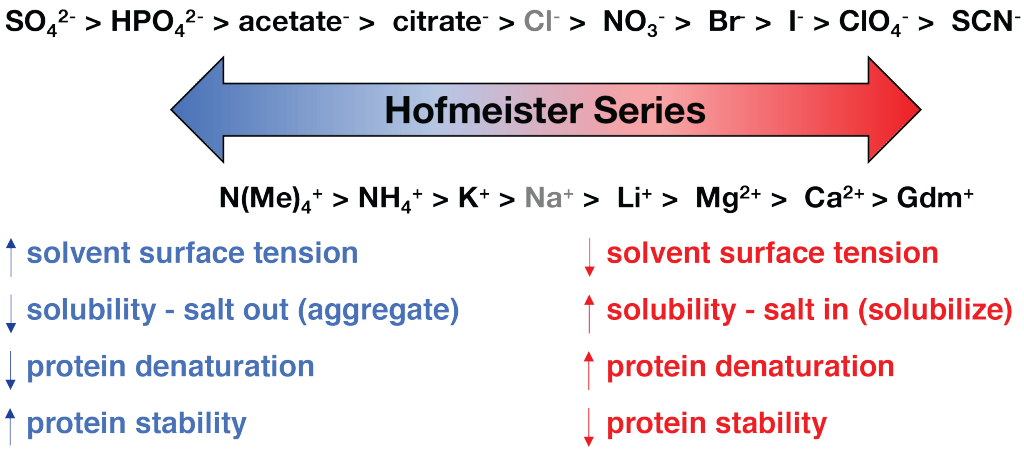



**Figure 1.3. The Hofmeister series.**

Modern version of the anionic (top) and cationic (bottom) Hofmeister series and the accompanying physical properties including the salting out ability. Adapted from Zhang and Cremer[48].

The initial explanations for the ionic ordering were eventually framed in the concept of structure-breaking ions, in which, ions are divided into structure makers (kosmotropes) and structure breakers (chaotropes), with the former having and the latter lacking the ability to order water molecules beyond their immediate solvation shells[50]. According to this view, strongly hydrated kosmotropes effectively 'steal' water from the protein, leading to

a salting-out effect, whereas weakly hydrated chaotropes do not possess this ability. While this explanation of the Hofmeister phenomena is appealing because of its simplicity, it brings in serious problems. Mainly because strongly hydrated ions at physiological and even at higher ionic strengths do not significantly influence water (i.e., long-range water ordering) beyond their immediate hydration layer[51]. Therefore, the whole concept of "kosmotropes" and "chaotropes" were soon disproved by spectroscopic and thermodynamics studies[48]. Moreover, the protein solute needs to be brought explicitly into the picture since salting out behaviour cannot be explained by considering ions and water only. As the most notable example, precipitation of positively charged lysozyme follows the Hofmeister series only at high pH values and high ionic strength but follows an apparently reversed series under neutral and acidic conditions, and up to moderate (~ 0.2 – 0.3 M) salt concentrations[52].

In fact, since the 1960s there has been explicit interest in understanding how ions interact with proteins, in particular using a reductionist approach of dividing polypeptide effects into backbone and side-chain contributions[53]. Following simple thermodynamic considerations that attractive ion–backbone interactions lead to salting in (and destabilisation) of the protein, it has been demonstrated that the amide groups interact favourably with weakly hydrated anions (e.g., $I^-$, $ClO_4^-$ or $SCN^-$) and, to a much lesser extent, with strongly hydrated cations (for example, $Mg^{2+}$, $Ca^{2+}$ or $Li^+$). About 25 years ago, in an attempt to understand Hofmeister ordering at molecular level, Baldwin[54] argued that specific ion effects on protein stability could be described in terms of the ability of ions to salt in the polar peptide group and salt out the nonpolar side chains. At first glance, it seems plausible that salting-out ions strengthen hydrophobic interactions (protein stabilisation), whereas salting-in ions decrease them (destabilisation). Yet, a series of studies, suggest specific ion effects on hydrophobic interactions are more complicated, and that we need to account with all interfacial ion-protein interactions, as well as ion-ion interactions at the protein surface and in the surrounding solvent[49]. For an extensive review, I refer the reader to Lo Nostro and Ninham[55].

The cationic and anionic Hofmeister series can now be rationalized primarily in terms of specific interactions of salt ions with the backbone and charged side chain groups at the protein surface in solution[56]. The crucial regions of protein surfaces interacting with ions are the backbone and charged side chains, with polar and hydrophobic side chains playing a much smaller role. Overall, in comparison with anions, the backbone interactions of cations are less pronounced. It should be pointed out that:

(1) Cations follow standard Hofmeister ordering with strongly hydrated cations interacting more strongly and thus being less efficient in protein stabilisation (salting out) than weakly hydrated ones. This is true both at the protein backbone and at negatively charged side chains (also, it prefers pairing with smaller, rather than larger cations), with the former interactions being significantly weaker than the latter.

(2) The picture for anions is more complex, with backbone and side chain interactions being oppositely ordered. At the backbone, they follow the normal anionic Hofmeister series with weakly hydrated anions interacting more strongly and thus being less efficient in protein stabilisation (can even lead to denaturation) than strongly hydrated ones. However, at the positively charged side chains, the anionic ordering is reversed, as it pairs more efficiently with smaller, rather than larger anions.

As a result of the above considerations, cations follow the Hofmeister series for protein stabilisation behaviour, while for anions this is only true for proteins where the backbone effect is stronger than that of the positively charged side chains. For strongly positively charged proteins such as lysozyme at low to neutral pH values, the anions can actually follow a reversed Hofmeister series[52]. In addition, it is not only the interaction of a given ion with the protein surface that must be considered, but also its interactions with counterions (e.g., strong cation-anion interactions (ion-pair)). As an example, concentrated guanidinium chloride (GdmCl) acts as a strong denaturant of model peptides, whereas concentrated guanidinium sulfate ($Gdm_2SO_4$) does not[57]. This difference in behaviour comes from the inhibition of the denaturant power of guanidinium cations by appreciable ion-pairing with sulfate anions in the solution.

Moreover, the ion-specificity is concentration dependent: while electrostatic effects are expected to be most significant at low concentrations, ion-specific effects have been observed both at low and high concentrations[13,58]. Charged proteins undergo electrostatic interactions with ions at very low concentrations, which result in protein destabilisation at 0.25 and 0.5 M by almost all salts. At high salt concentrations (≥ 1 M) electrostatic forces are screened and ion-specific effects become dominant. For stronger destabilising salts, such as $Na[ClO_4]$, the stability of the protein decreases monotonously.

**Ionic Liquids**

The past few decades have witnessed an explosion of interest in ionic liquids (ILs). The term "ionic liquids" replaced the older phrase "molten salts", and it refers to liquids that are composed solely of ions, by an asymmetric organic cation and a symmetric/asymmetric organic/inorganic anion. Mainly related with the asymmetry of the structure of these organic salts and the type of interactions between cation and anion, are their melting points (mp), that can be much lower than inorganic salts, with many combinations rendering liquids below 100°C (373 K)[5], and therefore amenable of being used as solvents. Their great reputation in several fields[59,60] is due to their attractive properties such as low vapor pressure, high thermal stability, high conductivity, non-flammability, and biocompatibility.

While the first report of a room-temperature molten salt, ethylammonium nitrate (EAN, $[EtNH_3][NO_3]$, mp 13-14ºC), goes back to 1914 by Walden[61], at the time, it was not recognized that chemistry in such solvents could become of widespread interest. Only in the 1990s it became increasingly clear that many ion combinations form air- and water-stable ILs, by way of Wilkes and Zaworotko[62] works, who reported on the preparation and characterisation of a new range of ILs containing 1-ethyl-3-methylimidazolium cation (denoted as $C_2mim^+$ or $Emim^+$) with alternative anions, [Emim]X (X = $[CH_3CO_2]$, $[NO_3]$ or $[BF_4]$). Since then, ILs have become increasingly popular in academia and industry. The perceived benefit of substituting traditional industrial solvents, most of which are volatile organic solvents (VOCs), with ILs as the "green" alternative due to their low volatility, was one of the primary driving forces behind research ILs. But this is not the only reason for using ILs. As there at least a million simple ionic liquids (and a trillion ternary systems) that can be easily prepared in the laboratory, they provide access to a wide range of physicochemical properties which

can be tailored for virtually any application[63]. This means that the solvent can be fine-tuned by the independent selection of cation and anion, so that ILs are called "designer solvents".

However, the bulk structure and interfacial behaviour of ILs is very complex, and their molecular-based interpretations often gives rise to controversies and speculations[3]. ILs are solvents that can participate in a variety of attractive interactions ranging from the weak, nonspecific, and isotropic forces (e.g., van der Waals, solvophobic, dispersion forces) to strong (Coulombic), specific, and anisotropic forces (e.g., hydrogen bonding, halogen bonding, dipole−dipole, magnetic dipole, electron pair donor/acceptor interactions). The diversity and strength of intermolecular forces in ILs fine-tunes local arrangements in the bulk and near interfaces. Nevertheless, some interactions have an entropic contribution, paving the way for complex, higher order self-assembled structures. Undeniably, ILs exist as liquids under standard ambient conditions due to the anion/ cation chemical structure. Since the liquid state is thermodynamically favourable due to the large size and conformational flexibility of the ions involved (that destabilises the solid-phase crystal), the small lattice enthalpies and large entropy changes favour melting. For instance, the cation alkyl chain must be long enough to reduce Coulombic forces and disrupt lattice packing, but not be too long ($\sim n < 12$) as this will increase salt melting point despite the enhanced asymmetry (cohesive interactions increase with length of nonpolar groups as per linear alkanes)[64]. The (nano)structure in ILs (including some ion-pair considerations) has been extensively reviewed[4].

ILs, like general solvents, are usually classified on the basis of chemical structure. Protic and aprotic ILs (AILs and PILS, respectively) are the two most common IL types, based on the well-established division between proton-donating (protic) and nonproton-donating (aprotic) molecular solvents. The H-bond donating ability is usually a property of the cation, while anions act as H-bond acceptors. For a long-time, research has mainly focused on AILs, represented by derivatives of pyrrolidinium, pyridinium or imidazolium cations in typical combination with halide, tetrafluoroborate, hexaflurophosphate, acetate or bis-(trifluoromethanesulfonyl)-imide anions. Despite of their classification as aprotic species, some cations can form weak H-bonds with proton-accepting anions. For example, the acidic hydrogen at carbon C-2 of the imidazolium ring[3]. In contrast to AILS, PILs consist of Brønsted acids and bases and often include ammonium or phosphonium derivatives as a cation, combined with a phosphate, sulfate or nitrate anion. Importantly, PILS possess higher strength of Coulomb- and hydrogen bonding-interactions as compared to their aprotic counterparts and form an extensive network of hydrogen bonding, which, in some regards, resembles the water H-bond network[65]. A prime example is cholinium ($Ch^+$), which possess an OH-terminated alkyl chain. Several other IL subclasses are reported in the literature[4], based upon distinct structural features (e.g., chiral, magnetic, divalent, polymeric, or fluorous ILs) or named from specific functional groups introduced on the ions (amino acid ILs and aryl alkyl ILs).

In the last two decades, we assisted to a growing awareness of the environmental impact of manufactured chemicals. While, the potential of ionic liquids to conceptually fulfil the requirements of environmental sustainability can be remarkable, the truth is that many ILs, including those based on imidazolium or pyridinium cations and halide-containing anions, if released into the environment, are indeed very toxic and poorly biodegradable[66].

Therefore, and to answer the increasing demand for environmentally friendly solvents, a significant amount of research is currently being done to determine the consequences and the environmental risks of ILs towards environmental systems and to synthetize active ILs with low toxicity and high biocompatibility[67]. The ideal strategy for synthesizing biocompatible ILs (Bio-ILs) uses both anionic and cationic counterparts derived from natural (i.e., non-toxic, biodegradable, biocompatible and renewable) sources, such as amino acids (AAs), non-nutritive sweeteners, glucose, and carboxylic acids. The first ILs derived from AAs (that is, amino acid ILs, AAILs) were introduced by Fukumoto et al. in 2005[68], through coupling the $emim^+$ cation with 20 different natural AA. Also noteworthy is the attention given to ILs with cholinium as the cation (cholinium-based ILs)[66]. Choline ($Ch^+$), an essential micronutrient for normal functioning of all cells[69], found in the head groups of cell-membrane phospholipids and precursor of the neurotransmitter acetylcholine. It is also known to be non-toxic and biodegradable (and inexpensive). Eventually, the above considerations led to the successful synthesis of cholinium-based amino acids ILs ([Ch][AA])[70]. The selection of a benign cation, such as $Ch^+$, combined with benign anions (e.g., lactates, acesulfamates, levulinate, carbohydrate derivatives, or AAs) constituted a major breakthrough in the conscious design of ionic liquids.

**Ionic Liquids – Protein interactions**

Due to the potential use of ILs as designer solvents for biocatalytic reactions or the formulation and storage of proteins, their effect on the stabilisation of proteins has attracted considerable attention. Undeniably, depending on their nature (i.e. neat or aqueous state), ILs are able to tune/modulate protein stability and activity[71–73].

Because the thermodynamic stability of native proteins is marginal, changes in the solvent environment can dramatically affect their properties. While unfolding equilibria are often reversible, folding intermediates and misfolds can promote irreversible protein aggregation into amorphous precipitates or highly ordered amyloid states[74]. Addition of ILs offers intriguing prospects for stabilising native proteins against these toxic off-pathways[75,76]. It may be possible that addition of such (benign) ILs facilitates conformational transitions via energetically favourable interactions. The protein fluctuates between multiple potential energy minima, overcoming free energy barriers between them, toward a thermally stable conformation. In practical terms, it smooths the ruggedness of the funnel-shaped folding energy landscape (above-mentioned in Fig. 1.2), lowering free-energy barriers and avoiding conformations to fall into kinetically trapped misfolded states[77].

In this section, the focus is on aqueous solution of ILs and specific protein-IL interactions towards protein (de)stabilisation. In 2013 and 2014, our research group established that imidazolium-based ILs can form contact ion-pairs in solution with consequences for protein stability, structure, and dynamics[78,79]. The reported data supports that protein destabilisation is a consequence of direct preferential binding of anion and/or cations to the protein surface (hydrophobic and electrostatic interactions), in combination with ion hydration. Although denaturing interactions with the anion are dominant and dictate the overall stability, the anion effect can be partially overcompensated by combination with a suitable cation,

through cation–anion pair and cation–protein interactions). For instance, it was found that longer alkyl side chains of imidazolium-cation based ILs ($Bmim^+$, more hydrophobic) are associated with a higher destabilisation effect of Human Serum Albumin (HSA) than short-alkyl groups ($Emim^+$, less hydrophobic), as shown in Fig. 1.4a. The reason for such destabilisation lies on the increased surface contact area of the cation with the protein, particularly on the hydrophobic contacts promoted by the terminus of the alkyl chain.

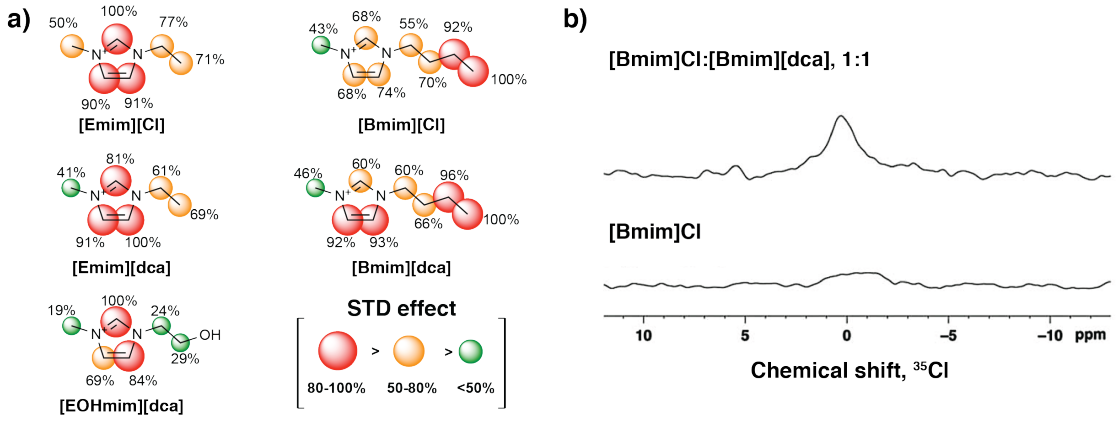



**Figure 1.4. Imidazolium-based ILs interactions with HSA protein.**

**a)** Epitope mapping obtained for the imidazolium cations of different ILs in the presence of HSA. While for $[Emim]^+$ the interaction is predominantly with the core of the aromatic imidazolium ring, for $[Bmim]^+$ the pattern of interaction shifts to the extended alkyl chain, with the maximum STD-NMR observed for the terminal methyl group of the butyl chain. **b)** $^{35}Cl$ NMR competition experiment between [Bmim][Cl] and [Bmim][dca]. In a [Bmim][Cl] solution containing sufficient HSA to broaden completely the $^{35}Cl$ signal (due to $Cl^-$/HSA interactions), the same molar amount of [Bmim][dca] was added. Remarkably, the $^{35}Cl$ NMR signal was partially recovered after addition of [Bmim][dca], meaning that $Cl^-$ binding sites were occupied by a stronger interacting $dca^-$ anion with a concomitant increase in the free $Cl^-$. Adapted from Silva et al[79].

In addition, Fig. 1.4b shows that direct anion interactions with the protein are the origin of destabilising effects. Proof of this is the strong denaturing capability of $dca^-$ (due to direct interaction with the protein), confirmed by comparison with an anion with moderately destabilising effect ($Cl^-$). Overall, for the negatively charged, small alpha-helical proteins Im7 and HSA, the anions and cations follow a direct Hofmeister series: [Emim]Cl > [Bmim]Cl > [Emim][dca] > [Bmim][dca] (the strongest denaturant).

Recently, protein stabilisation in aqueous IL solutions has been the subject of numerous and extensive reviews[80–86]. Nevertheless, investigations of IL effect on protein structure and stability at a molecular level are still scarce, and a series of parameters must be considered: IL composition (cation and/or anion with presence or lack of functional groups), protein surface charge and composition, solution pH, and specific cation/anion effects (ion-pair strength). In these studies, the observations of specific ion-protein interactions[49,56] (above discussed to explain Hofmeister series) were confirmed as valid, therefore site-specific interactions (e.g., electrostatic *versus* hydrophobic character) are crucial.

Based on the change in the $T_m$ (melting temperature) of positively charged ribonuclease A (RNase A) in various ionic liquids and inorganic salts, Weingärtner's group observed the trend of ions, in terms of decreasing protein stability, similar to the Hofmeister series[75,76,87], in which, the cation and anion series read:

$K^+ > Na^+ > [Me_4N]^+ \,||\, Li^+ > [Ch]^+ > [Et_4N]^+ \approx [Emim]^+ \approx [Gdm]^+ > [Bmpyr]^+ > [Bmim]^+ \approx [Pr_4N]^+ > [Hmim]^+ \approx [Bu_4N]^+$,

$[SO_4]^{2-} > [dhp]^- > [Ac]^- > F^- > Cl^- \,||\, [EtSO_4]^- > [BF_4]^- \approx Br^- > [OTf]^- > I^- > [SCN]^- \approx [dca]^- \gg [Tf_2N]^-$.

The double bar (||) indicates the crossover from stabilising to destabilising character. Abbreviations: $[Ch]^+$, Choline; $[Et_4N]^+$, Tetraethylammonium; $[Emim]^+$, 1-Ethyl-3-methylimidazolium; $[Gdm]^+$, Guanidinium; $[Bmpyr]^+$, *N*-butyl-*N*-methylpyrrolidinium; $[Bmim]^+$, 1-Butyl-3-methylimidazolium; $[Pr_4N]^+$, Tetra-*n*-propylammonium; $[Hmim]^+$, 1-Hexyl-3-methylimidazolium; $[Bu_4N]^+$, Tetra-*n*-butylammonium; $[dhp]^-$, Dihydrogenphosphate; $[Ac]^-$, acetate; $[EtSO_4]^-$, Ethylsulfate; $[OTf]^-$, Triflate or Trifluoromethanesulfonate; $[dca]^-$, Dicyanamide; $[Tf_2N]^-$, Bis(trifluoromethanesulfonyl)imide.

While the series suggest that weak hydration and high hydrophobicity are correlated with the decrease in $T_m$, this should not be taken as a rule of thumb. As outlined before, there are noticeable exceptions from Hofmeister behaviour[52], some ions are difficult to integrate into the Hofmeister scheme and have concentration dependent effects, with electrostatic (coulombic) effects dominating at low concentrations (< 0.5 M) and ion specific (Hofmeister) effects at high concentrations (≥ 1 M)[13,58].

Following the unique and remarkably enthalpic-entropic fingerprint of each cosolute (nonelectrolytes and salts from simple alkali halides to complex ILs), Senske et al.[13] found that certain cosolutes act through thermodynamically similar mechanisms that correlate with their chemical and physical characteristics. Differential Scanning Calorimetry (DSC) was used to determine the excess folding free energy $\Delta\Delta G_u$ of Rnase A, and its enthalpic and entropic contributions, $\Delta\Delta H_u$, $T\Delta\Delta S_u$, from the cosolute induced changes relative to the cosolute-free, buffered solution:

$$\Delta\Delta G_u^{0\prime} = \Delta G_{u,cosolute}^{0\prime} - \Delta G_{u,buffer}^{0\prime} = \Delta\Delta H_u^{0\prime} - T\Delta\Delta S_u^{0\prime}. \qquad [1.4]$$

First, for most salts / ILs with moderately (de)stabilising ion-specific effects, the trend of $\Delta\Delta G_u$ is parallel to the non-monotonic behaviour of $\Delta T_m$, with a shallow minima observed at concentrations between 0.25 and 0.5 M. Second, both $\Delta\Delta H_u$, $T\Delta\Delta S_u$ are cosolute dependent with further insight being gained by drawing an entropy-enthalpy compensation plot, $T\Delta\Delta S_u$ versus $\Delta\Delta H_u$ as shown in Fig. 1.5. The sections I and II and V and VI are the most populated due to the largely compensating enthalpic or entropic contributions, respectively.

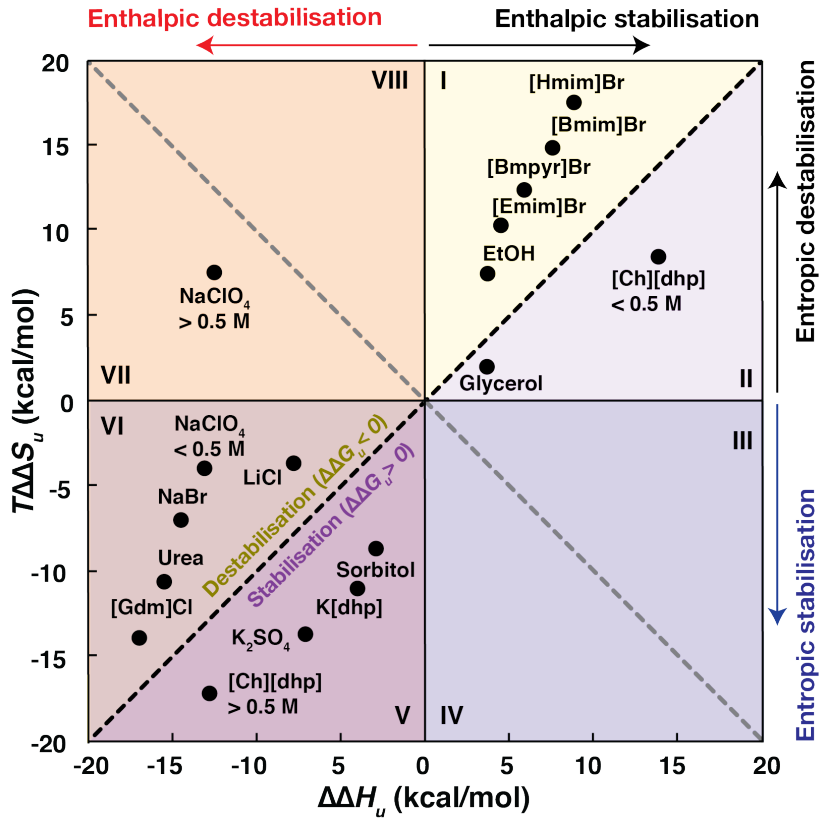



**Figure 1.5. Enthalpy-entropy compensation plot.**

The different segments correspond to different contributions of $\Delta\Delta G_u^{0'}$, $\Delta\Delta H_u^{0'}$, $T\Delta\Delta S_u^{0'}$. The different signs of these excess functions define eight different thermodynamic regimes with diverse background colours and numbers, e.g., region I indicates a destabilisation due to an increase in entropy which is not fully compensated by a positive $\Delta\Delta H_u^{0'}$. The black diagonal corresponds to a complete enthalpy-entropy compensation and separates the protein destabilising region ($\Delta\Delta G_u^{0'} < 0$) from the stabilising region ($\Delta\Delta G_u^{0'} > 0$). Positive (negative) values of $\Delta\Delta H_u^{0'}$ imply stabilisation (destabilisation) by the cosolute. The entropy term acts in the opposite direction: a positive $T\Delta\Delta S_u^{0'}$ supports destabilisation. Except where stated, distribution of cosolutes in the plot were adapted from data points at ~1 M of concentration, as reported by Senske et al[13]. The values are based on stability curves of Rnase A and calculated at $T_{m,\,buffer}$. Salts that populate near the black borderline (between I and II or V and VI fields) are omitted for clarity.

All hydrophobic ionic liquids (represented by the cations with butyl side chain or longer) fall into the yellow segment I, where $\Delta\Delta H_u$ is positive, but $T\Delta\Delta S_u$, is even more positive, leading to enthalpic stabilisation offset by entropic destabilisation (overall destabilisation, $\Delta\Delta G_u < 0$). Hydrophilic salts (e.g., alkali halides, [Ch]Cl), osmolytes (for example, sorbitol and glycerol) populate states near the borderline between sections V and VI, where both the enthalpic and the entropic contribution are negative. Depending on their magnitude, one obtains stabilising (section V) or destabilising (section VI) IL / salts. Interestingly, organic salt [Ch][dhp] reveals a strongly non-monotonous behaviour with a sharp transition from predominant enthalpic to entropic stabilisation near 0.5 M. Na[$ClO_4$] shows a non-monotonous behaviour as well, but is strongly destabilising, close to a purely enthalpic effect, and covers the segments VI and VII.

The chemical denaturants [Gdm]Cl and urea fall in section VI that is enthalpically dominated and incurs an entropic penalty (negative $\Delta\Delta H_u$ and $T\Delta\Delta S_u$). This thermodynamic fingerprint is consistent with their mechanism of denaturation *via* direct chemical interactions[88,89]. Since the unfolded ensemble has a larger preferential interaction with the denaturant than the folded state, it provides an enthalpic driving force for unfolding (that is, the unfolded state is energetically more favourable). In contrast, the hydrophobic (and destabiliser) ILs are mainly located in field I, which indicates a different molecular mechanism compared to a typical protein denaturant. Here, the increasing hydrophobicity of the cosolute promotes both the stabilising enthalpic and the destabilising entropic contributions (positive $\Delta\Delta H_u$ and $T\Delta\Delta S_u$). The observed $\Delta\Delta S_u > 0$ is explained by the reduced entropy of water in an aqueous solution of a hydrophobic cosolute that could therefore diminish the loss of entropy of the protein upon unfolding due to solvation of exposed hydrophobic groups. Other recent findings on enthalpic and entropic structure modification mechanisms of cosolutes have been reviewed elsewhere[90].

The above observations are not exclusive for positively charged proteins and can be applied for negatively charged proteins or to uncharged proteins (overall salt effect could be weakened[89]). However, as the magnitude of thermodynamic parameters of cosolute-protein interactions depends on the nature of the protein in question (e.g., isoelectric point, charge on the protein and hydrophobicity), the arrangements of ionic liquid ions into the direct or reverse Hofmeister ordering is inconclusive.

The site-specific IL-protein interactions were highlighted in the recent work of Bui-Le et al[91]. A set of analytical techniques such as circular dichroism (CD), fluorescence, ultraviolet-visible (UV/Vis), NMR and small-angle X-ray scattering (SAXS) were used to probe the specific and non-specific interactions between GFP (green fluorescent protein) and a range of $[Bmim]^+$ and $[Bmpyr]^+$ cation based ILs with $Cl^-$, $[Ac]^-$ and $[OTf]^-$ as anions. The protein is less stable in the presence of these ILs, with the anion effect being dominant, from $[Ac]^-$ to $[OTf]^-$ (the most destabiliser). Conversely, $[OTf]^-$ is the least interacting anion tested (preferential interaction with hydrophobic residues at the protein surface) but is the one that induces a structural contraction, which leads to a larger protein destabilisation. Other example is derived from theoretical studies of Lesch et al[92] concerning the influence of aqueous [Emim][Ac] on the stability of a β-hairpin peptide (the C-terminal from GB1 protein – this protein is studied in chapter 2), using atomistic molecular dynamic simulation and Kirkwood-Buff theory. The simulation work suggested that [Emim][Ac] ions induce the denaturation of the native conformation. In particular, the $[Emim]^+$ cation binds to both folded and unfolded peptide whereas $[Ac]^-$ anion exclusively binds to the unfolded conformation. These results indicate that the presence of acetate anions can be regarded as the main driving force for the denaturation of the hairpin structure. Moreover, it shows that the unfolded ensemble should not be overlooked in protein-IL studies. In line with these observations, recently, Shmool et al[93] pointed out that ILs can be used to stabilise a specific protein conformation from a heterogenous conformational ensemble. Avidin, a homo-tetrameric glycoprotein, became more compact with increasing choline-based IL concentration, interpreted as an increase of molecular confinement. Despite the use of a wide range of techniques, such as DLS (Dynamic Light Scattering), SAXS, DSF (Differential Scanning Fluorimetry), CD (Circular

Dichroism) spectroscopy, it was not possible to solve the specific structures of the avidin conformers.

A full molecular understanding of the IL-protein interaction, as well as the structural elucidation of a specific conformation can be achieved with NMR-based experiments, which is one of the main objectives of this thesis.

**Design of Bioinspired ILs**

In the search for biocompatible ILs, small charged molecules (osmolytes, metabolites) found in the intracellular milieu[94] offer a suitable choice for candidate ions since they are a natural, biodegradable and biocompatible source. *In vivo*, these small molecules are involved in regulation of diffusion, reaction rates, biochemical processes, cellular organization and protein stability[95], and can reach nearly molar concentrations[96]. In this cellular context, there is a real possibility that some of these charged molecules can behave as ion-pairs (i.e., naturally ILs) with unpredictable effects beyond Hofmeister chemistry on the protein stability. Therefore, their action needs to be studied in a broader context.

Biologically relevant charged metabolites, related with metabolic profiles of protein conformational diseases, were identified in the human metabolome database[97]. The potential of these compounds to associate as hydrated ionic liquids is related with their propensity to form stable ion-pairs in solution[98]. For the studies of proteins in solution, the cation-anion combination should also result in an ionic liquid that, in pure water, achieves proton activity buffering action (neutral pH control) from the IL itself[99]. Consequently, just the charged molecules that end in a transient ion-pair at solution at physiological pH were chosen by considering the charge of the micro-species distributed at that pH (due to the pKa values of acidic groups). This pool (Fig. 1.6a) contains a series of anions: such as, 5-oxoproline, acetate, aspartate, bicarbonate, citrate, formate, glutamate, glutarate, glycolate, glycerate, homovanillic acid, kynurenic acid, lactate, malate, maleate, malonate, oxoglutarate, propionate, pyruvate, salicylate, and succinate; and a few cations: acetylcholine, arginine, choline, diethanolammonium, dopamine, ethanolammonium, lysine, and tris(2-hydroxyethyl)methylammonium. Most of the identified anions (e.g., lactate, levulinate, carbohydrate derivatives and amino acids) have been explored as biocompatible ILs using cholinium as the cationic moiety[84,85]. Since these ions are derived from natural sources, they yield biocompatible ILs with toxicities much lower than traditional ILs. Furthermore, the presence of additional hydroxyl, carboxylic acid, and amide groups on the side chain of anions in cholinium-based ILs has been shown to improve the thermal stability of proteins and extend their long-term storage[100].

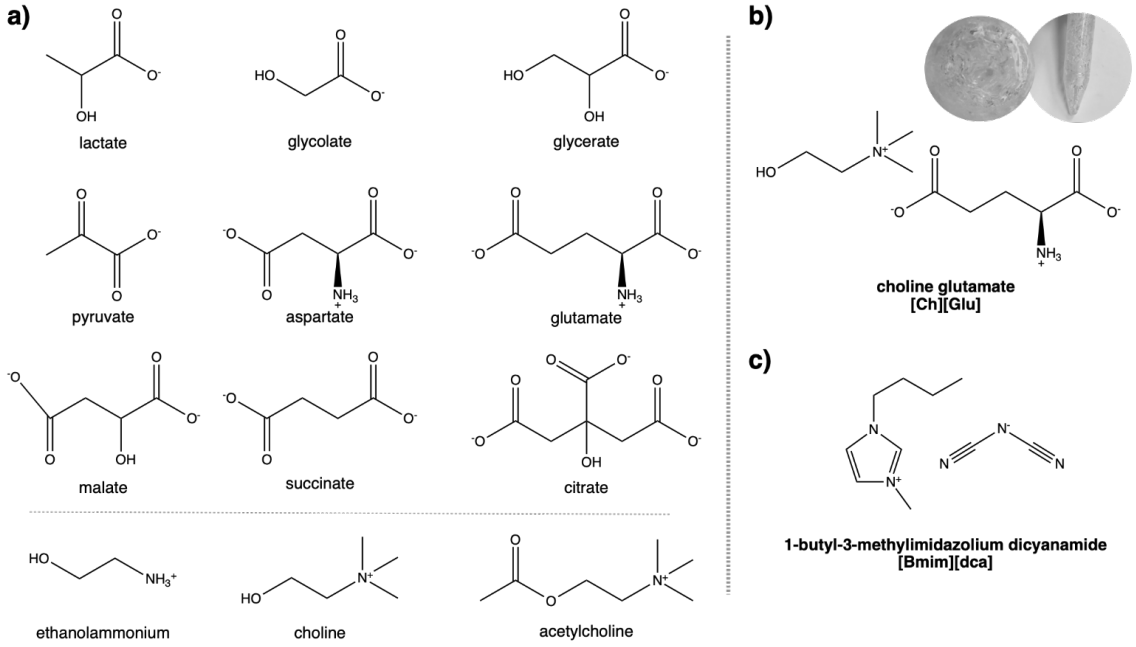



**Figure 1.6. Biocompatible ions and selected ionic liquids.**

**a)** Structure of a few anions (top) and cations (bottom) identified as potential Bio-ILs. **b)** Structure of the Bio-IL, [Ch][Glu], explored in this thesis. The inlay shows the aspect of a dry [Ch][Glu] IL, as a transparent, nearly colourless oily or glassy compound at room temperature. **c)** Structure of the [Bmim][dca] IL.

One of the most abundant intracellular metabolite found *in vivo* is the glutamate anion, for example in a exponentially growing *E. coli*, the combined concentration of metabolites has been estimated to be ~300 mM with $Glu^-$ at 96 mM[94,96]. In my master thesis[101], I synthetized and extensively explored the IL choline glutamate [Ch][Glu] (Fig. 1.6b). Results confirmed the existence of a transient ion-pair for a dilute [Ch][Glu] solution (0.5 M [Ch][Glu] in $D_2O$ at 298.2 K) with a measured intermolecular cation-anion cross-correlation distance of about 5 Å, using selective 1D $^1$H-NOESY.

In the present work, to understand the stabilising and denaturing effects of ILs on protein stability, besides the stabiliser [Ch][Glu], also the denaturing 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca], Fig. 1.6b), are explored. Our previous studies showed [Bmim][dca] leads to a generalized protein destabilisation and this effect is mainly derived from the direct and hydrophobic ion interactions with the protein but these are influenced by the ion specific ion-pair interactions which also have repercussions in ion-solvent interactions[78,79].

## Model proteins: native state to unfolded ensemble

Protein stability, in terms of structure and dynamics, was investigated in the presence of two cosolutes: the ILs [Ch][Glu] and [Bmim][dca]. To investigate these effects at a molecular level, the model proteins GB1 and drkN SH3 (Fig. 1.7) were selected based on their

distinct stability and structure. Here, I summarize the main features of these proteins. Detailed studies on GB1 and drkN SH3 can be found in chapters 2 and 3, respectively.

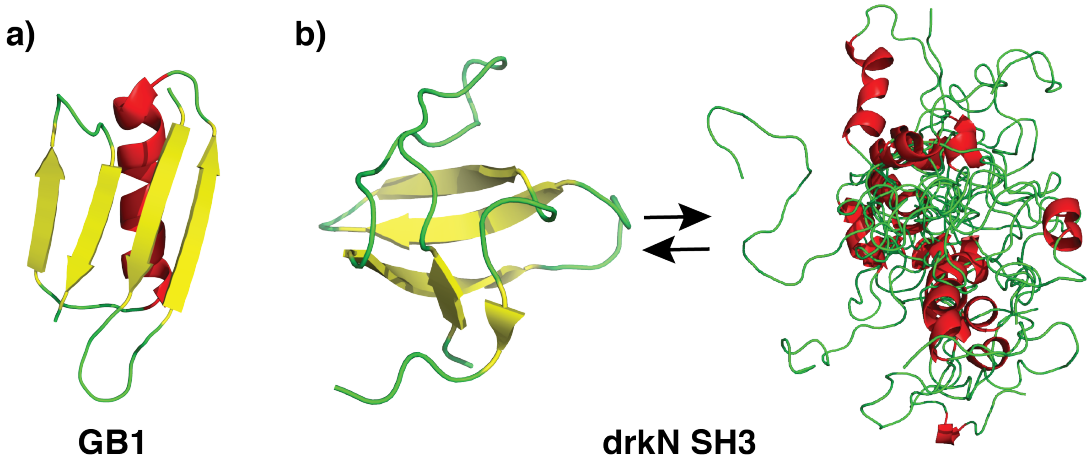



**Figure 1.7. Model proteins.**

Cartoon representation of the 3D structure of **a)** GB1 (PDB:2JSV[102]), **b)** drkN SH3 in folded state (PDB: 2A36[103]) and a few schematics of the unfolded ensemble (PED8AAC[104]). Picture was rendered by using PyMOL (coloured by secondary structure: red, helix; yellow, sheet; green, loop).

Protein G is a large multidomain cell surface protein of group *G streptococcus* and contains repeats of two or more immunoglobulin-binding domains, each comprising 55 or 56 residues. In particular, the small B1 domain (56 residues, 6.2 kDa) exhibits high thermal stability ($T_m$ of 87.5 ºC, at pH 5.4), despite the absence of any disulphide bridges. The structure is composed by a four-stranded β-sheet consisting of two antiparallel β-hairpins connected by an $\alpha$-helix, arranged in an -1, +3x, -1 topology (Fig. 1.7a)[105]. This unusual topology, coupled with an extensive hydrogen bonding network and a tightly packed and buried hydrophobic core in which 95% of the residues participate in regular secondary structure, accounts for its high stability. This protein also provides the ideal starting point for the analysis of the basic principles of folding because it folds rapidly by a simple two-state kinetics, that is, F ⇌ U, with a free energy of unfolding ($\Delta G_u$) ≈ 6 kcal/mol at 37ºC[106]. Over the past three decades, the immunoglobulin-binding domain B1 of *streptococcal* protein G, generally called GB1, has been widely used as a model system for studying protein structure, stability, dynamics, and folding kinetics[107–109].

Src homology 3 (SH3) domains are among a number of small (~60 residues), modular domains involved in cellular signalling pathways[110]. These domains mediate interactions between proteins in tyrosine kinase signalling cascades by recognizing short peptide motifs bearing one or more proline residues. The *Drosophila melanogaster* Enhancer of sevenless 2B or adapter protein Downstream of receptor kinase (Drk) is a 23-kDa protein containing a central SH2 domain surrounded by two SH3 domains[111]. *In vivo*, Drk mediates signalling between receptor tyrosine kinases (RTKs) and *ras* G-proteins by the binding to a phosphotyrosine of activated RTK with its SH2 domain and to the *ras*-guanine nucleotide exchange factor Sos (Son of sevenless) via several repeats of the consensus motif Pro-X-X-Pro at the C-terminus of Sos with its SH3 domains. The isolated N-terminal SH3 domain of the *Drosophila*

adapter protein Drk, denoted here as drkN SH3 or simply SH3, is a 6.9-kDa polypeptide (59 residues) and unlike other known SH3 domains, is marginally stable ($\Delta G_u$ ~ 1 kcal/mol) and exists in equilibrium between a folded ($F_{exch}$) and unfolded ($U_{exch}$) state, with equal populations ($pF/pU$ ~ 1, pH 6.0, 30 ºC) under non-denaturing buffer conditions[112]. Since the interconversion of the two states is slow on the NMR chemical shift timescale ($k_{ex}$ of ~ 2 $s^{-1}$, pH 6.0, 20 ºC)[113,114], distinct sets of resonances can be observed simultaneously. This allows the direct comparison between $F_{exch}$ and $U_{exch}$ under the same conditions, as well as the study of chemically or thermally denatured states, making the SH3 domain an excellent model for the study of disordered states, protein folding, and stability[103,114–116].

While certain proteins have a well-defined 3D structure, such as GB1, many others are intrinsically unstructured or possess significant unfolded domains (e.g., SH3) under physiological conditions. The drkN SH3 domain fold ($F_{exch}$) is highly conserved and forms an incomplete five-stranded antiparallel β-barrel (Fig. 1.7b)[103]. The long RT loop between β1 and β2 has a fairly regular structure with the possibility to form two additional irregular twisted β-hairpins. The long and flexible N-Src loop, and the distal loop connect strands β2 and β3, and β3 and β4, respectively. Strands β4 and β5 are connected by a short three-residue $3_{10}$ helix. Mimicking *in vivo* conditions, this domain is fully stabilised upon binding to the positively charged proline-rich peptide derived from the protein Sos[117]. Such observation provides a good example of protein stabilisation due to their binding with its biological target. Although the drkN SH3 unfolded state ($U_{exch}$) is an ensemble of rapidly interconverting conformers (Fig. 1.7b), there is considerable evidence suggesting that it possesses significant non-random or residual structural properties (i.e., non-native helical propensity)[104,115,117,118], in contrast to a full random-coil behaviour found in chemically denatured proteins. Furthermore, the observation that both bacterial and human Hsp70 chaperones interact with drkN SH3 by selecting their unfolded state, and thus generating a highly conformational heterogeneous chaperone-bound ensemble[119,120], supports the notion that the SH3 unfolded ensemble is representative of a physiologically relevant disordered state.

## Biomolecular NMR spectroscopy

In this last section, some fundamental principles of NMR spectroscopy applied to proteins are described through a simple HSQC experiment. Backbone and side chain assignment as well as protein dynamics methods are reviewed.

### HSQC and related experiments

In protein NMR spectroscopy, the assignment of backbone amides is a prerequisite for studying protein structure and dynamics, therefore $^{1}H$-$^{15}N$ HSQC (Heteronuclear Single-Quantum Correlation) spectroscopy is one of the most important and common experiments for proteins and peptides which can quite easily be enriched in $^{15}N$. This experiment, originally proposed by Bodenhausen and Ruben[121] in the early 1980s, is widely used for recording one-bond correlation spectra between $^{15}N$ (or $^{13}C$) and its directly coupled proton. Typically, in an $^{1}H$-$^{15}N$ HSQC spectrum, one signal is expected for each bonded N-H pair (roughly one peak per residue), as exemplified in Figure 1.8, with the exception of proline which has no amide-hydrogen due to the cyclic nature of its backbone. The signals also include the indole side chain $N^{\varepsilon}$-$H^{\varepsilon}$ group of tryptophan, and the asparagine and glutamine $N^{\delta}$-$H^{\delta 2}$ and $N^{\varepsilon}$-$H^{\varepsilon 2}$ side chains, respectively. In samples prepared at low pH (< 6.5) the arginine $N^{\varepsilon}$-$H^{\varepsilon}$ peaks can also be visible (if exchange is sufficiently slow), but because the $N^{\varepsilon}$ chemical shift (~ 85 ppm) is outside the region usually recorded (spectral width of ~20 to 36 ppm, centred at ~118 ppm), the peaks appear folded (i.e., a peak that appears at the position mirrored about the spectrum boundary). At low pH and temperature, the Lys $N^{\zeta}$-$H^{\zeta}$ and Arg $N^{\eta}$-$H^{\eta}$ groups (while broadened and of low intensity) can be visible but are also folded.

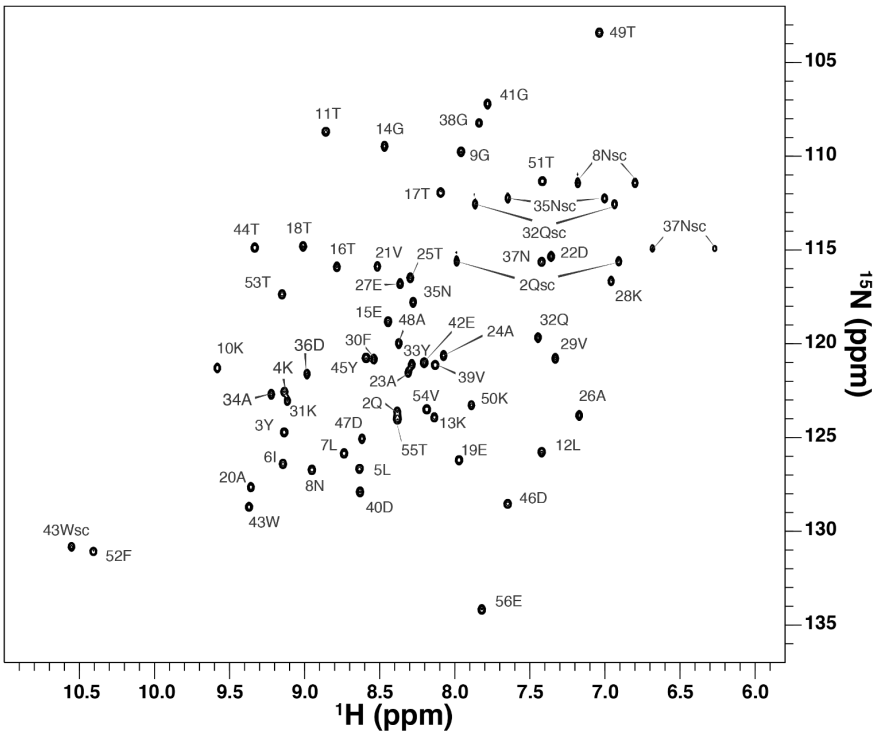



**Figure 1.8. $^{1}H$-$^{15}N$ HSQC spectrum of protein GB1.**

$^{1}$H-$^{15}$N heteronuclear correlation spectra (HSQC) of $^{15}$N-labelled GB1, recorded using standard Bruker HSQC pulse sequence with sensitivity-enhanced pulsed field gradient[122]. Backbone amide assignments are shown, as well as those of tryptophan, glutamine, and asparagine side chains [sc] in which $NH_2$ correlations are indicated by the label connecting the two non-equivalent proton resonances. The pattern of relatively well-dispersed peaks in chemical shift is indicative of a folded protein.

As each protein has a unique pattern of signal positions, the $^{1}$H-$^{15}$N HSQC spectrum is often referred to as the "fingerprint" of a protein (Fig. 1.8) and it is used for mapping protein-ligand or protein-protein interactions. In the $^{1}$H- $^{13}$C HSQC each C-H will give a cross-peak which also leads to a more crowded spectrum.

Given the importance of HSQC experiments and their association in other more complex schemes, a simple HSQC pulse sequence (Fig. 1.9) is further explained below[123–125]. Broadly speaking, the sequence works by first transferring magnetization from the *I* spin ($^{1}$H) to the *S* spin (typically $^{13}$C or $^{15}$N) using an INEPT-like sequence (Insensitive Nuclei Enhanced by Polarization Transfer – uses the magnetization of high-γ nuclei to enhance the weak NMR signals of low-γ nuclei). The *S* spin magnetization then evolves for $t_1$, during which time it acquires a frequency label according to the offset of *S* (with the centrally placed *I* spin 180°$_y$ pulse refocusing the evolution of the coupling). Finally, this magnetization is transferred back to *I*, where it is observed. The resulting spectrum thus has peaks centred at the offset of the *S* spin in $\omega_1$ dimension, and at the offset of the *I* spin in the $\omega_2$ dimension.

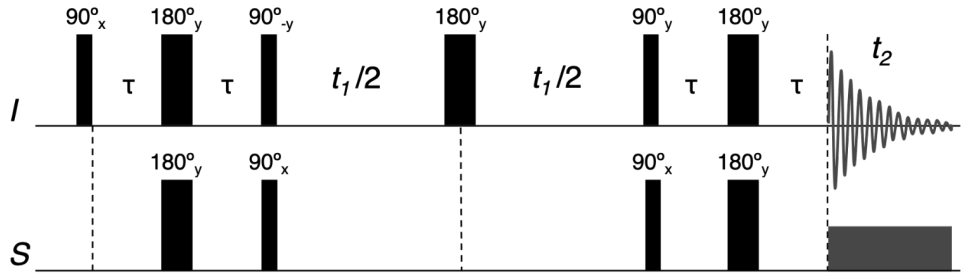



**Figure 1.9. The HSQC pulse sequence.**

The optimum value for τ is $1/(4J_{IS})$ (for one bond $^{1}$H-$^{15}$N $J_{NH}$= 92 Hz, and for one bond $^{1}$H-$^{13}$C $J_{CH}$= 140 Hz). Decoupling during $t_2$ is achieved by using GARP-1, WALTZ-16, or other decoupling sequences.

In the following, only the product terms-operator terms that illustrate the principles of the method are retained and to keep in a simple way, the coefficients of the operator terms are not shown. The first element of an HSQC is an INEPT sequence that is used to transfer *I* spin polarization ($I_z$) into antiphase heteronuclear single-quantum (SQ) coherence ($2I_zS_y$),

$$\boldsymbol{I_z + S_z} \xrightarrow{90^{o}I_x} -I_y + S_z \xrightarrow{\tau,180^{o}I_y,180^{o}S_y,\tau} 2I_xS_z - S_z \xrightarrow{90^{o}I_{-y},90^{o}S_x} -2I_zS_y + S_y, \qquad [1.5]$$

in which, for example, $90^{o}I_x$ represents a pulse with a rotation angle 90° and phase *x* operating on the *I* spins, τ is $1/(4J^{1}_{IS})$, and $J^{1}_{IS}$ is the one bond heteronuclear scalar coupling constant. The $S_y$ term is normally removed by phase cycling. The INEPT block is followed by a $t_1$

evolution period, in the middle of which a $180°_y$ pulse is applied to spin *I* that forms a spin echo. The evolution arising from *J*-coupling of *I* and *S* will be refocused, leaving only evolution under the *S*-spin chemical shift together with the effects of the 180° pulse on spin *I*. Thus, at the end of the evolution period, the state of the system will be

$$-2I_zS_y + S_y \xrightarrow{\frac{t_1}{2},180°I_y,\frac{t_1}{2}} 2I_zS_y cos\Omega_s t_1 - 2I_zS_x sin\Omega_s t_1. \quad [1.6]$$

The final element of HSQC is simply a reversal of the initial INEPT sequence (i.e., all the pulses and delays applied in reverse order, with the phase of the 90° pulses inverted), except that the first pulse was removed. After the 90° pulses to both spins, these transfer the term $2I_zS_y$ to $2I_xS_z$ which is anti-phase magnetization on the I spin, and the term $2I_zS_x$ becomes $2I_xS_x$ which is unobservable multiple-quantum coherence:

$$2I_zS_y cos\Omega_s t_1 - 2I_zS_x sin\Omega_s t_1 \xrightarrow{90°I_y,90°S_x} 2I_xS_z cos\Omega_s t_1 - 2I_xS_x sin\Omega_s t_1. \quad [1.7]$$

Next follows another spin-echo in order to allow the anti-phase terms to become in phase. As before, during this spin echo the coupling evolves, but the offset does not, so $2I_xS_z$ term becomes $I_y$. At the end of the HSQC pulse sequence, we will have:

$$2I_xS_z cos\Omega_s t_1 - 2I_xS_x sin\Omega_s t_1 \xrightarrow{\tau,180°I_y,180°S_y,\tau} \boldsymbol{I_y cos\Omega_s t_1} - 2I_xS_x sin\Omega_s t_1, \quad [1.8]$$

where the $I_y cos\Omega_S t_1$ term represents the observable spin *I*-magnetization that is labelled by the chemical shift of the heteronucleus during $t_1$ and is detected during $t_2$, which results in a single peak at $\Omega_I$ in the $\omega_2$ dimension, and $\Omega_S$ in the $\omega_1$. If the signal is now observed while broadband decoupling is applied to the *S* spin, just the term in $I_y$ contributes, thus the other term that represents multiple-quantum coherence is unobservable during $t_2$. The HSQC pulse sequence analysis step-by-step can be found originally in Palmer et al[126].

Interestingly, simply by nesting one HSQC sequence inside another, the frequencies of three different nuclei (e.g., $^1$H, $^{15}$N, and $^{13}$C) are possible to correlate. Two very important examples of this approach are the complementary three-dimensional experiments HNCA[127] and HN(CO)CA[128]. The HNCA experiment correlates the amide $^1H^N$ and $^{15}$N chemical shifts with the intra-residue $^{13}C^\alpha$ shift, by making use of the relatively small one-bond $^{15}$N–$^{13}C^\alpha$ scalar coupling (8–13 Hz) to establish correlations between the $^{15}$N and $^{13}C^\alpha$ spins. In addition, this experiment also provides sequential connectivity by transferring coherence from the $^{15}$N spins to the $^{13}C^\alpha$ of the preceding residue via the inter-residue two-bond $^{15}$N–$^{13}C^\alpha$ scalar coupling, which can be as large as 7 Hz. In contrast, the HN(CO)CA experiment provides sequential correlations between the amide $^1$H and $^{15}$N chemical shifts of one amino acid residue and the $^{13}C^\alpha$ chemical shift of the preceding residue alone, by transferring coherence via the intervening $^{13}$CO spin. By comparing HNCA and HN(CO)CA spectra, it is possible to achieve a complete assignment of the amide and C$\alpha$ resonances (Fig. 1.10)[129].

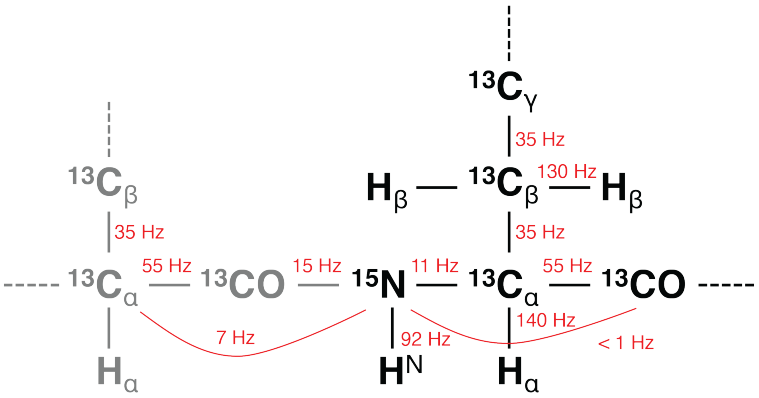



**Figure 1.10. Scalar coupling constants between the different nuclei in a polypeptide.**

Spin system of the peptide backbone and the size of the $^{1}J$ and $^{2}J$ coupling constants that are used for magnetization transfer in $^{13}$C-, $^{15}$N-labelled proteins[129].

The HNCA and HN(CO)CA experiments feature multiple "out-and-back" INEPT transfers (in which the $^{1}$HN magnetization excited initially in the pulse sequence also is detected during $t_3$) from $^{1}$H to $^{15}$N (and vice versa) and from $^{15}$N to $^{13}$C (and vice versa). The CO spins are manipulated separately from C$\alpha$ using 180° refocusing and frequency-selective $^{13}$C-pulses (since $^{13}$CO or $^{13}$C$^{\alpha}$ spins resonate at ~175 and ~55 ppm, respectively). In the HN(CO)CA experiment, INEPT transfer occurs from $^{1}$H to $^{15}$N, then from $^{15}$N to $^{13}$CO, and finally from $^{13}$CO to $^{13}$C$^{\alpha}$, before reversing the whole process. The magnetization is not allowed to evolve at the $^{13}$CO frequency, but is instead immediately transferred to the next spin, and so only three separate frequency dimensions are recorded.

### Structural characterisation: backbone and specific side chain assignment

Backbone correlations

A complete assignment of the backbone resonances of a protein forms the basis for further detailed conformational studies. Thirty years ago, Ad Bax and coworkers[127] proposed the earliest three-dimensional triple resonance experiments of uniformly $^{13}$C- and $^{15}$N-enriched proteins which have evolved tremendously over the years[130] thanks to advances not only in instrumentation but also due to an improvement in pulse sequence design and isotope labelling schemes. These pulse sequences can be used to successfully solve backbone assignment using one-bond and two-bond ($^{1}J/^{2}J$) scalar coupling interactions (Fig. 1.10), without recurring to NOE information. As $^{1}J/^{2}J$ couplings are generally larger than the linewidth ($J > \Delta\nu_{1/2}$) of the nuclei under consideration, they yield fast coherence transfers that can compete with the loss of magnetization as a result of relaxation during the pulse sequence.

Typically, 3D triple resonance HNCO, HN(CA)CO, CBCA(CO)NH or HN(CO)CACB, and HNCACB are sufficient to provide backbone resonance assignments (i.e., $^{1}$H$^{N}$, $^{15}$N, $^{13}$C$^{\alpha}$, $^{13}$C$^{\beta}$ and $^{13}$CO) for a protein with molecular weight (MW) below 20 kDa[129], as shown in Table 1.1. For larger proteins (MW > 20 kDa), the sensitivity of all heteronuclear triple resonance experiments declines as a result of increasing $T_2$ relaxation rates. For such proteins, fractional (~ 75 %) or complete (100 %) deuteration was proposed[131] since the dipolar coupling in a C–D fragment compared to a C–H fragment is scaled down because of the lower gyromagnetic ratio, $\gamma(^{2}\text{H}) = 1/6.5\ \gamma(^{1}\text{H})$, thus dilution of protons by deuterons leads to longer relaxation

times for the $^{13}C$ spins. For a deuterated protein, the combination of 3D or 4D experiments with $^{2}H$ decoupling yields dramatic sensitivity improvements when compared to a fully protonated sample[132]. In addition, minimization of $T_2$ relaxation losses during multi-step transfer experiments can be obtained by using a TROSY-scheme (transverse relaxation optimized spectroscopy)[133].

**Table 1.1. Pulse sequences typically used for backbone protein assignment.**

| **Experiment** | **HNCO** | **HN(CA)CO** | **CBCA(CO)NH** | **HNCACB** |
|---|---|---|---|---|
| **Nuclei observed** | H(*i*)- N(*i*), CO(*i*-1) | H(*i*), N(*i*), CO(*i*), CO(*i*-1) | H(*i*), N(*i*), Cα(*i*-1), Cβ(*i*-1) | H(*i*), N(*i*), Cα(*i*), Cβ(*i*) Cα(*i*-1), Cβ(*i*-1) |
| **Relative S/N [%]** | 100 | 13/4 α/β | 13/9 α/β | 1.3/0.5 α/β(*i*-1) |

The parentheses in the experiment name indicate a nuclear spin that is involved in the coherence transfer pathway, but not frequency labelled. The relative sensitivity of backbone assignment experiments was compared to the HNCO experiment[129].

The HNCO experiment correlates the amide $^{1}H^{N}$ and $^{15}N$ chemical shifts of one amino acid with the $^{13}CO$ chemical shift of the preceding residue, by using the one-bond $^{15}N$–$^{13}CO$ *J* coupling (~ 15 Hz) to establish the sequential correlation[127] (Fig. 1.11a, b).

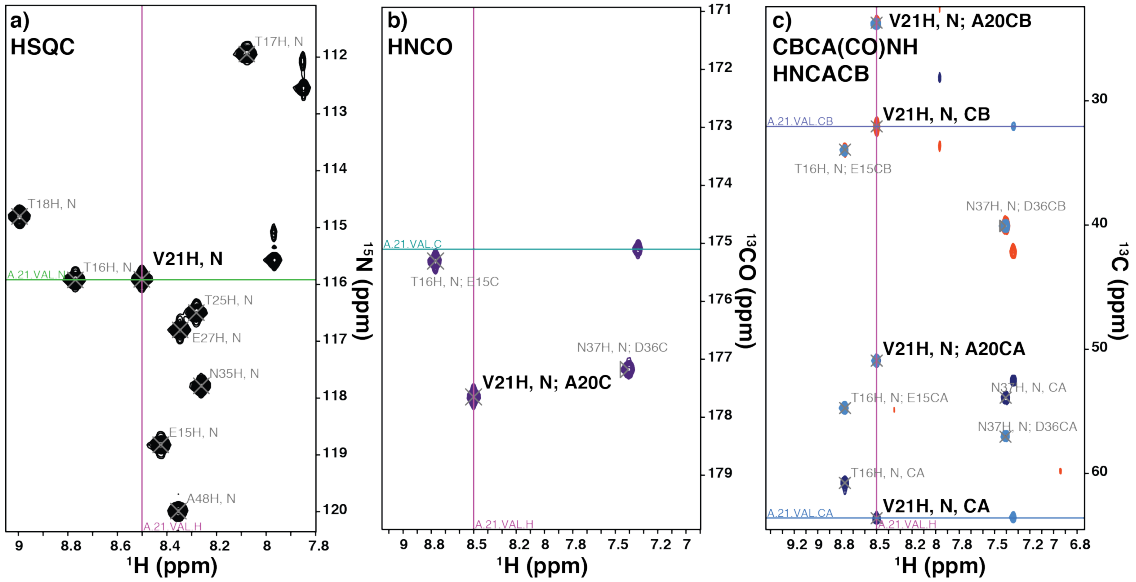



**Figure 1.11. Spectra typically used for backbone protein assignment.**

Identifying the CO(*i*-1), Cα(*i*), Cβ(*i*), Cα(*i*-1), and Cβ(*i*-1) backbone resonances of the selected amide group, Val21 of GB1. **a**) Part of the $^{1}H$-$^{15}N$ HSQC spectrum, **b**) plane of HNCO spectrum and **c**) plane of CBCA(CO)NH and HNCACB (blue, positive; orange, negative sign peaks) spectra of GB1 at the amide frequency of Val21 ($^{15}N$ δ = 115.92 ppm).

The HN(CA)CO experiment provides intra-residue correlations between the amide $^{1}H$, $^{15}N$, and $^{13}CO$ chemical shifts by using the one-bond $^{15}N$–$^{13}C^{\alpha}$ and $^{13}C^{\alpha}$–$^{13}CO$ *J* couplings to transfer coherence. In addition, this experiment can also provide sequential connectivity from the $^{15}N$ spins to the $^{13}CO$ of the preceding residue via the inter-residue two-bond $^{15}N$–$^{13}C^{\alpha}$ J coupling. When used in conjunction with the HNCO experiment, which gives the sequential correlations only, the HN(CA)CO experiment provides a method for sequentially assigning the amide $^{1}H$, $^{15}N$, and $^{13}CO$ resonances, and also to resolve some ambiguities associated with chemical shift degeneracy and/or overlap in other spectra (e.g., $^{13}C^{\alpha}$ spins). The CBCA(CO)NH (or HN(CO)CACB) experiment correlates both the $^{13}C^{\alpha}$ and the $^{13}C^{\beta}$ resonances of an amino acid residue with the amide $^{1}H$ and $^{15}N$ resonances of the following residue. The out-and-back HNCACB experiment correlates the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ resonances with the amide $^{1}H$ and $^{15}N$ resonances of the same residue and the amide resonances of the succeeding residue via the $^{1}J_{C\alpha N}$ and $^{2}J_{C\alpha N}$ couplings, respectively. In this spectrum (Fig. 1.11c), the correlation peaks for $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ spins have opposite signs which facilitates their distinction.

Although not performed in this thesis, the assignment of all side-chain aliphatic $^{1}H$ and $^{13}C$ resonances could be carried out with HCCH-TOCSY (via $^{1}J_{CH}$ and $^{1}J_{CC}$ couplings) in combination with other experiments (e.g., CBCA(CO)NH, HBHA(CO)NH, (H)CC(CO)NH-TOCSY, H(CC)(CO)NH-TOCSY). Detailed explanation for the experiments described here can be found in review of Sattler et al.[129] and textbook of Cavanagh et al. *Protein NMR spectroscopy*[132].

In my work, although conventional pulse schemes were used to collect 3D triple-resonance NMR spectra (Table 1.1), nonuniform sampling (NUS) was used in order to reduce measurement time and improve spectrum resolution[134]. Since NMR spectra consist of discrete resonances that are often quite sparse (i.e., the number of spectral components is many orders of magnitude smaller than the total number of data points in the final multi-dimensional spectrum). Instead of relying on simple discrete Fourier transformation (FT), the frequency domain spectrum can be reconstructed from a much smaller number of time domain data points by using iterative or other non-linear methods (NUS "skips" some fraction of data that would usually be measured). Generally, spectra of low dynamic range (sparse data), including most common triple resonance experiments, where peaks have similar intensities and spectra contain fewer peaks, are very suitable for NUS and allow low density sampling. It only requires an optimized sampling schedule and a reconstruction method (e.g., SMILE – Sparse multidimensional iterative lineshape-enhanced[135]).

Specific side chain resonances

The assignment of nuclei at the side chain end groups is quite important for understanding molecular folding, molecular interactions (such as electrostatic forces), and catalytic activities. Negatively charged groups (carboxyl groups of Glu and Asp residues, Fig. 1.12a), and positively charged groups (imidazole rings of His residues, amino groups of Lys residues, and guanidino groups of Arg residues, Fig. 1.12b) can be followed as sensitive probes of local electrostatic environment[136,137].

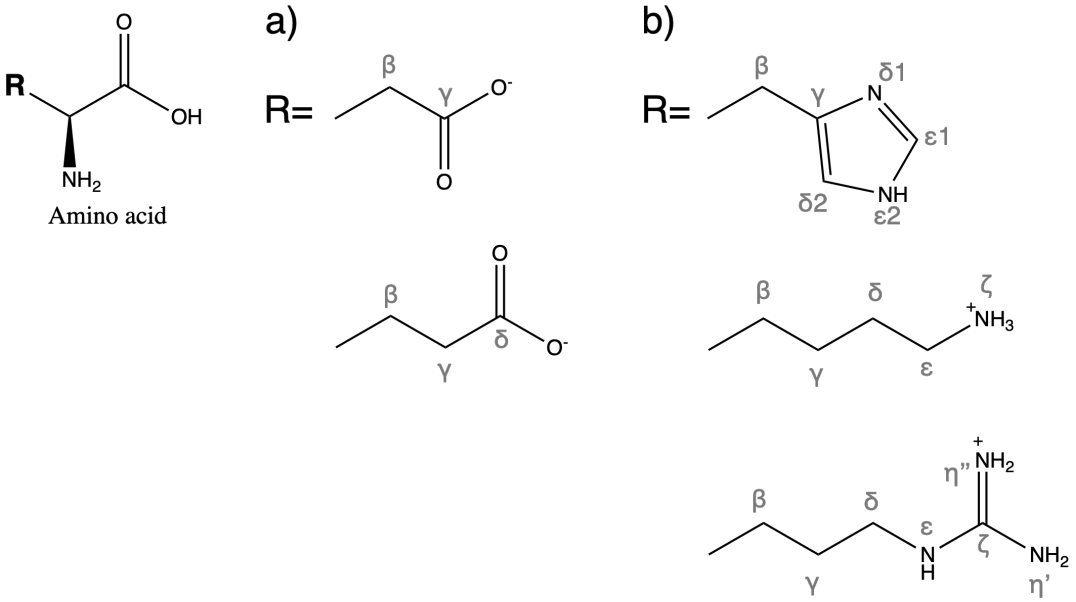



**Figure 1.12. Schematic diagram for aminoacids with a) negative and b) positive charged side chains.**

Aminoacids with charged side chains and charges according with physiological pH. **b**) Asp and Glu with their side chains $pk_a$ values of ~ 3.7 and ~ 4.2, respectively. **b**) Hist, Lys, Arg, with their side chains $pk_a$ values of ~ 6.0, ~ 10.7, ~ 12.1, respectively. The schematic diagram shows the nomenclature used to describe atoms along sidechain.

Pulse sequences have been introduced to identify specific polar and charged side chain spin systems (Asp, Glu, Asn and Gln)[138,139]. These pulse sequences are derived from backbone experiments by way of 3D HCACO experiment[127] and are optimized for side chain spin systems. A $^{1}H^{\beta/\gamma}$ -$^{13}CO^{\gamma/\delta}$ H(C)CO spectrum has been recorded with observable peaks from side chain carboxyl carbon nuclei (Asp and Glu) and carbonyl amide carbon nuclei (Asn and Gln residues) by an adapted 2D version of constant time-H(CA)CO (where $J_{C\alpha CO}$ coupling is eliminated from the spectrum) using as offset the side chain $^{13}CO$ chemical shift average[138]. Typically, these carbon nuclei are assigned from side chain methylene groups through single $^{13}C$-$^{13}CO$ *J* connectivities (HCCH-TOCSY). Other experiments, namely 2D $H(C)CO_2$ and 2D $HCCO_2$[139] are specifically select for carboxylates ($^{13}CO_2$), i.e., acidic side-chain spin systems are observed by selective correlation of $^{13}CH_2^{\beta}$ and $^{13}CO_2^{\gamma}$ of Asp, and $^{13}CH_2^{\gamma}$ and $^{13}CO_2^{\delta}$ of Glu (Fig. 1.12a). Another pulse sequence correlates Asp/Glu side-chain carboxyl $^{13}C$ chemical shifts with the backbone amide proton $^{1}H^{N}$ shift of the following residue, this is particularly useful when $^{1}H^{\beta/\gamma}$ chemical shift overlap is limiting[136]. Based on $^{1}H^{\beta/\gamma}$ -$^{13}CO^{\gamma/\delta}$ H(C)CO spectrum, Asx/Glx side chain carboxyl and carbonyl $^{13}C$ nuclei relaxation methods have been reported[140,141].

For histidine's, the sidechain $^{15}N^{\varepsilon 2}$ and $^{15}N^{\delta 2}$ chemical shifts are correlated with the nonexchangeable ring protons (Fig. 1.12b) via two- and three-bond $J_{NH}$ coupling[142]. The canonical $^{15}N$ chemical shifts for the tautomeric and cationic states of the side chain thus provides a direct estimation of the relative populations of each conformer.

With regard to probing arginine and lysine sidechains, conventional methods (2D $^{1}H$-$^{15}N$ HSQC, HMQC and TROSY) are often impeded by rapid exchange of the detected protons with the bulk solvent, which can lead to extensive line-broadening and effectively undetectable signals at and above physiological pH. Indeed, just at low pH and temperature arginine $^{1}H^{\varepsilon}$-$^{15}N^{\varepsilon}$ and lysine $^{1}H^{\zeta}$-$^{15}N^{\zeta}$ may become detectable (i.e., $k_{ex}^{water}$ is slow enough to permit their detection by $^{1}H$ NMR), whereas arginine $^{1}H^{\eta}$-$^{15}N^{\eta}$ cross-peaks remain broadened by slow rotation of the guanidino head group[143] . Direct observation of $NH_3^+$ groups is also possible at low pH in the context of hydrogen or ion pairs, specifically using a $^{1}H$-$^{15}N$ HISQC (heteronuclear in-phase single quantum coherence spectroscopy) experiment optimized for the $N^{\zeta}$ groups[144]. This experiment leads to a significant sensitivity gain because it is independent of scalar relaxation in the $^{15}N$ dimension via rapid water exchange and as a result exhibits strikingly sharper $^{15}N$ line shapes and higher intensities for $NH_3$ cross-peaks than either HSQC or HMQC. Further details can be found in Zandarashvili et al[145].

To alleviate the pH and temperature problem to detect Arg $N^{\varepsilon}$ (~ 85 ppm) and Lys $N^{\zeta}$ (~ 33 ppm) nuclei, André et al.[137] introduced an elegant approach relying on detection of aliphatic protons. As non-exchanging carbon-bound protons ($^{1}H^{\delta}$-$^{13}C^{\delta}$ of Arg and $^{1}H^{\varepsilon}$-$^{13}C^{\varepsilon}$ of Lys) are used for excitation and detection, the H2CN experiment can be performed over the full pH range and at ambient or elevated temperature without signal loss or broadening due to exchange with solvent water. 2D H2(C)N experiments are preferred over 2D H2C(N) since potential spectral overlap of Arg $^{1}H^{\delta}$ and Lys $^{1}H^{\varepsilon}$ is easily resolved along the $^{15}N$ dimension, whereas the chemical shift separation of $^{13}C^{\delta}$ in Arg and $^{13}C^{\varepsilon}$ in Lys is much smaller. In addition, since the charge of the guanidinium group is delocalized over the entire $N^{\varepsilon}$-$C^{\zeta}$-$(N^{\eta}H_2)_2$ moiety, the H2C(N) experiment was extended with an additional transfer step to record Arg $^{13}C^{\zeta}$ chemical shifts. Whilst this method circumvents the rapid exchange, it is less beneficial for nuclear spin relaxation measurements, where the two bound protons enhance the transverse relaxation of the aliphatic carbons. Based on $^{13}C^{\zeta}$-detection, Hansen and co-workers have been developing pulse schemes to obtain $^{13}C^{\zeta}$–$^{15}N^{\varepsilon}$ spin pairs information[146,147], and $^{13}C^{\zeta}$–$^{15}N^{\eta}$ HDQC (double-quantum correlation) spectra[148]. In this last approach, the $^{15}N^{\eta}$ chemical shifts (~ 71 ppm) characterisation is possible through a creation of double-quantum coherence that is insensitive to the rotation about the $C^{\zeta}$ –$N^{\varepsilon}$ bond and thus unaffected by the line broadening resulting from this exchange process.

**Exploring molecular motions: from picoseconds to seconds timescale**

In structural biology, motional fluctuations of biomolecules are crucial to understand their function[149]. That is, proteins are more accurately described by an equilibrating set of time-dependent fluctuations in structure. However, it is important to recognize that a protein conformation will be time-averaged within the time window and ensemble-averaged within the set of molecules under observation. The shape of this energy landscape is unique for each protein and may change with conditions (e.g., ionic strength, ligand binding, pH, pressure, temperature)[74].

In terms of the energy landscape theory (mentioned above), as protein samples different conformations in dynamic equilibrium, the population of each state is determined by its

relative free energy, while interconversion rates are determined by inter-state energy barriers. Considering a biomolecule that exchanges between the native state, A, and other conformation that is thermally accessible, B, with population $p_A$ and $p_B$,

$$A \underset{k_{BA}}{\overset{k_{AB}}{\rightleftharpoons}} B, \qquad [1.9]$$

where equilibrium between populations is established and maintained by kinetic processes that satisfy $p_B/\ p_A = k_{AB}/k_{BA}$. The sum of all populations is unity ($p_A + p_B = 1$) and sum of forward and reverse rates, $k_{BA} + k_{AB} = k_{ex}$ is defined by the energy barriers along the path connecting the two states and contains temporal information. The population of the tier-0 states A and B, are defined as Boltzmann distribution based on their difference in Gibbs free energy ($\Delta G_u^{0\prime}$),

$$\left(\frac{p_B}{p_A}\right)_{equilibrium} = exp\left(\frac{-\Delta G_{AB}^{0\prime}}{k_B T}\right), \qquad [1.10]$$

where $k_BT$ is the product of the Boltzmann constant and the absolute temperature. The exchange rate constants ($k_{AB}$ and $k_{BA}$) are related to the barriers between these states ($\Delta G^{0\prime\ddagger}$) through the Arrhenius equation [$k = A \exp(-\Delta G^{0\prime\ddagger}/RT)$, where A is a reaction-specific quantity and $\Delta G^{0\prime\ddagger}$ the Gibbs free energy of activation] that is recast by transition state theory to

$$k_{AB\ or\ BA} = \kappa \frac{k_B T}{h} exp\left(\frac{-\Delta G_{AB\ or\ BA}^{0\prime\ddagger}}{RT}\right), \qquad [1.11]$$

where $k_{AB\ or\ BA}$ is the interconversion rate constant at absolute temperature $T$, $\Delta G_{AB\ or\ BA}{}^{0\prime\ddagger}$ is the mean difference in energy between the conformations at the saddle point of the reaction and the ground state at $T$, $R$ is the gas constant, $h$ is Planck's constant, $k_B$ is the Boltzmann constant, and $\kappa$ is a transmission coefficient related to the probability that the reaction proceeds to product from the transition state.

An array of NMR dynamics techniques is used to characterize protein motions on timescales spanning many orders of magnitude (Fig. 1.13) and have been extensively reviewed by Palmer and co-workers[150–152]. The choice of which technique to apply is based on the time scale of the kinetic or conformational dynamical process[153,154]. Despite all the methods mentioned, here, three distinct methods are underlined: spin relaxation (ps-ns time scale), magnetization-/ZZ-exchange and hydrogen-deuterium (H/D) exchange (~ s).

Faster processes in the picosecond-to-nanosecond (ps-ns) range have small free-energy barriers ($\ll RT$) and typically can be attributed to more local intramolecular motions. These ps-ns timescale motions are quite important since contribute significantly to the overall conformational entropy[155]. Kinetic processes in microsecond-to-second range includes major conformational changes, molecular recognition, folding and allosteric events. Significant structural differences in states interconversion are commonly identified (e.g., partial unfolding) as the exchange that involves conformers are separated by larger free-energy barriers. These slower processes often play significant roles in biology because transiently populated conformational states can have distinct functionalities from the native state[32,156–158].

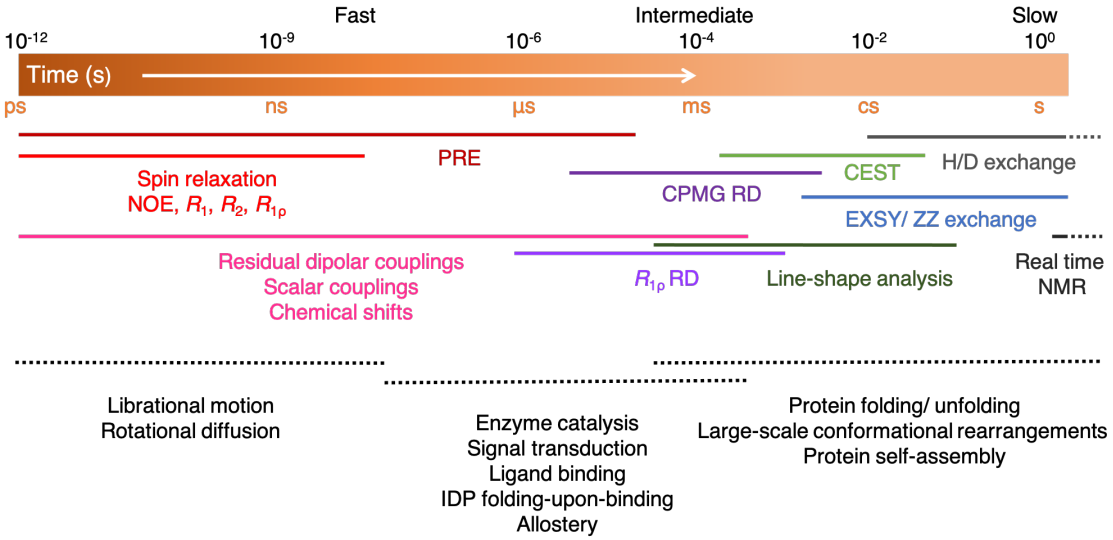



**Figure 1.13. Survey of the NMR methods for studying biomolecular dynamics.**

A wide range of NMR techniques are available to probe protein motions (bottom) on a broad range of time scales (top). The fast, intermediate, and slow time scale corresponds to the different chemical exchange regimes in NMR. Abbreviations: CEST, chemical exchange saturation transfer; CPMG, Carr-Purcell-Meiboom-Gill; IDP, intrinsically disordered protein; H/D exchange, hydrogen–deuterium exchange; NOE, nuclear Overhauser enhancement; PRE, paramagnetic relaxation enhancement; RD, relaxation dispersion; ZZex, magnetization exchange. Adapted from Sekhar and Kay[157].

For each spin probe, analysis of dynamic NMR experiments (Fig. 1.13) provides site-specific motional (thermodynamic and kinetic) as well as structural information (chemical shift differences between reporter spins in each of the exchanging sites, $\Delta\omega = \omega B - \omega A$)[150]. The NMR timescale for conformational exchange is thus defined by $k_{ex}$ relative to the chemical-shift timescale ($\Delta\omega$), whether an exchange process is referred to as fast ($k_{ex} \gg \Delta\omega$), intermediate ($k_{ex} \approx \Delta\omega$), or slow ($k_{ex} \ll \Delta\omega$). Figure 1.14 illustrates the classic two-site chemical exchange scenario in which a probe (nucleus) attached to a molecule reports on conformational exchange between a pair of states. Provided that the NMR resonance frequency of the probe in each of the two states is different, then large spectral changes are observed as the exchange rate varies. At low exchange frequencies, a pair of peaks derived from the individual states is observed in NMR spectra, whereas the signals broaden and merge into a single resonance at higher rates.

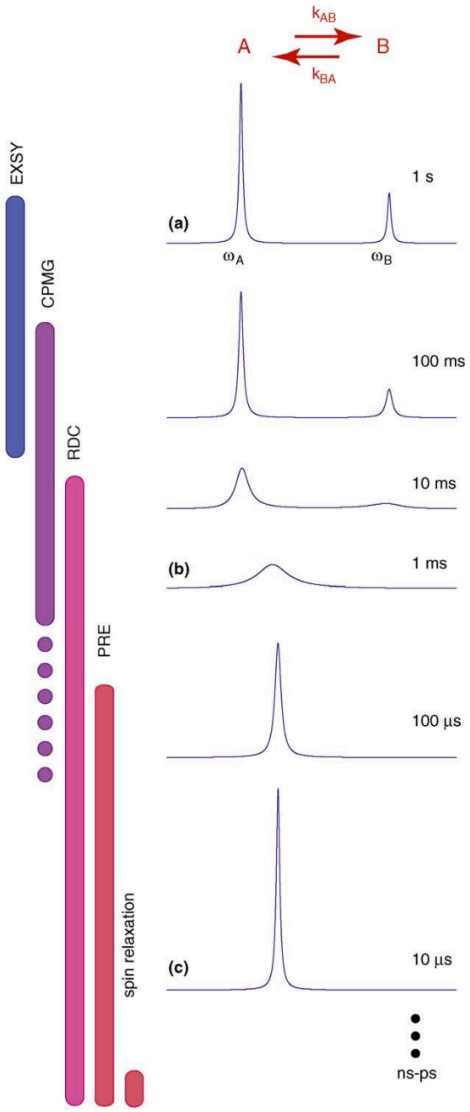



**Figure 1.14. Overview of NMR dynamic experiments with the classic two-site chemical exchange.**

Simulated one-dimensional NMR spectra for nuclei exchanging between two distinct chemical environments, populated in the ratio 0.75: 0.25 ($p_A$: $p_B$) and associated with chemical shifts $\omega_A$ and $\omega_B$, respectively ($\Delta\omega_{AB}$ = 100 Hz)[150]. The indicated times on the right side correspond to $k_{ex}$= $k_{AB}$ + $k_{BA}$ ($s^{-1}$). **(a)** In the slow exchange regime ($k_{ex} \ll \Delta\omega$), signals from both states are observed reflecting their distinct chemical shifts, intensities, and linewidths. EXSY (or ZZex) method can be used to determine the kinetics of interconversion. In this slow regime, transient conformational excursions could be followed by H/D exchange and Real time NMR. Importantly, the stochastic variations in chemical shift produced by exchange can cause the NMR signal to decay more rapidly, leading to broader, weaker signals (coalescence condition). **(b)** Intermediate exchange regime ($k_{ex} \approx \Delta\omega$). Exchange broadening resulting from $k_{ex} < \approx$ 2 kHz can be analysed by CPMG and $k_{ex} \approx$ 20 kHz by $R_{1\varrho}$ relaxation dispersion methods. **(c)** In fast exchange regime ($k_{ex} \gg \Delta\omega$), a single sharp peak is obtained at the population-weighted average chemical shift ($\omega^{obs} = p_A\omega_A + p_B\omega_B$). Under these conditions, the decay of the NMR signal is dictated primarily by the rotational diffusion of the molecule in solution and nano- to picosecond timescale internal motions. NMR spin relaxation provides a measure of the amplitudes and frequencies of these fast time dynamics. In addition, structure-based measurements (insensitive to the timescale), such as RDC, PRE, NOEs and scalar couplings provide data that can be interpreted in terms of dynamics. Adapted from Mittermaier and Kay[140].

Methods for ps-ns timescale dynamics

Motions on ps-ns time scales are accessible to spin relaxation resulting from modulation of dipole-dipole, chemical shift anisotropy (CSA), and quadrupolar Hamiltonians[132]. While the latter arises from reorientation of bonds from molecular rotation on ns timescale with respect to the static $B_0$ field, and from site-specific internal motions on the ps timescale, motions on $\mu$s-ms time scales are accessible to spin relaxation resulting from modulation of isotropic chemical shifts (i.e., chemical exchange phenomena), *vide infra*. Nuclear spin relaxation can be used to constrain a model of protein dynamics at this timescale by measuring site-specific observables $R_1$ and $R_2$ rate constants, which describe the build-up of longitudinal (spin-lattice) and the decay of transverse (spin-spin) magnetization of a set of excited nuclear spins, respectively, as the system returns to equilibrium, and the heteronuclear nuclear Overhauser effect (hetNOE). Although the majority of applications of spin relaxation methods in proteins utilize the amide $^{15}$N nuclei as a spin probe of backbone motions (initially proposed by Kay, Torchia and Bax[159]), it is shown that quantification of $^{13}$C and $^{2}$H relaxation rates at

backbone as well as at side chain positions (e.g., methyl groups) are feasible with reasonable accuracy, in which the appropriate isotope labelling scheme simplify the relaxation process[160,161].

In this thesis, the $^{15}N$ spin relaxation of GB1 backbone is explored towards the correlation between fast dynamics and stability (see details in chapter 2). Interpretation of $^{15}N$ $R_1$, $R_2$ and HetNOE values are commonly shown along the sequence of the protein to identify regions of anomalous dynamic behaviour. If the protein is internally rigid and tumbles isotropically, then these values will be the same at each site. Therefore, values far from the average reflect site-specific internal motions and/or anisotropic molecular rotation. In general, $R_1$ increases with faster tumbling and is decreased by ps–ns flexibility; while $R_2$ increases with slower tumbling, μs-ms contributions from chemical exchange, $R_{ex}$, inhomogeneity in the static magnetic field, and it is decreased with ps–ns flexibility. $R_2$ is also affected by anisotropic rotation, therefore elevated $R_2$ values require careful interpretation since anisotropic tumbling can cause the same apparent effect as chemical exchange. The site-specific ratio $R_2/R_1$ is constant for spherical molecules and therefore global structure-dependent differences can indicate anisotropic molecular rotation. The $\{^1H\}$–$^{15}N$ hetNOE ranges from −4 to 1 and is reduced in the presence of internal flexibility. Since these $^{15}N$ relaxation rates are governed almost entirely by motions of the NH bond vector, which modulate both the $^{15}N$ chemical shift and the strength of the $^1H/^{15}N$ dipole-dipole interaction, their analysis usually separates amplitude of NH bond-vector motions into independent contributions from picosecond-timescale internal dynamics and nanosecond-timescale rotational diffusion of the molecule as a whole.

Briefly, a series of 2D HSQC-type spectra is collected with a variable relaxation delay, and, to obtain relaxation rates, single exponential equations are fit to the decays in resonance intensities, in addition to the $\{^1H\}$-$^{15}N$ hetNOE that is obtained by comparing signal intensities in the presence and absence of dipolar coupling (in practice, spectra are recorded with and without saturation of $^1H^N$ spins during the recycle delay). To better restrain fitting parameters, experiments are ideally carried out at two static magnetic field strengths. The resulting relaxation rates are subsequently analysed using the model-free approach introduced by Lipari & Szabo equations and extended by Clore et al[162]. Several key parameters can be quantified via their relationship with spectral density functions [$J(0)$, $J(\omega_H)$, $J(\omega_N)$][132], including: (1) the rigidity of the bond vector under investigation, which is expressed by the order parameter $S^2$, that can be further partitioned into fast and slow components $S^2_{fast}$ and $S^2_{slow}$; (2) the time scales of intra-molecular motions (range 0-100 ps), expressed by the correlation time $\tau_e$ and partitioned into $\tau_{e,fast}$ and $\tau_{e,slow}$ ; (3) the overall level of molecular tumbling (or rotational diffusion tensor $D$), expressed by the correlation time $\tau_c$; (4) the contributions to chemical exchange, $R_{ex}$. The goal is to model the dynamics of each bond vector as simply as possible by prudently selecting the site-specific parameters. Fast time scale dynamics of protein backbones have been also reviewed[160].

The order parameter $S^2$ gathers the most attention in fast relaxation dynamics since it is often mapped to the protein structure to identify local fluctuations. Its site-specific values can range from 0 for a complete flexible to 1 for a rigid bond-vector. Empirically, secondary structural elements tend to have high ($S^2 > 0.8$) amide N–H order parameters, whereas loops,

turns, and termini typically exhibit reduced values ($0.5 < S^2 < 0.8$). Changes in $S^2$ upon ligand binding or conformational transition (or similar) reflect altered flexibility and can be quantitatively related to a change in conformational entropy $\Delta S_p$[155]. Moreover, the temperature-dependence of $S^2$ can be used to obtain information about contributions to heat capacity ($C_p$) from such conformational fluctuations[163]. However, these thermodynamic interpretations of $S^2$ have numerous caveats[160], because we are forced to assume some simplicity (since each bond vector moves independently of each other bond vector). Despite their limitations, NMR-derived order parameters are indeed a proxy for conformational entropy with a significant role in allosteric regulation[164].

Methods for μs-ms timescale dynamics

Protein conformational dynamics on μs-to-ms timescales (intermediate-fast regime, $k_{ex} \approx \Delta\omega$) manifest as chemical exchange line broadening that is quantified by an excess contribution to the transverse relaxation rate [in the rotating frame ($R_{1\rho}$) and in the laboratory frame ($R_2$)], by $R_2^{obs} = R_2^0 + R_{ex}$, in which $R_2^{obs}$ is the observed free-precession transverse relaxation rate, $R_2^0$ is the relaxation rate constant in the absence of chemical exchange, and $R_{ex}$ is the chemical exchange relaxation contribution. For the slower millisecond exchange processes (0.25 ms $< \tau_{ex} = 1/k_{ex} <$ 5 ms) Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion (RD)[150,165,166] is the method of choice. Here, the effective transverse relaxation rate, $R_{2,eff}$, of a spin probe is monitored as a function of the number of refocusing pulses applied in a fixed time. That is, in the case of an exchanging system, the interconversion between states leads to a stochastic modulation of chemical shifts that is a source of signal dephasing. The extent of the dephasing, in turn, can be systematically modulated by the application of chemical shift refocusing pulses at specific repetition frequencies ($\nu_{CPMG} \approx$ 25 - 500 Hz) during a constant-time exchange period ($T_{relax}$), as shown in Fig. 1.15a. CPMG data are often presented as plots of signal dephasing rate/exchanging broadening ($R_{2,eff}$) versus $\nu_{CPMG}$ (Fig. 1.15b), where its shape depends on the parameters $p_B$, $k_{ex}$, and $\Delta\omega_{AB}$. These can be extracted from fits of CPMG profiles acquired at multiple static magnetic field strengths by modelling them globally using the so-called Bloch-McConnell equations. In the absence of conformational exchange, there is no observable modulation of $R_{2,eff}$.

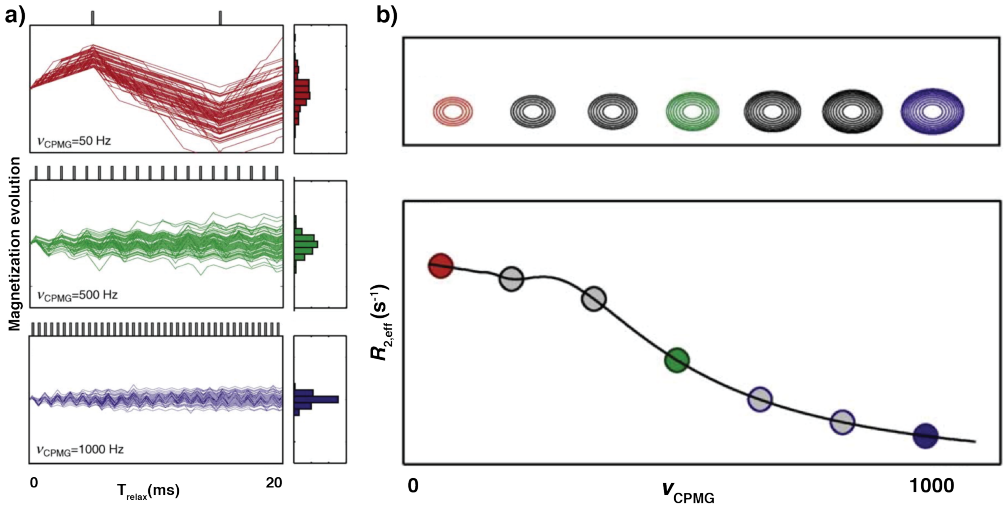



**Figure 1.15. CPMG relaxation dispersion experiments.**

(a) Simulated signal trajectories for an ensemble of nuclei exchanging stochastically between two environments with different chemical shifts during a 20 ms relaxation delay. Vertical bars at the top of each panel represent refocusing pulses applied during the relaxation delay. (b) (top) Simulated peak that increases in intensity at higher pulse repetition rates and (bottom) calculated transverse relaxation rates ($R_{2,eff}$) plotted as a function of $\nu_{CPMG} = 1/(2\tau)$, where $\tau$ is the delay between successive refocusing pulses. Adapted from Mittermaier and Kay[140].

For faster microsecond timescale chemical exchange processes, it becomes impossible to apply refocusing pulses sufficiently fast so as to modulate $R_{2,eff}$. Beyond this limit, rotating frame relaxation ($R_{1\rho}$) RD experiments are employed[167], where magnetization is "locked" along an effective field (typically 1-6 kHz), the direction of which depends on the spin-lock field strength ($B_1$) as well as on the offset ($\Omega$) of the applied field relative to the resonance position of the major state peak. The effective rate of relaxation is governed by both the intrinsic relaxation properties of the target nucleus as well as by conformational exchange processes. $R_{2,eff}$ is systematically modulated by altering $B_1$ and $\Omega$, the resulting off-resonance $R_{1\rho}$ relaxation dispersion profile maps $R_{2,eff}$ as a function of $\Omega$ . The relaxation profile can be fit to appropriate models of conformational exchange using the Bloch-McConnell equations. In the absence of exchange, flat profiles are obtained (no $\Omega$ dependence) that do not vary with $B_1$.

Methods for ms-s timescale dynamics

One approach to study slowly exchanging systems within a window of 2 ms < $\tau_{ex}$ < 25 ms is CEST (chemical exchange saturation transfer)[168,169]. This experiment amplifies the NMR signals of "invisible" excited conformations with populations as low as 0.5% (which typically cannot be directly observed in NMR spectra) and rates of exchange lie between ~200 and 2000 $s^{-1}$. In the CEST experiment a weak $B_1$ field, typically between 20–50 Hz, is applied, one frequency at a time, so as to cover the frequency range of the resonances of interest (e.g., between 105–135 ppm for amide $^{15}N$ spins) and the intensities of peaks from the major conformational (ground) state are quantified relative to the intensity in a reference spectrum. Intensity decreases when the weak $B_1$ field is applied at the resonance frequency of a spin in

the invisible state as well when the field is positioned at the resonance frequency of the ground state. A plot of $I/I_0$ generates a CEST profile (Fig. 1.16) with a pair of dips at the resonance positions of the ground ($\omega_G$) and excited states ($\omega_E$).

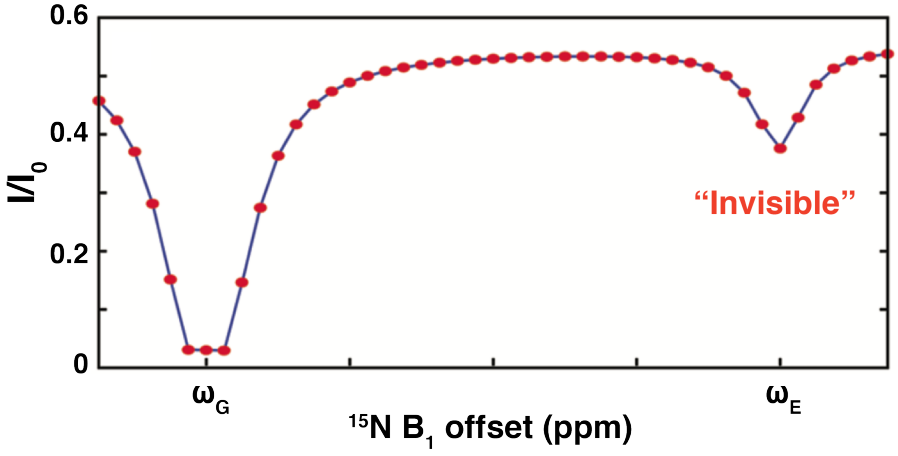



**Figure 1.16. $^{15}$N CEST typical profile.**

Intensity profile obtained by quantifying the intensity of the visible-state peak as a function of position of the weak $B_1$ irradiation field. The ratio $I/I_0$ is plotted, where I is the intensity after an irradiation period of duration $T_{EX}$ (exchange time) and $I_0$ is the intensity when $T_{EX} = 0$. There is a loss of intensity when the weak continuous-wave field is resonant with the major and minor states. Adapted from Vallurupalli et al[168].

The resulting profile can be fit to the Bloch-McConnell equations to extract $p_B$, $k_{ex}$, and chemical shifts of rare conformers[168]. Much like CEST, both CPMG and $R_{1\rho}$ RD briefly mentioned can be used to define sparsely populated conformations that are not visible in standard NMR spectra[170]. These data can be complemented with structural information: chemical shifts; orientation of bond vectors via residual dipolar coupling (RDCs)[171] ; and relaxation enhancements (PRE) that occur upon addition of paramagnets to the system of interest[172].

One of the most popular methods for probing biomolecular dynamics on slower timescales ($10\ ms < \tau_{ex} < 1\ s$) is referred to as magnetization exchange, or ZZex, in which, the best-known pulse sequence is the 2D $^1$H-$^{15}$N heteronuclear correlation experiment developed by Farrow and his colleagues in 1994[113]. This experiment, along with CEST, exploits the slow relaxation of longitudinal magnetization that enables quantitation of these slower processes. Like $^1$H EXSY (magnetization exchange spectroscopy), first proposed by Jeener et al. in 1979[173], the so-called ZZex is useful in cases where both states A and B are comparably populated (at the limit, the minor peak should be visible which means at least $p_B \geq 10\ \%$) so that peaks from both states are available for quantification in NMR spectra (Fig. 1.17). These experiments are the analogues to the $^1$H NOESY experiments but have the advantage of lacking numerous confounding NOE cross peaks. Furthermore, $k_{ex}$ must not be much less than $R_{1A}$ and $R_{1B}$, otherwise, the signals decay due to relaxation faster than population transfer occurs. Because $R_1$ of $^{15}$N and $^{13}$C spins in macromolecules is much slower than that of $^1$H spins, the X-nucleus (X = $^{15}$N or $^{13}$C) relaxation rate constant is more favourable than for two-spin order. For proteins, X-nucleus $R_1$ values typically are on the order of 1 $s^{-1}$, consequently ZZ-exchange measurements are applicable to chemical exchange processes with $k_{ex}$ in the range 0.1 to 10

$s^{-1}$. Numerous pulse sequences have been recently reviewed[152]. Briefly, in these schemes, the chemical exchange occurs between $2S_zI_z$ ($I$ = $^1$H, $S$ = $^{13}$C or $^{15}$N) longitudinal two-spin order, or longitudinal magnetization ($S_z$), or single transition operators (e.g., $S_zI^\beta$), in which detection of the signal utilizes HSQC or TROSY-detection/selection.

In a typical two-dimensional $^1$H-$^{15}$N HSQC spectrum (Fig. 1.17a), one peak would be observed for each backbone $^1$H/$^{15}$N, spin pair. In the case of an exchanging system, a pair of peaks (referred to as diagonal peaks) is obtained for each amide at frequencies of ($\omega_A{}^H$, $\omega_A{}^N$) and ($\omega_B{}^H$, $\omega_B{}^N$), with $^{15}$N resonance frequencies ($\omega_{A,B}{}^N$) recorded prior to detection of $^1$H resonance frequencies ($\omega_{A,B}{}^H$). In a conventional 2D $^1$H-$^{15}$N ZZex experiment[113], a delay is included between these frequency-recording periods during which longitudinal nitrogen magnetization ($N_z$) exchange occurs (i.e., $N_z$ magnetization is utilized for monitoring exchange rates and then is subsequently transferred back to $^1$H for detection by an INEPT scheme). This gives rise to an additional set of peaks (see Fig. 1.17b for cross-peaks) that result from conformational exchange between A and B during this interval so that a fraction of state A magnetization during the first chemical shift encoding period is transferred to state B for the second period and vice versa.

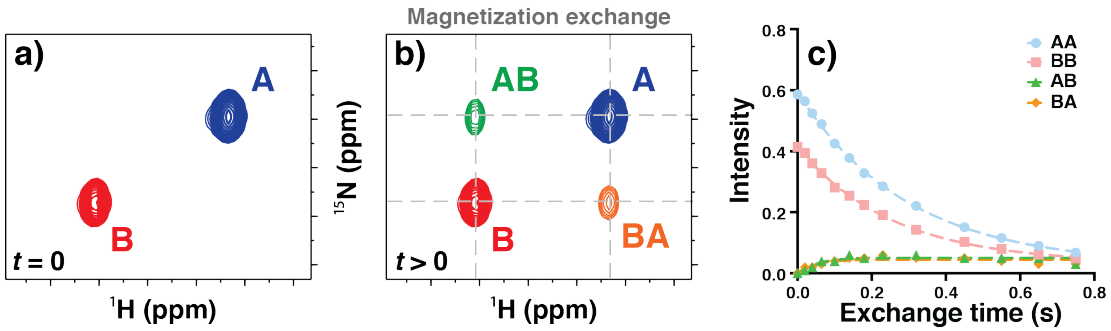



**Figure 1.17. ZZ-exchange experiment.**

ZZex spectra for a $^1$H/$^{15}$N spin pair undergoing slow two-site exchange for **a)** absence of chemical exchange and **b)** mixing period $t$ applied. Blue and red peaks correspond to the auto-peaks of states A and B, respectively, and green and orange cross-peaks are produced by conformational exchange during the delay. **a)** When $t$ = 0, two peaks corresponding to states A and B are obtained. **b)** When $t$ > 0, some of the nuclei initially in state A convert to state B during the delay, giving rise to a cross-peak at the $^{15}$N chemical shift of A and the $^1$H chemical shift of B, and vice versa. **c)** Plot of the time dependence of the intensities of the four peaks. In the simplified case shown here, profiles were simulated using parameter values of $p_B \approx 0.4$, $k_{ex} \approx 1.7$ $s^{-1}$, and $R_{1A} = R_{1B} \approx 2.6$ $s^{-1}$, in which cross-peaks AB and BA exhibit the same time dependence. The intensities of the cross-peaks initially increase as a function of $t$ (as the molecules are given more opportunity to exchange), however for longer delays, the intensities of all peaks decrease, due to longitudinal relaxation.

To extract populations, longitudinal $^{15}$N $T_1$ relaxation and interconversion rate constants, the time evolution of the resulting four peaks (Fig. 1.17c) is fit to the appropriate exchange model[114,150] using the solution of Bloch-McConnell equations. To obtain some insight into the form of the equations, assuming that $R_{1A} = R_{1B} = R_1$ (though this need not be the case), the time dependence of the longitudinal magnetization is given by[150]

$$a_{AA}(t) = [p_A + p_B \exp(-k_{ex}t)] \exp(-R_1 t),$$

$$a_{BB}(t) = [p_B + p_A \exp(-k_{ex}t)] \exp(-R_1 t),$$
$$a_{AB}(t) = p_A[1 - \exp(-k_{ex}t)] \exp(-R_1 t),$$
$$a_{BA}(t) = p_B[1 - \exp(-k_{ex}t)] \exp(-R_1 t). \quad [1.12]$$

Here, ZZex method was applied to the slow folding-unfolding interconversion of drkN SH3 as demonstrated previously[113,114] (chapter 3). A recent example of the use of conventional $^{1}$H-$^{15}$N ZZex experiment in folding kinetics investigation was obtained from two autonomously folding protein domains derived from ribosomal protein L9 using high-pressure NMR spectroscopy[174]. From this report, direct global fitting of the time dependence of the two auto- and cross-peak intensities is obviously prone to systematic errors from the effects of differential relaxation during parts of the experiment other than the relation delay, *t*, particularly the $t_2$ acquisition period. Therefore, a simplified analysis was proposed by Miloushev et al[175],

$$\Xi(t) = \frac{I_{AB}(t) I_{BA}(t)}{I(t) I_{BB}(t) - I_{AB}(t) I_{BA}(t)} \approx k_{AB} k_{BA} t^2, \quad [1.13]$$

where $I_{AB,BA}(t)$ and $I_{AA,BB}(t)$ are the intensity of the cross-peak and auto-peak, respectively, in the spectrum recorded with a mixing time *t*. In particular, for data at shorter mixing times, the simplicity of this global analysis is evident[152]. Other robust approach[176] consists in the recording of a second experiment without resolving exchange cross-peaks at each mixing time (i.e., without frequency-labelling in an indirect dimension). By combining the two datasets, systematic errors that arise from differential line broadening effects are eliminated.

Methods for timescale dynamics above seconds to hours/days

Dynamic processes on the second timescale are directly detected by quantifying the time-dependence of NMR signal intensities[177]. Physical processes on this timescale include protein folding, solvent hydrogen-exchange and relatively slow conformational changes. The real-time (RT) NMR experiment is performed by initiating the physical process of interest commonly using a special apparatus (e.g., pH, temperature or pressure jump, addition of a binding partner, denaturants or stabilising cosolutes) and then rapidly acquiring a sequence of NMR spectra. Fast RT NMR methods have been reviewed in detail elsewhere[178].

RT NMR can be used to study the dynamic behaviour of proteins via hydrogen–deuterium exchange ($H^N/D^N$ or HX) by monitoring the exchange of hydrogen atoms between the protein and the solvent[179,180]. In a typical experiment, a protein bearing exchangeable $^{1}$H is dissolved in a fully deuterated solvent (such as $D_2O$, D= $^{2}$H), $^{1}$H is replaced by $^{2}$H and hence the observable $^{1}$H signals decay in intensity over time (for example individual backbone amide protons $H^N$ are followed by a series of $^{1}$H-$^{15}$N HSQC spectra). The observed rate of amide proton (N–H) exchange, $k_{ex}^{obs}$, is related to equilibrium stability by considering a protein in which each N–H exists in a closed (protected, exchange-incompetent), or an open (exposed, exchange-competent) state, accordingly with Linderstrøm-Lang and his colleagues[181]

$$closed\ (N-H) \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}}\ open\ (N-H) \xrightarrow{k_{int}}\ exchange\ (N-D), \qquad [1.14]$$

each position opens and closes with rate constants, $k_{op}$ and $k_{cl}$ (equilibrium constant for structural opening by $K_{op} = k_{op}/k_{cl}$), and exchange from the open state occurs with intrinsic rate constant, $k_{int}$ that can be approximated as the intrinsic rate constant for a random coil (depends on local amino acid sequence, temperature, and pH)[182]. Measured hydrogen exchange rates are often recast in terms of site-specific protection factors ($P_f$), where $P_f = k_{int} / k_{ex}^{obs}$. The simple two-state kinetic model (Eq. 1.14) specifies that when structure is stable (that is, $k_{cl} \gg k_{op}$) the exchange occurs within two distinct regimes denoted EX1 (fast and unimolecular) and EX2 (slow and bimolecular), accordingly to[181],

$$k_{ex}^{obs} = \frac{k_{op} k_{int}}{k_{cl} + k_{int}}. \qquad [1.15]$$

At the so-called EX1 limit the closing is rate determining ($k_{cl} \ll k_{int}$) and the observed exchange rate reflects the rate of opening ($k_{ex}^{obs} \approx k_{op}$), this regime is typically sampled at high pH/temperature or in the presence of denaturants since all opening events lead to an exchange. At the usually observed EX2 kinetics, the exchange from the open state is rate limiting, $k_{cl} \gg k_{int}$ and $k_{ex}^{obs}$ reflects the equilibrium between the *open* and *closed* states: $k_{ex}^{obs} \approx (k_{op}/k_{cl})\ k_{int} = K_{op}\ k_{int}$. Based on this equilibrium, local stabilities are measured as the free energy of exchange or opening ($\Delta G_{HX}$) which depends on both measured $k_{ex}^{obs}$ and predicted $k_{int}$ values,

$$\Delta G_{HX} = -RTln(K_{op}) = -RTln\left(\frac{k_{ex}^{obs}}{k_{int}}\right) = -RTln\left(\frac{1}{P_f}\right), \qquad [1.16]$$

where RT is the molar gas constant multiplied by the absolute temperature. The nature of the events leading to the exchange can also be studied, distinguishing between full unfolding and local fluctuation[183]. Although amides in the EX1 limit often exchange too quickly to be detected (seconds to minutes) via conventional RT-NMR, $^{1}H$-$^{15}N$ 2D SOFAST-HMQC experiment has been successfully applied to study EX1 of human ubiquitin at high pH[184]. It is worth noting that in the case where a "state" corresponds to an average over multiple conformers that interconvert rapidly on the NMR chemical shift timescale, the measured $k_{ex}^{obs}$ values would correspond to a weighted average over exchange rates in the ensemble.

Faster exchange processes can also be studied with a variety of solvent EXSY methods. Examples of such experiments are phase-modulated clean chemical exchange (CLEANEX-PM)[185] and solvent exchange spectroscopy (SOLEXY)[186]. These techniques have been developed to reduce artifacts introduced by alternative magnetization transfer mechanisms and made it possible to study hydrogen exchange processes on the millisecond time scale. In the limit of slow exchange it is even possible to quantify hydrogen exchange rate constants for minor/invisible excited states using $^{15}N$-based CEST schemes[187].

HX exchange using HSQC pulse scheme as well as (CLEANEX-PM)-FHSQC technique are explored with GB1 protein, which has a EX2 mechanism of exchange very well established[109,188] (discussed in chapter 2).

## Outline of the thesis

This introductory chapter laid a background for the work presented in the following chapters, which are divided by the protein under study (GB1 and drkN SH3). Along the thesis, [Ch][Glu] or [Bmim][dca] ILs are used as a stabiliser or as a denaturant cosolute, respectively. From the training perspective, the key objective of this thesis has been to acquire knowledge on the application of biomolecular NMR methods and other biophysical approaches to study protein stability. The data are analysed to yield thermodynamic and kinetic information. The two main goals of this thesis are: (1) Identification of relevant charged metabolites protect protein stability; and (2) Molecular understanding of ILs effects on protein stability. To accomplish the last goal, two tasks were defined: (2.1) Interactions studies of model protein system and selected ILs, and (2.2) Identification of key protein residue clusters and ILs effects involved in protein stability.

In chapter 2, I study the effects of ILs on GB1 stability and structure, exploring chemical shift perturbations, protein fast dynamics and HX exchange features to clarify IL effects. The thermal stability of GB1 is investigated by monitorization of its intrinsic tryptophan fluorescence. Since the unfolded ensemble is expected to be affected by the IL but it is not properly addressed with GB1, in chapter 3, I explore the modulation of drkN SH3 folding pathway with ILs. Because SH3 is marginally stable, a thermodynamic analysis is implemented, and each population (folded or unfolded conformers) is evaluated in terms of structural and kinetic changes. The characterisation of the slow folding/unfolding interconversion rates of SH3 is performed with ZZ-exchange method. Finally, chapter 4 summarizes and correlates the findings of each chapter. The data gathered provides a thorough understanding of the IL-protein interactions and their effect on the equilibrium thermodynamics and kinetics.

# References


1. Jungwirth, P. Biological Water or Rather Water in Biology? *J. Phys. Chem. Lett.* **6**, 2449–2451 (2015).
2. Hofmeister, F. Zur Lehre von der Wirkung der Salze. *Arch. für Exp. Pathol. und Pharmakologie* **24**, 247–260 (1888).
3. Weingärtner, H. Understanding Ionic Liquids at the Molecular Level: Facts, Problems, and Controversies. *Angew. Chemie Int. Ed.* **47**, 654–670 (2008).
4. Hayes, R., Warr, G. G. & Atkin, R. Structure and Nanostructure in Ionic Liquids. *Chem. Rev.* **115**, 6357–6426 (2015).
5. Welton, T. Room-Temperature Ionic Liquids. Solvents for Synthesis and Catalysis. *Chem. Rev.* **99**, 2071–2084 (1999).
6. Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331 (1999).
7. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science (80-. ).* **181**, 223–230 (1973).
8. Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. *Nat. Struct. Mol. Biol.* **4**, 10–19 (1997).
9. Oliveberg, M. & Wolynes, P. G. The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* **38**, 245–288 (2005).
10. Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **29**, 7133–7155 (1990).
11. Lumry, R. & Rajender, S. Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous properly of water. *Biopolymers* **9**, 1125–1227 (1970).
12. Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation. in *Advances in Protein Chemistry* **14**, 1–63 (1959).
13. Senske, M. *et al.* The temperature dependence of the Hofmeister series: thermodynamic fingerprints of cosolute–protein interactions. *Phys. Chem. Chem. Phys.* **18**, 29698–29708 (2016).
14. Tanford, C. Protein Denaturation. in *Advances in protein chemistry* **23**, 121–282 (1968).
15. Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. (W. H. Freeman and Company, New York, 1999).
16. Makhatadze, G. I. & Privalov, P. L. Contribution of Hydration to Protein Folding Thermodynamics. *J. Mol. Biol.* **232**, 639–659 (1993).
17. Makhatadze, G. I. & Privalov, P. L. Energetics of Protein Structure. in *Advances in Protein Chemistry* **47**, 307–425 (1995).
18. Privalov, P. L. & Makhatadze, G. I. Heat capacity of proteins. *J. Mol. Biol.* **213**, 385–391 (1990).
19. Becktel, W. J. & Schellman, J. A. Protein stability curves. *Biopolymers* **26**, 1859–1877 (1987).
20. Eaton, W. A. *et al.* Fast Kinetics and Mechanisms in Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359 (2000).
21. Fersht, A. R. Characterizing transition states in protein folding: an essential step in the puzzle. *Curr. Opin. Struct. Biol.* **5**, 79–84 (1995).
22. Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122–126 (1989).
23. Dill, K. A. & MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science (80-. ).* **338**, 1042–1046 (2012).
24. Englander, S. W. & Mayne, L. The nature of protein folding pathways. *Proc. Natl. Acad. Sci.* **111**, 15873–15880 (2014).
25. Levinthal, C. How to fold graciously. *Mössbauer Spectrosc. Biol. Syst. Proc.* **24**, 22–24 (1969).
26. Bryngelson, J. D. & Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci.* **84**, 7524–7528 (1987).
27. Leopold, P. E., Montal, M. & Onuchic, J. N. Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8721–8725 (1992).
28. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins Struct. Funct. Bioinforma.* **21**, 167–195 (1995).
29. Onuchic, J. N., Luthey-Schulten, Z. & Wolynes, P. G. THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
30. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. & Socci, N. D. Toward an outline of the topography of a realistic protein-folding funnel. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3626–3630 (1995).
31. Baldwin, A. J. & Kay, L. E. NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5**, 808–814 (2009).
32. Sekhar, A. & Kay, L. E. NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc. Natl. Acad. Sci.* **110**, 12867–12874 (2013).
33. Ferreiro, D. U., Hegler, J. A., Komives, E. A. & Wolynes, P. G. Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19819–19824 (2007).
34. Ferreiro, D. U., Komives, E. A. & Wolynes, P. G. Frustration in biomolecules. *Q. Rev. Biophys.* **47**, 285–363 (2014).
35. Chen, M. *et al.* Surveying biomolecular frustration at atomic resolution. *Nat. Commun.* **11**, 1–9 (2020).

36. Sutto, L., Lätzer, J., Hegler, J. A., Ferreiro, D. U. & Wolynes, P. G. Consequences of localized frustration for the folding mechanism of the IM7 protein. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19825–19830 (2007).

37. Englander, S. W. & Mayne, L. The case for defined protein folding pathways. *Proc. Natl. Acad. Sci.* **114**, 8253–8258 (2017).

38. Eaton, W. A. & Wolynes, P. G. Theory, simulations, and experiments show that proteins fold by multiple pathways. *Proc. Natl. Acad. Sci.* **114**, E9759–E9760 (2017).

39. Guinn, E. J., Jagannathan, B. & Marqusee, S. Single-molecule chemo-mechanical unfolding reveals multiple transition state barriers in a small single-domain protein. *Nat. Commun.* **6**, 1–8 (2015).

40. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–42 (2000).

41. Senior, A. W. *et al.* Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).

42. Fink, A. L. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* **15**, 35–41 (2005).

43. Van Der Lee, R. *et al.* Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 (2014).

44. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).

45. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).

46. Uversky, V. N. Intrinsically Disordered Proteins and Their "Mysterious" (Meta)Physics. *Front. Phys.* **7**, 8–23 (2019).

47. Kunz, W., Henle, J. & Ninham, B. W. 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. *Curr. Opin. Colloid Interface Sci.* **9**, 19–37 (2004).

48. Zhang, Y. & Cremer, P. S. Interactions between macromolecules and ions: the Hofmeister series. *Curr. Opin. Chem. Biol.* **10**, 658–663 (2006).

49. Jungwirth, P. & Cremer, P. S. Beyond Hofmeister. *Nat. Chem.* **6**, 261–263 (2014).

50. Marcus, Y. Effect of ions on the structure of water: Structure making and breaking. *Chem. Rev.* **109**, 1346–1370 (2009).

51. Omta, A. W. Negligible Effect of Ions on the Hydrogen-Bond Structure in Liquid Water. *Science (80-. ).* **301**, 347–349 (2003).

52. Zhang, Y. & Cremer, P. S. The inverse and direct Hofmeister series for lysozyme. *Proc. Natl. Acad. Sci.* **106**, 15249–15253 (2009).

53. Arakawa, T. & Timasheff, S. N. Preferential Interactions of Proteins with Salts in Concentrated Solutions. *Biochemistry* **21**, 6545–6552 (1982).

54. Baldwin, R. L. How Hofmeister ion interactions affect protein stability. *Biophys. J.* **71**, 2056–2063 (1996).

55. Lo Nostro, P. & Ninham, B. W. Hofmeister Phenomena: An Update on Ion Specificity in Biology. *Chem. Rev.* **112**, 2286–2322 (2012).

56. Okur, H. I. *et al.* Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **121**, 1997–2014 (2017).

57. Mason, P. E. *et al.* Specificity of Ion−Protein Interactions: Complementary and Competitive Effects of Tetrapropylammonium, Guanidinium, Sulfate, and Chloride Ions. *J. Phys. Chem. B* **113**, 3227–3234 (2009).

58. Sukenik, S., Sapir, L., Gilman-Politi, R. & Harries, D. Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday Discuss.* **160**, 225–237 (2013).

59. Plechkova, N. V & Seddon, K. R. Applications of ionic liquids in the chemical industry. *Chem. Soc. Rev.* **37**, 123–150 (2008).

60. Hallett, J. P. & Welton, T. Room-Temperature Ionic Liquids: Solvents for Synthesis and Catalysis. 2. *Chem. Rev.* **111**, 3508–3576 (2011).

61. Walden, P. Ueber die Molekulargrösse und elektrische Leitfähigkeit einiger geschmolzenen Salze. *Bull. l'Académie impériale des Sci. St.-pétersbg.* **8**, 405–422 (1914).

62. Wilkes, J. S. & Zaworotko, M. J. Air and water stable 1-ethyl-3-methylimidazolium based ionic liquids. *J. Chem. Soc. Chem. Commun.* 965 (1992).

63. Rogers, R. D. & Seddon, K. R. Ionic Liquids - Solvents of the Future? *Science (80-. ).* **302**, 792–793 (2003).

64. López-Martin, I., Burello, E., Davey, P. N., Seddon, K. R. & Rothenberg, G. Anion and Cation Effects on Imidazolium Salt Melting Points: A Descriptor Modelling Study. *ChemPhysChem* **8**, 690–695 (2007).

65. Fumino, K., Wulf, A. & Ludwig, R. Hydrogen Bonding in Protic Ionic Liquids: Reminiscent of Water. *Angew. Chemie Int. Ed.* **48**, 3184–3186 (2009).

66. Petkovic, M., Seddon, K. R., Rebelo, L. P. N. & Silva Pereira, C. Ionic liquids: a pathway to environmental acceptability. *Chem. Soc. Rev.* **40**, 1383–1403 (2011).

67. Gomes, J. M., Silva, S. S. & Reis, R. L. Biocompatible ionic liquids: fundamental behaviours and applications. *Chem. Soc. Rev.* **48**, 4317–4335 (2019).

68. Fukumoto, K., Yoshizawa, M. & Ohno, H. Room Temperature Ionic Liquids from 20 Natural Amino Acids. *J. Am. Chem. Soc.* **127**, 2398–2399 (2005).

69. Zeisel, S. H. & Blusztajn, J. K. Choline and Human Nutrition. *Annu. Rev. Nutr.* **14**, 269–296 (1994).

70. Liu, Q.-P., Hou, X.-D., Li, N. & Zong, M.-H. Ionic liquids from renewable biomaterials: synthesis, characterization and application in the pretreatment of biomass. *Green Chem.* **14**, 304–307 (2012).

71. Fujita, K., MacFarlane, D. R. & Forsyth, M. Protein solubilising and stabilising ionic liquids. *Chem. Commun.* **70**, 4804 (2005).

72. Byrne, N., Wang, L.-M., Belieres, J. & Angell, C. A. Reversible folding–unfolding, aggregation protection, and multi-year stabilization, in high concentration protein solutions, using ionic liquids. *Chem. Commun.* 2714–2716 (2007).

73. Attri, P., Venkatesu, P. & Kumar, A. Activity and stability of $\alpha$-chymotrypsin in biocompatible ionic liquids: enzyme refolding by triethyl ammonium acetate. *Phys. Chem. Chem. Phys.* **13**, 2788–2796 (2011).

74. Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).

75. Constatinescu, D., Herrmann, C. & Weingärtner, H. Patterns of protein unfolding and protein aggregation in ionic liquids. *Phys. Chem. Chem. Phys.* **12**, 1756 (2010).

76. Weingärtner, H., Cabrele, C. & Herrmann, C. How ionic liquids can help to stabilize native proteins. *Phys. Chem. Chem. Phys.* **14**, 415–426 (2012).

77. Vendruscolo, M. & Dobson, C. M. Towards complete descriptions of the free–energy landscapes of proteins. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **363**, 433–452 (2005).

78. Figueiredo, A. M., Sardinha, J., Moore, G. R. & Cabrita, E. J. Protein destabilisation in ionic liquids: the role of preferential interactions in denaturation. *Phys. Chem. Chem. Phys.* **15**, 19632 (2013).

79. Silva, M., Figueiredo, A. M. & Cabrita, E. J. Epitope mapping of imidazolium cations in ionic liquid–protein interactions unveils the balance between hydrophobicity and electrostatics towards protein destabilisation. *Phys. Chem. Chem. Phys.* **16**, 23394–23403 (2014).

80. Zhao, H. Protein stabilization and enzyme activation in ionic liquids: specific ion effects. *J. Chem. Technol. Biotechnol.* **91**, 25–50 (2016).

81. Smiatek, J. Aqueous ionic liquids and their effects on protein structures: an overview on recent theoretical and experimental results. *J. Phys. Condens. Matter* **29**, 233001 (2017).

82. Schröder, C. Proteins in Ionic Liquids: Current Status of Experiments and Simulations. *Top. Curr. Chem.* **375**, 25 (2017).

83. Kumar, A., Bisht, M. & Venkatesu, P. Biocompatibility of ionic liquids towards protein stability: A comprehensive overview on the current understanding and their implications. *Int. J. Biol. Macromol.* **96**, 611–651 (2017).

84. Reslan, M. & Kayser, V. Ionic liquids as biocompatible stabilizers of proteins. *Biophys. Rev.* **10**, 781–793 (2018).

85. Schindl, A., Hagen, M. L., Muzammal, S., Gunasekera, H. A. D. & Croft, A. K. Proteins in Ionic Liquids: Reactions, Applications, and Futures. *Front. Chem.* **7**, 1–31 (2019).

86. Shukla, S. K. & Mikkola, J.-P. Use of Ionic Liquids in Protein and DNA Chemistry. *Front. Chem.* **8**, 1–23 (2020).

87. Constantinescu, D., Weingärtner, H. & Herrmann, C. Protein Denaturation by Ionic Liquids and the Hofmeister Series: A Case Study of Aqueous Solutions of Ribonuclease A. *Angew. Chemie Int. Ed.* **46**, 8887–8889 (2007).

88. Canchi, D. R. & García, A. E. Cosolvent effects on protein stability. *Annu. Rev. Phys. Chem.* **64**, 273–293 (2013).

89. Senske, M. *et al.* Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc.* **136**, 9036–9041 (2014).

90. Oprzeska-Zingrebe, E. A. & Smiatek, J. Aqueous ionic liquids in comparison with standard co-solutes. *Biophys. Rev.* **10**, 809–824 (2018).

91. Bui-Le, L. *et al.* Revealing the complexity of ionic liquid–protein interactions through a multi-technique investigation. *Commun. Chem.* **3**, 55 (2020).

92. Lesch, V., Heuer, A., Tatsis, V. a., Holm, C. & Smiatek, J. Peptides in the presence of aqueous ionic liquids: tunable co-solutes as denaturants or protectants? *Phys. Chem. Chem. Phys.* **17**, 26049–26053 (2015).

93. Shmool, T. A. *et al.* Exploring conformational preferences of proteins: ionic liquid effects on the energy landscape of avidin. *Chem. Sci.* **12**, 196–209 (2021).

94. Theillet, F.-X. *et al.* Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **114**, 6661–6714 (2014).

95. Harries, D. & Rösgen, J. A Practical Guide on How Osmolytes Modulate Macromolecular Properties. in *Methods in Cell Biology* **84**, 679–735 (Academic Press, 2008).

96. Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* **5**, 593–599 (2009).

97. Wishart, D. S. *et al.* HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **46**, D608–D617 (2018).

98. Hollóczki, O., Malberg, F., Welton, T. & Kirchner, B. On the origin of ionicity in ionic liquids. Ion pairing versus charge transfer. *Phys. Chem. Chem. Phys.* **16**, 16880–16890 (2014).

99. MacFarlane, D. R. *et al.* Ionic liquid "buffers"—pH control in ionic liquid systems. *Chem. Commun.* **46**, 7703 (2010).

100. Kumar, A., Bhakuni, K. & Venkatesu, P. Strategic planning of proteins in ionic liquids: future solvents for the enhanced stability of proteins against multiple stresses. *Phys. Chem. Chem. Phys.* **21**, 23269–23282 (2019).

101. Silva, M. S., Figueiredo, A. M. & Cabrita, E. J. Design of bio-inspired ionic liquids for protein stabilisation. (Universidade

Nova de Lisboa, 2015).

102. Franks, W. T. *et al.* Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. *Proc. Natl. Acad. Sci.* **105**, 4621–4626 (2008).

103. Bezsonova, I., Singer, A., Choy, W.-Y., Tollinger, M. & Forman-Kay, J. D. Structural Comparison of the Unstable drkN SH3 Domain and a Stable Mutant † , ‡. *Biochemistry* **44**, 15550–15560 (2005).

104. Marsh, J. A. & Forman-Kay, J. D. Structure and Disorder in an Unfolded State under Nondenaturing Conditions from Ensemble Models Consistent with a Large Number of Experimental Restraints. *J. Mol. Biol.* **391**, 359–374 (2009).

105. Gronenborn, A. *et al.* A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science (80-. ).* **253**, 657–661 (1991).

106. Alexander, P., Fahnestock, S., Lee, T., Orban, J. & Bryan, P. Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. *Biochemistry* **31**, 3597–3603 (1992).

107. McCallister, E. L., Alm, E. & Baker, D. Critical role of β-hairpin formation in protein G folding. *Nat. Struct. Biol.* **7**, 669–673 (2000).

108. Ding, K., Louis, J. M. & Gronenborn, A. M. Insights into Conformation and Dynamics of Protein GB1 During Folding and Unfolding by NMR. *J. Mol. Biol.* **335**, 1299–1307 (2004).

109. Monteith, W. B. & Pielak, G. J. Residue level quantification of protein stability in living cells. *Proc. Natl. Acad. Sci.* **111**, 11335–11340 (2014).

110. Pawson, T. Protein modules and signalling networks. *Nature* **373**, 573–580 (1995).

111. Olivier, J. P. *et al.* A Drosophila SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of Ras guanine nucleotide exchange, Sos. *Cell* **73**, 179–191 (1993).

112. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845–858 (1994).

113. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734 (1994).

114. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D. & Kay, L. E. Slow Dynamics in Folded and Unfolded States of an SH3 Domain. *J. Am. Chem. Soc.* **123**, 11341–11352 (2001).

115. Marsh, J. A. *et al.* Improved Structural Characterizations of the drkN SH3 Domain Unfolded State Suggest a Compact Ensemble with Native-like and Non-native Structure. *J. Mol. Biol.* **367**, 1494–1510 (2007).

116. Smith, A. E., Zhou, L. Z., Gorensek, A. H., Senske, M. & Pielak, G. J. In-cell thermodynamics and a new role for protein surfaces. *Proc. Natl. Acad. Sci.* **113**, 1725–1730 (2016).

117. Zhang, O. & Forman-Kay, J. D. NMR Studies of Unfolded States of an SH3 Domain in Aqueous Solution and Denaturing Conditions †. *Biochemistry* **36**, 3959–3970 (1997).

118. Lincoff, J. *et al.* Extended experimental inferential structure determination method in determining the structural ensembles of disordered protein states. *Commun. Chem.* **3**, 74 (2020).

119. Lee, J. H. *et al.* Heterogeneous binding of the SH3 client protein to the DnaK molecular chaperone. *Proc. Natl. Acad. Sci.* **112**, E4206–E4215 (2015).

120. Sekhar, A. *et al.* Conserved conformational selection mechanism of Hsp70 chaperone-substrate interactions. *Elife* **7**, 1–29 (2018).

121. Bodenhausen, G. & Ruben, D. J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185–189 (1980).

122. Kay, L. E., Keifer, P. & Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665 (1992).

123. Keeler, J. *Understanding NMR spectroscopy.* (John Wiley & Sons, 2011).

124. Levitt, M. H. *Spin Dynamics: Basics of Nuclear Magnetic Resonance.* (John Wiley & Sons, 2013).

125. Hore, P. J., Jones, J. A. & Wimperis, S. *NMR: The Toolkit: How Pulse Sequences Work*. (Oxford University Press, 2015).

126. Palmer, A. G., Cavanagh, J., Wright, P. E. & Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* **93**, 151–170 (1991).

127. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89**, 496–514 (1990).

128. Bax, A. & Ikura, M. An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. *J. Biomol. NMR* **1**, 99–104 (1991).

129. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).

130. Kay, L. E. The evolution of solution state NMR pulse sequences through the 'eyes' of triple-resonance spectroscopy. *J. Magn. Reson.* **306**, 48–54 (2019).

131. Sattler, M. & Fesik, S. W. Use of deuterium labeling in NMR: Overcoming a sizeable problem. *Structure* **4**, 1245–1249 (1996).

132. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Rance, M. & Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*. (Academic Press, 2006).

133. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. 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**, 12366–12371 (1997).

134. Mobli, M. & Hoch, J. C. Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **83**, 21–41 (2014).

135. Ying, J., Delaglio, F., Torchia, D. A. & Bax, A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **68**, 101–118 (2017).

136. Tollinger, M., Forman-Kay, J. D. & Kay, L. E. Measurement of Side-Chain Carboxyl p K a Values of Glutamate and Aspartate Residues in an Unfolded Protein by Multinuclear NMR Spectroscopy. *J. Am. Chem. Soc.* **124**, 5714–5717 (2002).

137. André, I., Linse, S. & Mulder, F. A. A. Residue-Specific p K a Determination of Lysine and Arginine Side Chains by Indirect 15 N and 13 C NMR Spectroscopy: Application to apo Calmodulin. *J. Am. Chem. Soc.* **129**, 15805–15813 (2007).

138. Yamazaki, T., Yoshida, M. & Nagayama, K. Complete assignments of magnetic resonances of ribonuclease H from Escherichia coli by double- and triple-resonance 2D and 3D NMR spectroscopies. *Biochemistry* **32**, 5656–5669 (1993).

139. Pellecchia, M., Iwai, H., Szyperski, T. & Wüthrich, K. The 2D NMR Experiments H(C)CO2 and HCCO2 for Assignment and pH Titration of Carboxylate Groups in Uniformly 15N/13C-Labeled Proteins. *J. Magn. Reson.* **124**, 274–278 (1997).

140. Paquin, R., Ferrage, F., Mulder, F. A. A., Akke, M. & Bodenhausen, G. Multiple-Timescale Dynamics of Side-Chain Carboxyl and Carbonyl Groups in Proteins by 13 C Nuclear Spin Relaxation. *J. Am. Chem. Soc.* **130**, 15805–15807 (2008).

141. Hansen, A. L. & Kay, L. E. Quantifying millisecond time-scale exchange in proteins by CPMG relaxation dispersion NMR spectroscopy of side-chain carbonyl groups. *J. Biomol. NMR* **50**, 347–355 (2011).

142. Pelton, J. G., Torchia, D. A., Meadow, N. D. & Roseman, S. Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated III Glc , a signal-transducing protein from escherichia coli , using two-dimensional heteronuclear NMR techniques. *Protein Sci.* **2**, 543–558 (1993).

143. Henry, G. D. & Sykes, B. D. Determination of the rotational dynamics and pH dependence of the hydrogen exchange rates of the arginine guanidino group using NMR spectroscopy. *J. Biomol. NMR* **6**, 59–66 (1995).

144. Iwahara, J., Jung, Y.-S. & Clore, G. M. Heteronuclear NMR Spectroscopy for Lysine NH 3 Groups in Proteins: Unique Effect of Water Exchange on 15 N Transverse Relaxation. *J. Am. Chem. Soc.* **129**, 2971–2980 (2007).

145. Zandarashvili, L., Esadze, A. & Iwahara, J. NMR Studies on the Dynamics of Hydrogen Bonds and Ion Pairs Involving Lysine Side Chains of Proteins. in *Biomolecular Spectroscopy: Advances from Integrating Experiments and Theory* (ed. Intergovernmental Panel on Climate Change) 37–80 (Cambridge University Press, 2013).

146. Werbeck, N. D., Kirkpatrick, J. & Hansen, D. F. Probing Arginine Side-Chains and Their Dynamics with Carbon-Detected NMR Spectroscopy: Application to the 42 kDa Human Histone Deacetylase 8 at High pH. *Angew. Chemie Int. Ed.* **52**, 3145–3147 (2013).

147. Gerecht, K., Figueiredo, A. M. & Hansen, D. F. Determining rotational dynamics of the guanidino group of arginine side chains in proteins by carbon-detected NMR. *Chem. Commun.* **53**, 10062–10065 (2017).

148. Mackenzie, H. W. & Hansen, D. F. A 13C-detected 15N double-quantum NMR experiment to probe arginine side-chain guanidinium 15Nη chemical shifts. *J. Biomol. NMR* **69**, 123–132 (2017).

149. Boehr, D. D., Dyson, H. J. & Wright, P. E. An NMR Perspective on Enzyme Dynamics. *Chem. Rev.* **106**, 3055–3079 (2006).

150. Palmer, A. G., Kroenke, C. D. & Patrick Loria, J. Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules. in *Methods in Enzymology* **339**, 204–238 (Elsevier Masson SAS, 2001).

151. Palmer, A. G. NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **104**, 3623–3640 (2004).

152. Palmer, A. G. & Koss, H. Chemical Exchange. in *Methods in Enzymology* **615**, 177–236 (Elsevier Inc., 2019).

153. Kleckner, I. R. & Foster, M. P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta - Proteins Proteomics* **1814**, 942–968 (2011).

154. Kovermann, M., Rogne, P. & Wolf-Watz, M. Protein dynamics and function from solution state NMR spectroscopy. *Q. Rev. Biophys.* **49**, e6 (2016).

155. Yang, D. & Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: Application to protein folding. *J. Mol. Biol.* **263**, 369–382 (1996).

156. Mittermaier, A. K. & Kay, L. E. Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.* **34**, 601–611 (2009).

157. Sekhar, A. & Kay, L. E. An NMR View of Protein Dynamics in Health and Disease. *Annu. Rev. Biophys.* **48**, 297–319 (2019).

158. Alderson, T. R. & Kay, L. E. NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. *Cell* **184**, 577–595 (2021).

159. Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).

160. Jarymowycz, V. A. & Stone, M. J. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. *Chem. Rev.* **106**, 1624–1671 (2006).

161. Igumenova, T. I., Frederick, K. K. & Wand, A. J. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chem. Rev.* **106**, 1672–1699 (2006).

162. Clore, G. M. *et al.* Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.* **112**, 4989–4991 (1990).

163. Yang, D., Mok, Y.-K., Forman-Kay, J. D., Farrow, N. A. & Kay, L. E. Contributions to protein entropy and heat capacity from bond vector motions measured by NMR spin relaxation. *J. Mol. Biol.* **272**, 790–804 (1997).

164. Wand, A. J. Dynamic activation of protein function: a view emerging from NMR spectroscopy. *Nat. Struct. Biol.* **8**, 926–31 (2001).

165. Loria, J. P., Rance, M. & Palmer, A. G. A Relaxation-Compensated Carr−Purcell−Meiboom−Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332 (1999).

166. Hansen, D. F., Vallurupalli, P. & Kay, L. E. An Improved 15 N Relaxation Dispersion Experiment for the Measurement of Millisecond Time-Scale Dynamics in Proteins †. *J. Phys. Chem. B* **112**, 5898–5904 (2008).

167. Palmer, A. G. & Massi, F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).

168. Vallurupalli, P., Bouvignies, G. & Kay, L. E. Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *J. Am. Chem. Soc.* **134**, 8148–8161 (2012).

169. Anthis, N. J. & Clore, G. M. Visualizing transient dark states by NMR spectroscopy. *Q. Rev. Biophys.* **48**, 35–116 (2015).

170. Korzhnev, D. M. *et al.* Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586–590 (2004).

171. Tolman, J. R. & Ruan, K. NMR residual dipolar couplings as probes of biomolecular dynamics. *Chem. Rev.* **106**, 1720–1736 (2006).

172. Clore, G. M. & Iwahara, J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **109**, 4108–4139 (2009).

173. Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553 (1979).

174. Zhang, Y. *et al.* High Pressure ZZ-Exchange NMR Reveals Key Features of Protein Folding Transition States. *J. Am. Chem. Soc.* **138**, 15260–15266 (2016).

175. Miloushev, V. Z. *et al.* Dynamic Properties of a Type II Cadherin Adhesive Domain: Implications for the Mechanism of Strand-Swapping of Classical Cadherins. *Structure* **16**, 1195–1205 (2008).

176. Kloiber, K., Spitzer, R., Grutsch, S., Kreutz, C. & Tollinger, M. Longitudinal exchange: An alternative strategy towards quantification of dynamics parameters in ZZ exchange spectroscopy. *J. Biomol. NMR* **51**, 123–129 (2011).

177. Zeeb, M. & Balbach, J. Protein folding studied by real-time NMR spectroscopy. *Methods* **34**, 65–74 (2004).

178. Rennella, E. & Brutscher, B. Fast Real-Time NMR Methods for Characterizing Short-Lived Molecular States. *ChemPhysChem* **14**, 3059–3070 (2013).

179. Englander, S. W. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521–655 (1983).

180. Bai, Y., Sosnick, T., Mayne, L. & Englander, S. Protein folding intermediates: native-state hydrogen exchange. *Science (80-. ).* **269**, 192–197 (1995).

181. Hvidt, A. & Nielsen, S. O. Hydrogen Exchange in Proteins. in *Advances in Protein Chemistry* **21**, 287–386 (1966).

182. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **17**, 75–86 (1993).

183. Skinner, J. J., Lim, W. K., Bédard, S., Black, B. E. & Englander, S. W. Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **21**, 996–1005 (2012).

184. Schanda, P., Forge, V. & Brutscher, B. Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11257–11262 (2007).

185. Hwang, T. L., Van Zijl, P. C. M. & Mori, S. Accurate quantitation of water-amide proton exchange rates using the Phase-Modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR* **11**, 221–226 (1998).

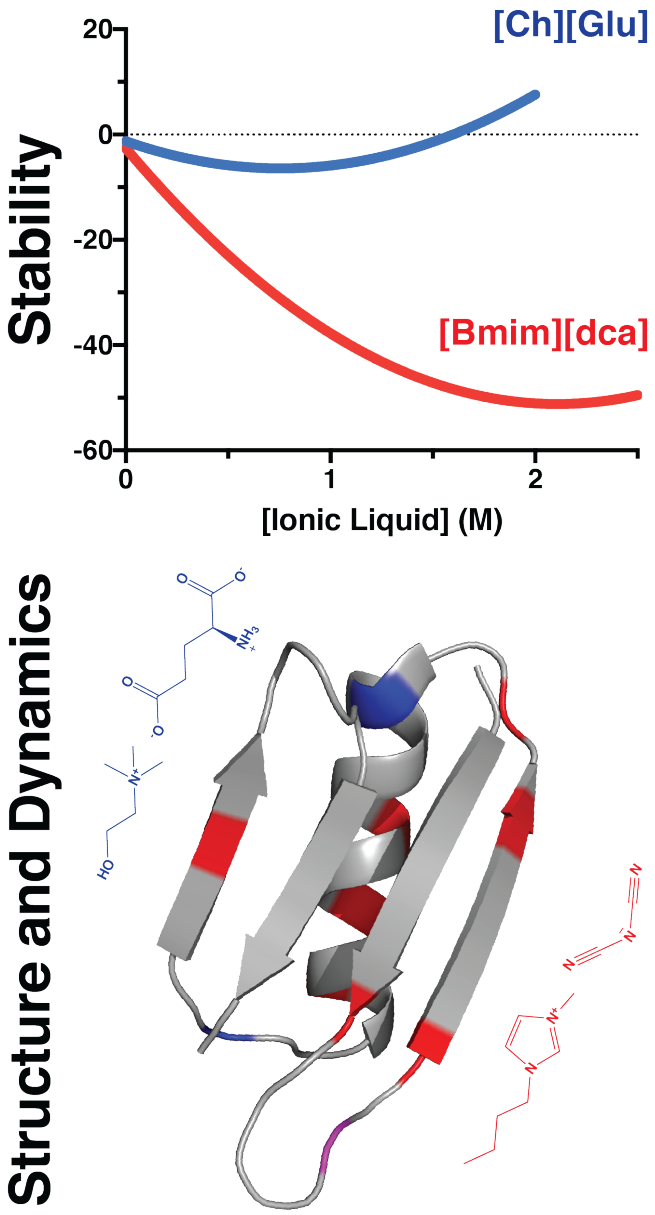
186. Chevelkov, V., Xue, Y., Krishna Rao, D., Forman-Kay, J. D. & Skrynnikov, N. R. 15NH/D-SOLEXSY experiment for accurate measurement of amide solvent exchange rates: application to denatured drkN SH3. *J. Biomol. NMR* **46**, 227–244 (2010).

187. Long, D., Bouvignies, G. & Kay, L. E. Measuring hydrogen exchange rates in invisible protein excited states. *Proc. Natl. Acad. Sci.* **111**, 8820–8825 (2014).

188. Orban, J., Alexander, P. & Bryan, P. Hydrogen-Deuterium Exchange in the Free and Immunoglobulin G-Bound Protein G B-Domain. *Biochemistry* **33**, 5702–5710 (1994).

2

# Chapter 2: Understanding the effects of ILs on GB1 stability



## Abstract

Ionic Liquids have emerged as powerful cosolutes in which depending on the nature of its ions are able to tune protein stability and activity. The mechanism of action towards protein (de)stabilisation is a balance between anion-cation, ion-solvent and ion-protein interactions. Despite much of this information revealed the tunability properties of ILs, there is still a lack of atomic resolution, in terms of how ions and ion-pairs interact with proteins towards (de)stabilisation. Based on that and bio-inspired by the most common small charged metabolites present in the cell milieu, in this chapter, I explored the Choline Glutamate ([Ch][Glu]) ionic liquid and studied its effects on the stability, structure and dynamics of the domain B1 of protein G (GB1) to obtain a complete mechanistic and molecular picture of all the forces in play. The interaction was investigated using a combination of fluorescence and NMR spectroscopy. First, I show that [Ch][Glu] affects GB1's stability by weakly binding to its charged, flexible and solvent-accessible residues. By comparing the data with that acquired in the presence of a known strong denaturant IL 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca]), I found a clearly distinct interaction pattern that allowed us to infer about the nature of the interactions in play in protein stabilisation and destabilisation in ionic liquids. The results from protein diffusion, exchange and dynamics derived from NMR-based experiments, as a function of [Ch][Glu] IL concentration, suggest a transition from a predominant preferential ions exclusion to its accumulation at the protein surface which possibly leads to an increase of GB1 stability at high IL concentrations. Furthermore, the data gathered is a clear evidence that charged metabolites found in cell milieu can form ion-pairs with the potential to modulate protein stability. This paves the way for further studies aiming at the understanding of the implications of variations in cellular homeostasis in health and disease.

Contributed to the publications:

**Fast NMR method to probe solvent accessibility and disordered regions in proteins.**

André F. Faustino, Glauce M. Barbosa, Micael Silva, Miguel A. R. B. Castanho, Andrea T. Da Poian, Eurico J. Cabrita, Nuno C. Santos, Fabio C. L. Almeida & Ivo C. Martins

*Scientific Reports*, **9,** 1647 (2019).

**Molecular basis of protein (de)stabilisation in Ionic Liquids.**

Micael S. Silva, Aldino Viegas, Angelo M. Figueiredo, Eurico J. Cabrita

*In preparation.*

## Introduction

In the last 15 years, a large number of ILs has been synthesized and their effects on protein stability have been thoroughly studied[1–6] leading to a substantial increase on our understanding of protein stability in aqueous-ILs. The observed effects have been attributed to characteristics such as polarity, hydrophobicity, hydrogen-bonding basicity, the ionic liquids' composition, structure, viscosity, or to the ions kosmotropicity/chaotropicity. Generally, the variations of the anionic moiety of ILs have a more dominant effect on protein properties than the cation variations, and its effects have mainly been linked to the Hofmeister series[4,7]. However, the Hofmeister effects tend to be protein-dependent and often contradictory[8–12]. For example, it has been noted that a particular IL that is favourable for a given protein may not be for the other. This highlights the importance of local interactions at the protein surfaces[13–15] (e.g., its composition and charge as well as site-specific interactions) and the need for a molecular understanding of the involvement of ILs in protein stability.

One of the reasons for some confusing results in the literature is the lack of attention to the IL concentration-dependent on protein stability[4,10,11,16]. Studies with lysozyme in varying concentrations of different cholinium-based ILs showed that while some anions stabilise the protein only above a certain threshold, others act as destabilisers at low concentrations and stabilisers at high concentrations[16]. This dual behaviour was reported for different proteins in different ILs, and experimental data shows that charged proteins undergo electrostatic interactions with ions at very low concentrations (e.g., valence-specific and non-specific coulombic effects), resulting in protein destabilisation at 0.25 and 0.5 M by almost all salts[10,17,18]. At high salt concentrations (≥ 1 M) electrostatic forces are screened and ion-specific effects become dominant (traditionally called Hofmeister effects as reviewed in the introduction). As a result, ions which bind to the protein surface and screen the surface charges at low concentration promoting salting-out, induce salting-in at higher concentrations by remaining hydrated at the protein surface[8]. In the case of stabilising salts, one therefore expects a low concentration dependence of $T_m$, while destabilising salts act in the same direction as the electrostatic effects and depress $T_m$. These observations seems to be true for ionic liquids effects.

Here, taking full advantage of the highly characterised immunoglobulin-binding domain B1 of *streptococcal* protein G (GB1)[19,20], I provide fundamental insights into the interactions between ILs and this protein. GB1 is well suited for these studies since it is small (56 residues, 6.2 kDa), highly stable, composed by a simple architecture (a four-strand β-sheet and an $\alpha$-helix surrounding its well-packed hydrophobic core – Fig. 2.1a and b, PDB: 2JSV[21]), and its structure, stability, dynamics, and folding kinetics have been extensively investigated[22–26].

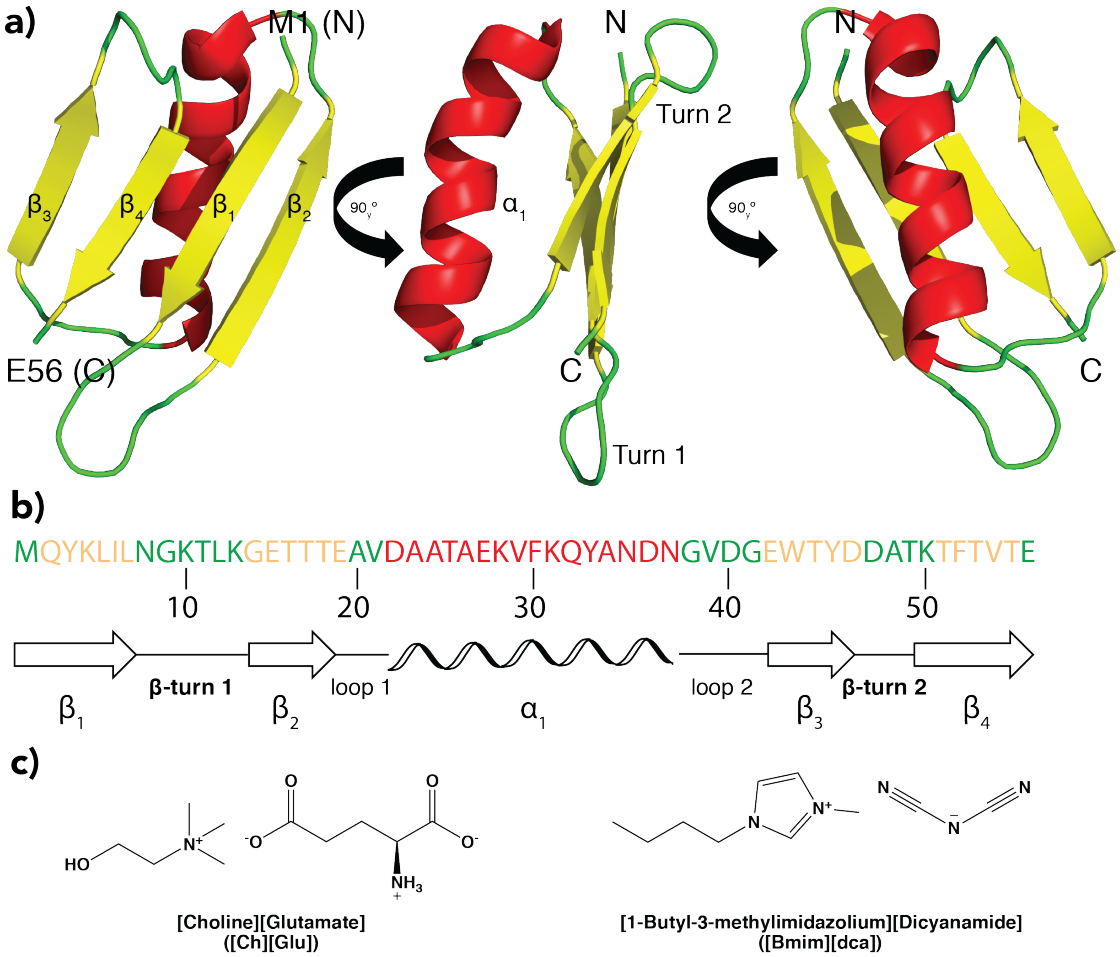



**Figure 2.1. The well-folded structure of GB1.**

**a**) Cartoon representation of the 3D structure of GB1 (PDB: 2JSV[21]) and coloured accordingly with its secondary structure (β-sheets, *α*-helix and loops/turns are coloured as green, red and yellow, respectively). Picture was rendered by using PyMOL. **b**) GB1 primary sequence (top) and its secondary structure diagram. β-sheets, *α*-helix and loops/turns are represented as arrows, spiral and lines, respectively, according to the cartoon representation **c**) Used ionic liquids: [choline][glutamate] ([Ch][Glu]), on the left, and [1-Butyl-3-methylimidazolium][dicyanamide] ([Bmim][dca]) on the right.

As GB1 is negatively charged at neutral pH, it is representative of most of the proteins in *E. coli*[27]. GB1's stability is pH-dependent, and it is most stable close to its isoelectric point (pI ≈ 4.3)[20,28,29], where a minimum of net electrostatic contributions are expected.

Using a GB1 variant with additional two negative charges, introduced by N8D and N37D mutations (see the primary wild-type sequence in Fig. 2.1b), Lindman et al[28] show that 2 M NaCl salt stabilises the protein at low pH (net charge +6, total charges 6) but not at neutral or high pH (net charge ≥ -6, total charges ≥ 18). Moreover, the addition of 0.15 M NaCl slightly decreases the stability the protein over the pH range (no change in $T_m$ for no-added and high salt conditions at neutral and high pH). The authors suggest that 2 M NaCl cannot eliminate electrostatic effects in highly charged systems because interactions between charged groups that are very close to each other are insensitive to salt screening and that 0.15

M destabilises this variant due to (1) an uneven shielding of charged in the folded and unfolded states (e.g., $pK_a$ values are different for the two populations) and/or (2) effects on hydrophobic effects, and/or that short- and long-range electrostatic effects that are opposing and screened to different extents. These observations corroborate the ideas debated in ref[10]. Protein stability is not just charge-dependent but also structure-dependent.

Although the existence of an on-pathway intermediate in the folding of GB1 has been previously proposed by continuous-flow ultrarapid mixing studies and molecular dynamics simulations[30–32], it is generally accepted that GB1 folds in a two-state manner. In its folding transition state ensemble, while the first β-turn and the α-helix (Fig. 2.1a and b) appear to be relatively disordered, the second β-hairpin (C-terminal) is largely formed with a native like-structure[33]. In fact, any non-native interactions present during the initial states of folding, such as the hydrophobic staple between V21 and A26 detected in a fragment of GB1[34], which can stabilise the unfolded state[35], can apparently be resolved without trapping the protein in long-lived states. Interestingly, at low pH and for a destabilised GB1 (i.e., the S14 variant which have L5M, L7V and V54I mutations), intermediate substates exist and are directly observed by NMR[25]. Backbone amide residual dipolar couplings (RDCs) indicate that: (1) prior to complete thermal unfolding, two melting hotspots (i.e., local backbone reorientations associated to destabilisation) are formed at the first β-turn around T11, L12 and K13 and the N-terminus of the helix at the A24 and T25; and (2) the first β-turn is thermally less stable than the α-helix and that helix unwinding starts at the N-terminus (see Fig. 2.1a).

In the present chapter, using mainly NMR spectroscopy-based methods and thermal denaturation studies, I have thoroughly investigated the specific and non-specific interactions between GB1 and ILs, in particular, using the stabiliser choline glutamate ([Ch][Glu], Fig. 2.1c-left) and the denaturant 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca], Fig. 2.1c-right). First, the (high) stability of GB1, in the presence of ILs and related salts, was monitored by its intrinsic tryptophan fluorescence changes upon thermal unfolding. Next, IL-GB1 interactions were probed on the backbone and on the charged side chains using chemical shift perturbations (e.g., based on $^{1}H$-$^{15}N$ HSQC, $^{1}H^{\varepsilon}$-$^{15}N^{\zeta}$ H2(C)N and $H^{\beta/\gamma}$-$^{13}CO^{\gamma/\delta}$ H(C)CO spectra). While the results show that denaturation induced by [Bmim][dca] is largely explained by preferential hydrophobic interactions with the hydrophobic core of GB1, the (de)stabilisation of GB1 with the increase of [Ch][Glu] concentration is a shallow parabola and it is not explained exclusively on the weak interactions observed for charged, flexible and solvent-accessible regions of the protein. Therefore, the GB1 structure, diffusion, dynamics and solvent accessibility were comprehensively investigated by NMR in the presence of diverse [Ch][Glu] concentrations at different temperatures and magnetic fields.

This work has enabled me to address that stabiliser ILs could act differently from protective osmolytes (acting uniquely by preferential hydration[36]), highlighting the presence of solvent-mediated interactions and that resulting changes on protein hydration could be very relevant in terms of improving protein stability. Interestingly, the local changes observed for A23 residue (melting hotspot region), in terms of chemical shift and its protection from solvent, induced by the anion seem to be crucial to maintain GB1's stability. On the other hand, as a residual α-helical structure is detected in the denatured-[Bmim][dca] state of GB1, the unfolded ensemble cannot be disregarded (approached in the next chapter 3). The

fundamental insights provided by this work will aid the understanding, development, and application of ILs as advanced reaction and processing media.

## Results and Discussion

### ILs and salts effects on protein stability

To understand the effects of aqueous solutions of [Ch][Glu] and [Bmim][dca] ILs on protein stability, I first investigated the thermal stability of GB1 under different concentrations of the ILs. This was achieved by monitoring GB1's intrinsic tryptophan fluorescence changes upon thermal unfolding in order to determine the melting temperature ($T_m$), as shown in Fig. 2.2.

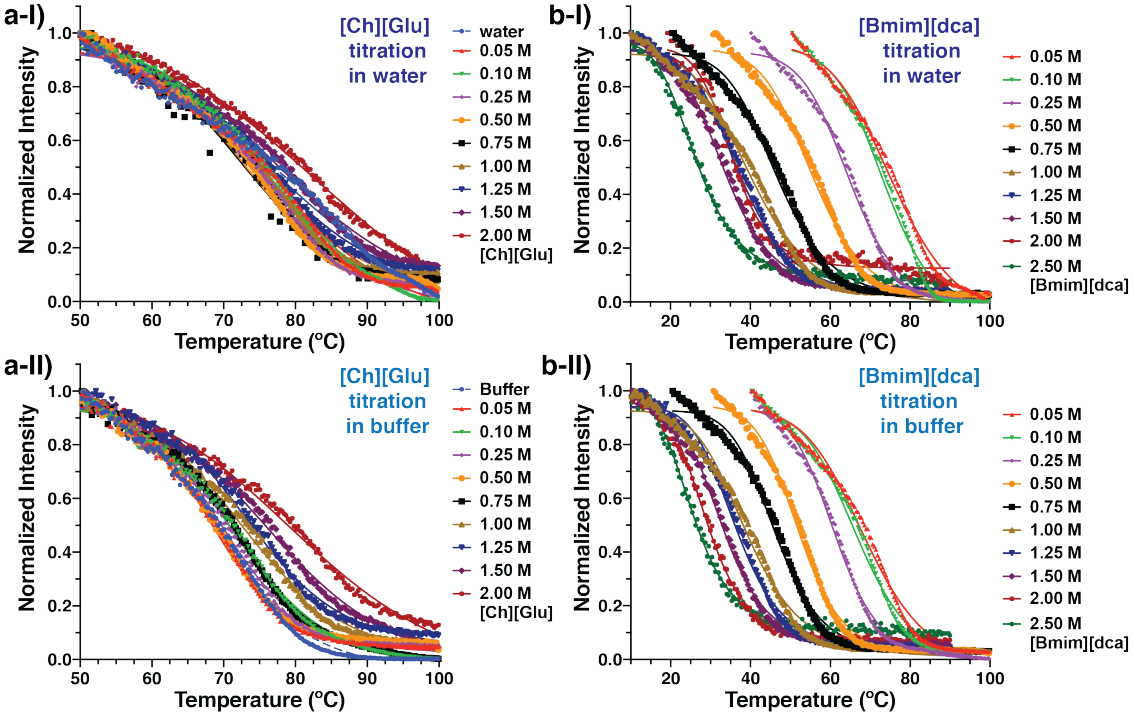



**Figure 2.2. Fluorescence thermal denaturation studies of GB1 with ILs.**

Normalized maximum Trp43 intrinsic fluorescence intensity ($\lambda_{ex}$=280 nm, $\lambda_{em}$=350 nm) from 10 or 50°C to 100°C with 1°C/min rate for GB1 at 52 μM of concentration in the presence of increasing **a)** [Ch][Glu] or **b)** [Bmim][dca] IL concentrations from 0 to 2.0 M or 2.5 M. Solutions were prepared from each solution-stock in **I)** *milliQ* water pH 6.5 or **II)** buffer 0.05 M sodium phosphate pH 7.2.

Assuming a two-state protein folding, $T_m$ is defined as the temperature at which both folded ($p_f$) and unfolded ($p_f$) states are equally populated at equilibrium ($p_f = p_u = 0.5$) and can be determined from the analysis of the plots of the normalized maximum Trp43 intrinsic fluorescence intensity as function of temperature. ILs titrations were performed both in water (pH = 6.5) and in a 0.05 M sodium phosphate buffered solution pH 7.2 (Fig. 2.2-I and 2.2-II, respectively). As can be seen in Fig. 2.3, the $T_m$ obtained in the presence of ILs for both

conditions (water and buffer) are different, particularly in the low and high concentration regions. When comparing the water *versus* buffer conditions in the absence of ILs, GB1 is slightly destabilised by the simple presence of sodium phosphate. Although GB1 stability is dependent on the pH, being more stable near its isoelectric point (pI = 4.3 ± 0.1[28,37]), the small difference in the water/buffer solution pH does not explain the change in $T_m$, and that should be a consequence of destabilising electrostatic interactions with sodium and phosphate ions[38]. However, when comparing the effects of ILs in water and buffer the overall profile with concentration is similar: (1) for [Ch][Glu] a shallow minima of $T_m$ near 0.5 M (Fig. 2.3a) is observed and (2) for [Bmim][dca] $T_m$ strongly decreases with increasing concentration which leads to a $T_m$ of ~ 26 ºC at 2.5 M (Fig. 2.3b).

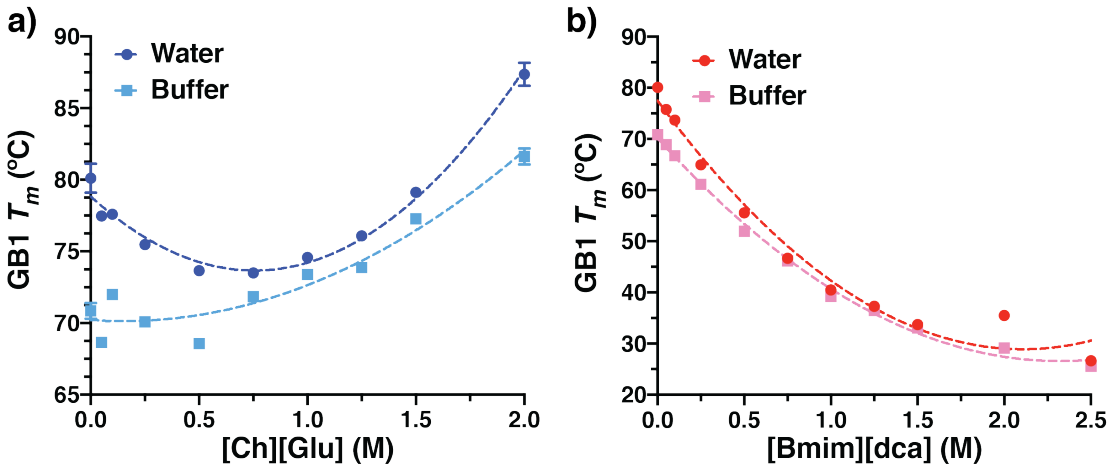



**Figure 2.3. Thermal stability of GB1 in the presence of ILs.**

$T_m$ of GB1 in water or buffer 0.05 M sodium phosphate pH 7.2 in the presence of increasing **a)** [Ch][Glu] (at blue) or **b)** [Bmim][dca] (at red) ILs concentration. $T_m$ values were determined by fitting the values from Trp43 intrinsic fluorescence intensity decay to a sigmoidal function and the errors were obtained directly from the fitting (if the error bar is shorter than the size of the symbol, it would not appear). The dashed lines are indicative only.

To differentiate between a specific IL effect (cation, anion, or ion-pair) or a general ionic strength effect, the experiments were repeated in water with related salts containing the relevant ionic moieties for each IL, i.e., in the presence of cholinium chloride ($[Ch]^+$), sodium glutamate ($[Glu]^-$), and sodium dicyanamide ($[dca]^-$) salts, using NaCl as a reference. For the 1-Butyl-3-methylimidazolium ($[Bmim]^+$) cation, the IL [Bmim]Cl was used. For each cosolute titration, $T_m$ values and their uncertainties can be found in Table A1 of Appendix A. By plotting the cosolute-induced shift, $\Delta T_m = T_{m,\,\text{cosolute}} - T_{m,\,\text{ref}}$, as a function of the IL or salt concentration, and comparing the obtained effects I can infer about the individual contribution of each ion for the observed effect on GB1's $T_m$ (Fig. 2.4). Unless otherwise stated, $T_{m,\,\text{ref}} = T_{m,\,\text{water}}$ corresponds to the protein $T_m$ measured in water (in the absence of cosolute).

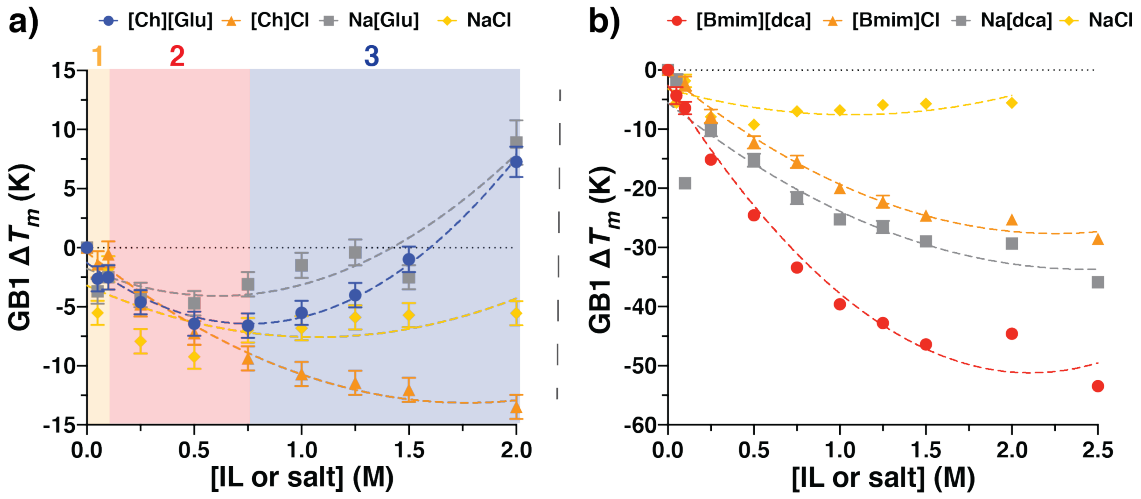



**Figure 2.4: Cosolute-induced (de)stabilisation of GB1.**

Effect of **a)** [Ch][Glu] or **b)** [Bmim][dca] and related salts in the $T_m$ of GB1 in water [$\Delta T_m = T_{m,\,\text{cosolute}} - T_{m,\,\text{water}}$], and the uncertainties were propagated from the fitting error for each condition [$\sigma = \sqrt{(\delta_{Tm,\text{cosolute}})^2 + (\delta_{Tm,\text{water}})^2}$]. The dashed lines are indicative only.

Considering the stabilisation of anionic GB1, the tested cosolutes can be ranked with the stability order (measured at 2 M of co-solute) Na[Glu] ≈ [Ch][Glu] > NaCl > [Ch]Cl > [Bmim]Cl > Na[dca] ≫ [Bmim][dca] (Table 2.1). The resulting stability series upon variation of the anion is given as $Glu^- > Cl^- > dca^-$ and with the cation by $Na^+ > Ch^+ > Bmim^+$. Although the "Hofmeister series" are discussed controversially and there are exceptions[13] (as previously reviewed in the introductory section of the role of cosolute-protein interactions), the resulting anion and cation rankings of ILs and salts for GB1 stability remains valid according to the Hofmeister ordering and are compatible with those obtained for other negatively and positively charged proteins with cholinium- and/or imidazolium-based ILs[3,4,6,11,16,39,40].

**Table 2.1: $\Delta T_m$ of GB1 at 2 M of cosolute concentration.**

| **Cosolute** | Na[Glu] | **[Ch][Glu]** | NaCl | [Ch]Cl | [Bmim][Cl] | Na[dca] | **[Bmim][dca]** |
|---|---|---|---|---|---|---|---|
| **$\Delta T_m$ $^{2\,M}$ (K)** | 9 ± 2 | **7 ± 1** | -6 ± 1 | -13 ± 1 | -25 ± 1 | -29 ± 1 | **-45 ± 1** |

$\Delta T_m$ of GB1 of each cosolute at 2.0 M of concentration was determined using $T_m$ of water as reference, and the uncertainty was propagated from the $T_m$ fitting error.

In general, even for a negatively charged protein as GB1 at neutral pH, the anionic effects seem to prevail over cationic effects. [Ch][Glu] IL stabilising effect seems to result mainly from the interaction between protein and anion and not from the cation or ion-pair (yet the existence of a transient Ch-Glu ion-pair in solution is possible[41]). Since $[Ch]^+$ has been shown to represent weak propensity to the protein surface and to prefers bulk hydration[42,43], I predict that the anion is more active in the proteins' surface. Accordingly, the stabilising effect of $[Glu]^-$ should results from unfavourable interactions with hydrocarbons groups and amide oxygens exposed in unfolding (general exclusion from the water of hydration of the protein) as suggested for the effects of K[Glu] and Na[Glu] salts on proteins[44]. Furthermore,

the higher viscosity of ILs (as revealed for [Ch][Glu]) suggests a strong strength of hydration by the ions (increasing surface tension and salting-out)[45]. On the other hand, both the $[Bmim]^+$ cation and the $[dca]^-$ anion seem to be equally destabilisers. Because the imidazolium cation is a hydrophobic surface-active specie and $[dca]^-$ is a large polarizable anion with low charge density, both ions have high propensity for non-polar residues and the ability to lower protein-water interfacial tension[40]. Interestingly, their ion combination results in an extraordinarily strong denaturant and a strong ion-pair must be considered to exist in the solution[6]. In a similar way that hetero-ion pairing of $[Gdm]^+$ with stabilising $SO_4^{2-}$ anions reduce its denaturant activity (e.g., by attenuating $[Gdm]^+$-peptide hydrogen bonds)[43,46], it is reasonable to assume that $[Bmim]^+$ cations have a higher denaturing power by significant ion-pairing with $[dca]^-$ anions, and vice-versa.

GB1 establishes electrostatic interactions with ions at low concentrations due to a screening of surface charges that results primarily from electronic polarization, reorientation of dipolar groups and changes in the concentrations of ions in the vicinity of dipolar groups[47]. As summarised in the introductory section of the role of cosolute-protein interactions, this charge screening could lead to protein destabilisation, as in this case, for concentrations between 0.25 M and 0.75 M of [Ch][Glu] and related salts) (Fig. 2.4a, region 2). Such behaviour could be explained by uneven shielding of favourable and unfavourable electrostatic interactions, that is, a balance between the shield of unfavourable charge repulsions and the screening of attractive electrostatic interactions, as observed for GB1 in the presence of NaCl salt[28,48]. As for other stabilising salts, a minimum in the concentration dependence of $T_m$ is expected for charged proteins[10,17,18]. Particularly, this non-monotonic and dual behaviour of $\Delta T_m$ has been reported for cholinium-based ILs[16]. Data for [Ch][Glu], Na[Glu], [Ch]Cl and NaCl shows that at high concentrations ($\geq$ 1 M) ion-specific effects become dominant as the $\Delta T_m$ values start to be significantly different (Fig. 2.4a, region 3). Because charges are effectively screened, these specific effects are now more dependent on the ion nature and subsequent mechanisms of interactions (direct binding, water surface tension effects). [Ch]Cl destabilises GB1 along the full titration by acting in the same direction as the electrostatic effects and depressing $T_m$. This outcome shows that the global effect of [Ch][Glu] (especially $\geq$ 1 M) is not simply the combination of the individual $Glu^-$ and $Ch^+$ ions, as their combination should result in a weaker stabiliser. For stronger destabilising ILs and salts, such as [Bmim][dca], [Bmim]Cl and Na[dca], $T_m$ decreases monotonously and leads to largely negative $\Delta T_m$ (Fig. 2.4b).

Other aspect to consider, as outlined in the introduction of this chapter, it is the fact that GB1 stability is pH-dependent (with the highest stability reached at pH close to the protein pI = 4.3 $\pm$ 0.1)[28]. In order to disregard the effects of cosolutes on protein stability due to a change in the overall solution pH, I measured the pH value for each condition (Fig. 2.5).

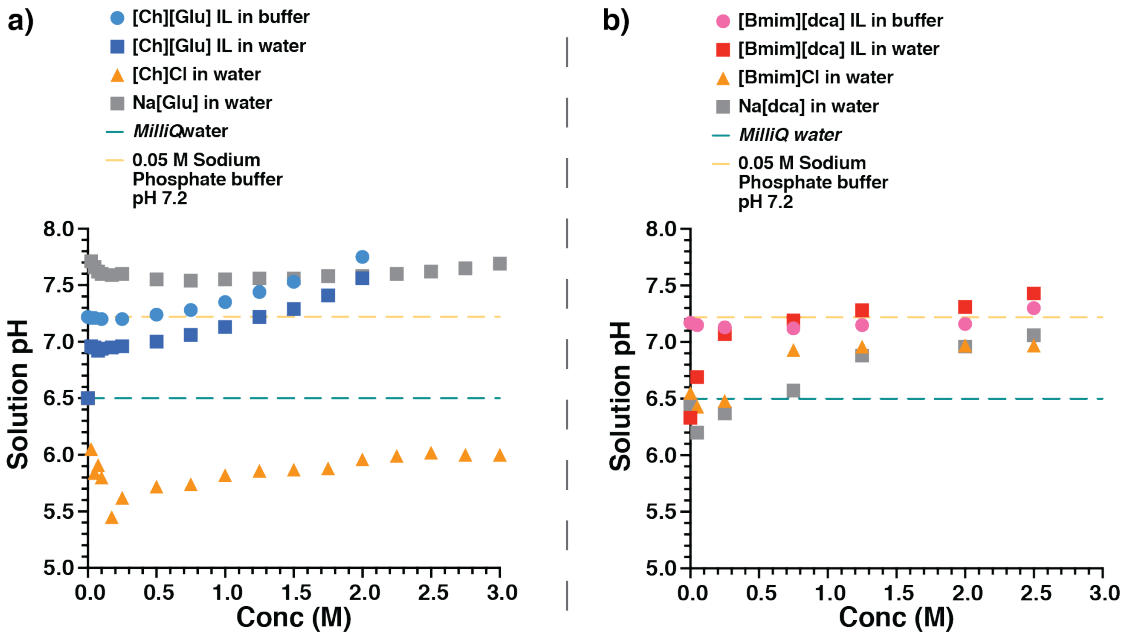



**Figure 2.5. Aqueous-solution pH variation with cosolutes.**

Similarly, to fluorescence experiments, stock solutions for each cosolute ([Ch][Glu], [Bmim][dca], [Bmim]Cl, [Ch]Cl, Na[Glu], Na[dca], and NaCl) were prepared in *milliQ* water and then diluted for the desired concentration. Additionally [Ch][Glu] and [Bmim][dca] stock solution) and their dilutions were also prepared in buffer 0.05 M sodium phosphate pH 7.22 ± 0.05. For those stock solutions with pH far from 7.2, it was adjusted to maintain the solution pH in the range of 7.2 ± 0.2 with negligible microliter addition of HCl or NaOH solutions.

Upon a small addition of [Ch][Glu] (≈ 25 mM), the pH of the protein solution in water (pH ≈ 6.5) jumps to a controlled solution pH (pH ≈ 7.0) similar to an addition of a traditional aqueous buffer. Since [Ch][Glu] is formed by neutralisation of glutamic acid (triprotic amino acid with $pK_a^{aq}$ values: 1.9, 4.3, 9.5) with [Ch][OH] (1:1 moles), the anion is deprotonated and provides an intrinsic proton buffering action, in a similar way that phosphate species ($H_2PO_4^{2-}/HPO_4^{-}$) does in [Ch][dhp] IL or [tartarate]$^{2-}$ anion in [Ch][Tar] IL[49]. Adding increasing amounts of [Ch][Glu] to a solution of GB1 in water leads to a constant increase in the pH from 7.0 to 7.6 (0.025 to 2.0 M, Fig. 2.5a). At high concentration of [Ch][Glu] (> 1.5 M), the protein is stabilised despite the increase of pH that would otherwise lead to protein destabilisation[28] (pH > 7.3, far from the pI). For Na[Glu] the pH remains relatively constant at about 7.6 throughout the titration. Furthermore, while [Ch]Cl leads to a general acidification of the solution, I observe this salt destabilises GB1 (Fig. 2.4a). On the other hand, [Bmim][dca] is slightly basic, by virtue of the distinct basicity of dicyanamide anion and can be expected to deprotonate acidic solute species[50]. [Bmim][Cl] is moderately acidic. For these imidazolium-based ILs and Na[dca] salt, the pH of each stock-solution was adjusted to pH ≈ 7.2 (see Materials and Methods for details). Their variation in water from pH ≈ 6.5 to ≈ 7.2 (0.05 M to 2.5 M, as shown in Fig. 2.5b, using the stock-solution with adjusted pH) are considered negligible due to the large variation of $T_m$.

The stability of a highly charged cationic protein, lysozyme (pI ≈ 11), was also studied in order to infer about the effect of the overall protein net charge. Lysozyme is a well-established model for protein stability studies, and it has been studied with cholinium-based and/or imidazolium-based ILs[16,39,51–53]. I monitored the effect of [Ch][Glu] and related salts ([Ch]Cl, Na[Glu] and NaCl) on lysozyme $T_m$ using differential scanning fluorimetry (DSF)[54] (Fig. 2.6). The lysozyme's thermal stability seems to increase linearly with [Ch][Glu] concentration in water and in buffer (Fig. 2.6a). As observed for GB1, sodium phosphate buffer destabilises the protein with respect to water and in dilute IL solution (< 0.1 M). For the related salts (all in buffer), lysozyme $T_m$ increases, even in the presence of [Ch]Cl, which after an insignificant decrease for concentrations < 1.5 M, leads to a positive $\Delta T_m$ – Fig. 2.6b. While the stability ranking for lysozyme is the same as that of GB1, the results show that [Ch][Glu] is a better stabiliser than Na[Glu], particularly at concentrations above 1.5 M, and that $[Glu]^-$ anion is the main responsible for the overall stabilisation. This is a further indication that the overall protein net charge is not a key factor for the observed stabilisation effect of [Ch][Glu] and emphasises the importance of hydration forces on ion specificity[55]. Because lysozyme's stability behaviour is linearly dependent with the IL or salt concentration, the ion-specific effects are not only dominant at low and high concentrations (comparing to electrostatic screening) but also too strong to trace a minimum in $T_m$.

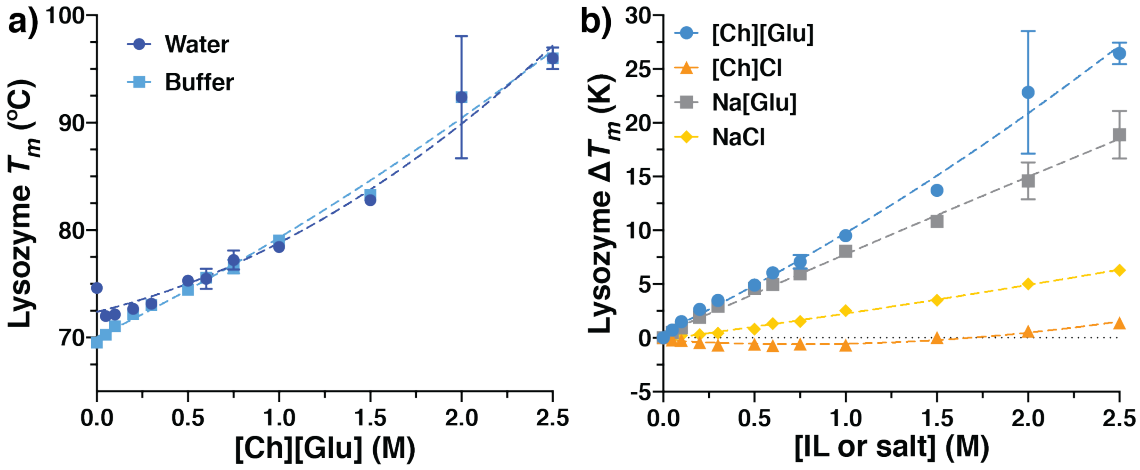



**Figure 2.6. Cosolute-induced (de)stabilisation of positively charged Lysozyme.**

(**a**) $T_m$ of Lysozyme from chicken egg white (15 μM of protein concentration) in water or buffer 0.05 M sodium phosphate pH 7.2 in the presence of increasing [Ch][Glu] IL concentration. For each condition, the protein's thermal shift was analysed with Protein Thermal Shift™ Software v1.1 (Applied Biosystems), the transition midpoint ($T_m$) of the sigmoidal curve was calculated automatically, and the error was measured from the triplicate of each condition. Note that the protein thermal shift arises from the fluorescence of the dye Sypro Orange that is dramatically increased upon binding to hydrophobic patches during protein unfolding, and the post-transition decrease in fluorescence is due to events occurring at high temperatures (e.g., temperature dependence of fluorescence, protein aggregation and probe dissociation). (**b**) Effect of [Ch][Glu] and related salts in the $T_m$ of Lysozyme in buffer [$\Delta T_m = T_{m,\,\text{cosolute}} - T_{m,\,\text{buffer}}$], and the uncertainties were propagated from the triplicate error for each condition. The dashed lines are indicative only.

## Ion specific interactions

Having established that ILs are able to modulate protein thermal stability, acting either as stabilisers ([Ch][Glu]) or destabilisers ([Bmim][dca]), I proceeded to study the molecular features that define this modulation. Previous work[6,56] showed that the nature of the interactions between proteins and imidazolium-based ILs include electrostatic and hydrophobic interactions (e.g., coulomb *versus* van der Waals forces), and that both are crucial to understand the effects of ILs on the stability of proteins. Thus, I proceed to identifying and mapping the IL-protein interactions. The specific ion effects are primarily described for [Ch][Glu] and, later, for [Bmim][dca].

### [Ch][Glu]-GB1 interactions

First, I followed the backbone chemical shift perturbations (CSP) of GB1 in the presence of increasing amounts of [Ch][Glu] (up to 1.5 M of concentration) by acquiring $^{1}H$-$^{15}N$ HSQC spectra at each titration point (Fig. 2.7a). In parallel, complementary 2D $^{1}H^{\varepsilon}$-$^{15}N^{\zeta}$ H2(C)N[57] and $H^{\beta/\gamma}$-$^{13}CO^{\gamma/\delta}$ H(C)CO[58] correlation spectra were also collected (Fig. 2.7b, c).

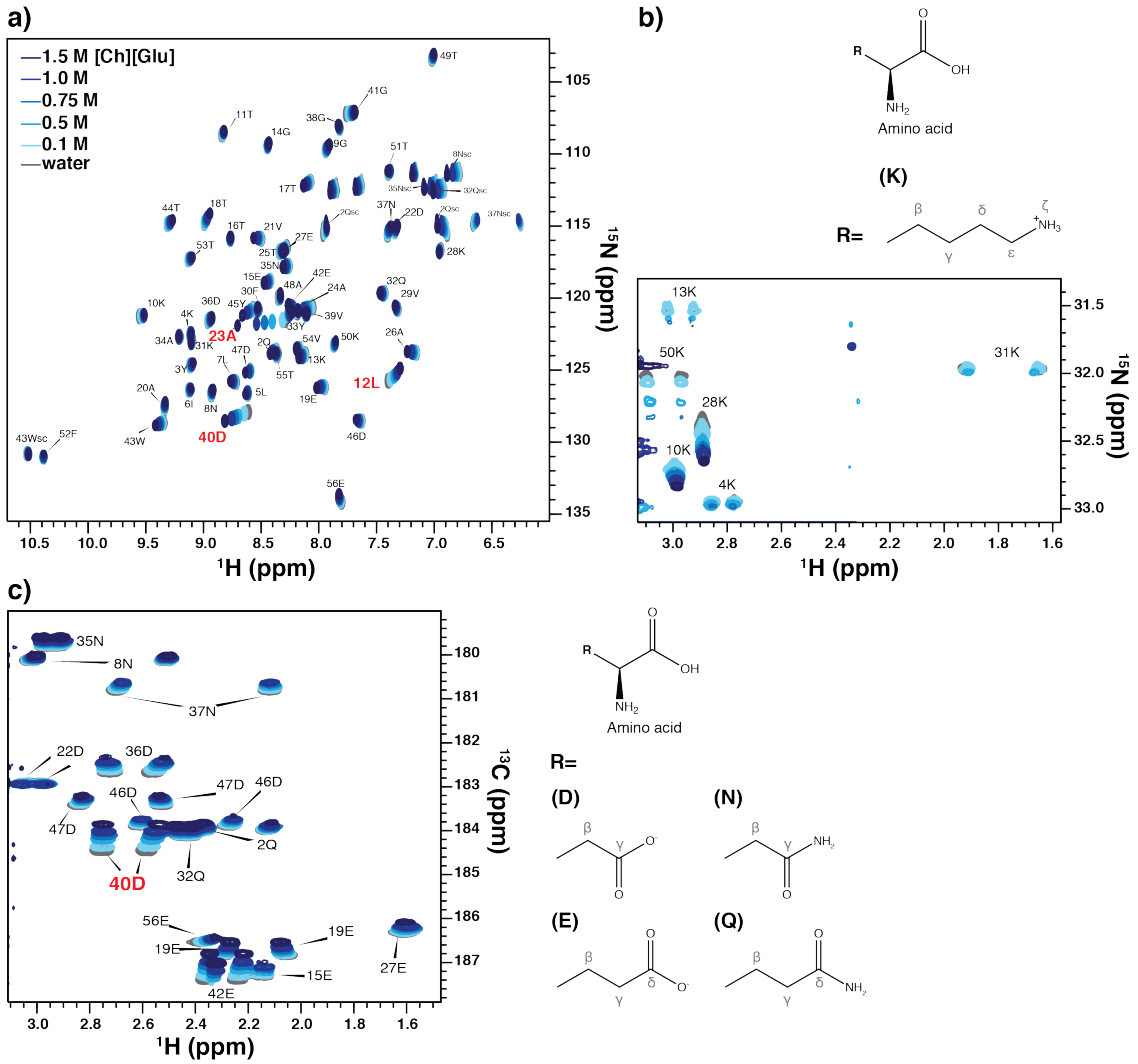



**Figure 2.7. HSQC and complementary spectra of GB1 in the presence of [Ch][Glu].**

**a**) Overlay of 2D $^{1}H$-$^{15}N$ HSQC spectra of GB1's backbone amide groups. **b**) Overlay of $^{1}H$-$^{15}N$ H2(C)N spectra of GB1's Lys $NH_3^+$ groups displaying the $^{1}H^{\varepsilon}$ and $^{15}N^{\zeta}$ correlation. **c**) Overlay of $^{1}H$-$^{13}C$ H(C)CO spectra of GB1's Asx and Glx carboxyl and carbonyl groups displaying the $^{1}H^{\beta/\gamma}$ and $^{13}CO^{\gamma/\delta}$ correlation. Data were acquired with 0.6 mM $^{15}N/^{13}C$ isotopically enriched GB1 with increasing amounts of [Ch][Glu] (0 (grey), 0.1 M, 0.5 M, 0.75 M, 1.0 M, 1.5 M (dark blue)) in 90% $H_2O$/ 10% $D_2O$, 0.1% $NaN_3$, pH 7.1 ± 0.1 with 50 μM DSS as internal reference at 298.2 K, 600.13 MHz. The labels show the assignments for all residues (sc, side chain), in which, the red labels indicate the most affected residues. The inlay in b) and c) indicates the schematic diagram for the side chains of amino acids under analysis.

Because protein/IL (or salt) interactions are, in general, weak and non-specific, these spectra were used to probe charged side chain perturbations (expected to be more sensible), following $^{15}N$ chemical shift of GB1's Lys side chain $NH_3^+$ groups through non-exchanging carbon-bound protons ($^{1}H^{\varepsilon}$-$^{13}C^{\varepsilon}$), or $^{13}C$ nuclei of GB1's Asx/Glx side chain carboxyl and carbonyl groups by way of $^{13}CH_2^{\beta,\gamma}$, respectively (see Materials and Methods for further details). Backbone CSPs were analysed in terms of the combined $^{1}H_N$ and $^{15}N_H$ chemical shift differences of the backbone amide, $\Delta\delta_{comb}$[59]. The chemical shifts were found to vary linearly with IL concentration with no evidence of saturation (Fig. 2.8a) indicating a weak interaction between [Ch][Glu] and GB1. To rank the more affected residues, a cut-off line[60] was calculated, and the $\Delta\delta_{comb}$ of backbone amide from G9, L12, T17, T18, E19, V21, A23, A24, A26, N37, D40, G41, E42, W43, Y45, and E56 residues were identified with a shift above the calculated threshold (Fig. 2.8b). Those residues more affected were mapped onto the 3D structure of GB1 (Fig. 2.8c), revealing that the affected residues are mainly solvent exposed and located in the helix terminals, β-strand 3 and in the loops, in or near charged surface areas. Of all the affected residues, there are three whose CSP stands out: L12, A23 and D40; which are located in the loop connecting β-strands 1 and 2, in the N-terminal end of the $\alpha$-helix, and in the loop connecting the $\alpha$-helix and β-strand 3, respectively.

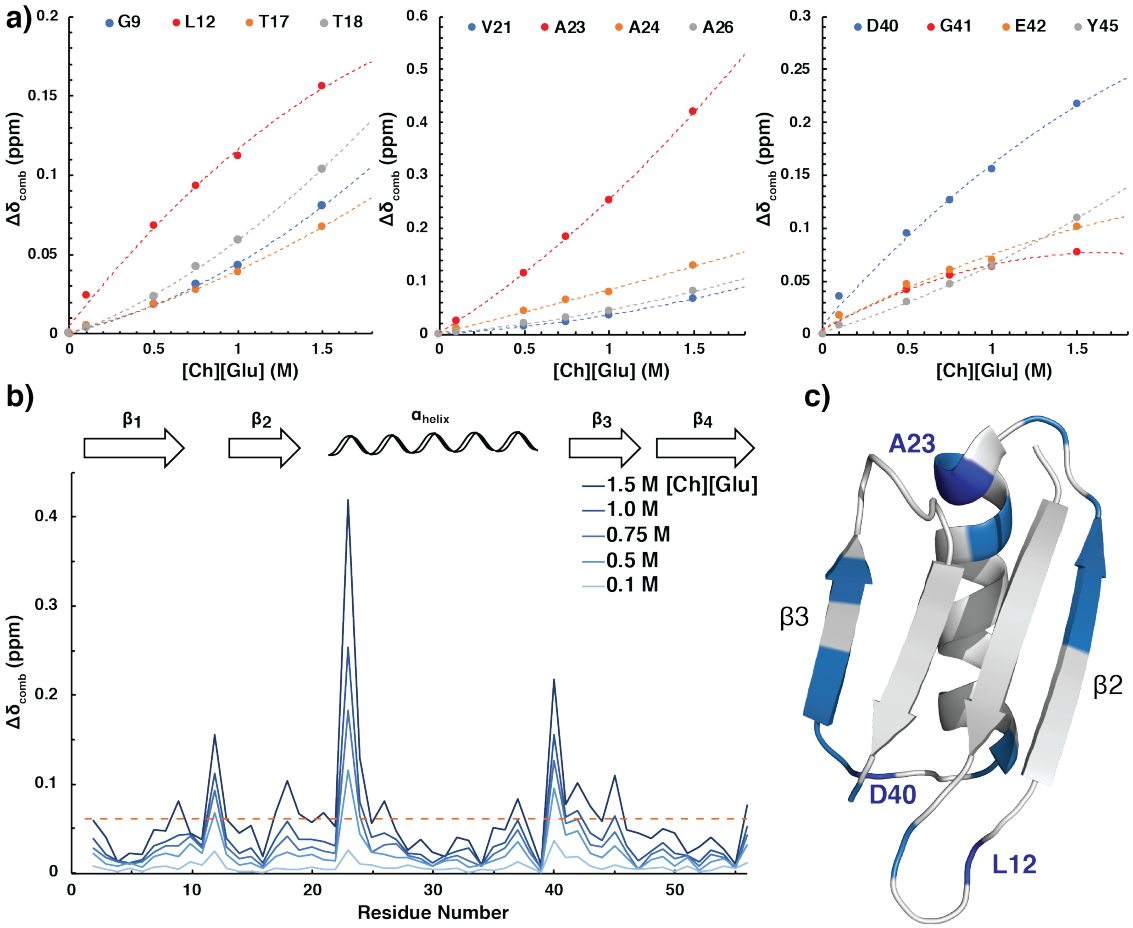



**Figure 2.8. Backbone CSPs as a function of [Ch][Glu] concentration.**

**a)** Plot of $\Delta\delta_{comb}$ as function of [Ch][Glu] for the most affected residues. The dashed lines are only indicative (polynomial with two variables were found to better fit the experimental data). **b**) Overlay of the combined chemical shift of GB1 in the presence of increasing [Ch][Glu] concentrations (dark blue). The threshold that distinguishes the affected from non-affected residues is represented by the orange dashed line for 1.5 M [Ch][Glu]. Above the plot it is depicted the secondary structure of the protein. The combined chemical shifts were calculated against the chemical shifts in water. **c**) The affected residues are mapped onto the respective 3D structure (PDB: 2JSV[21]).

Concerning the perturbations by the IL on the charged side chains (Lys and Asp/Glu residues), as shown in Fig. 2.9, both deviations in $^{15}N^{\zeta}$ chemical shifts from $NH_3^+$ groups of Lys and $^{13}CO^{\gamma/\delta}$ chemical shifts from carbonyl groups of Asp/Glu are in line with CSPs observed for the backbone amide.

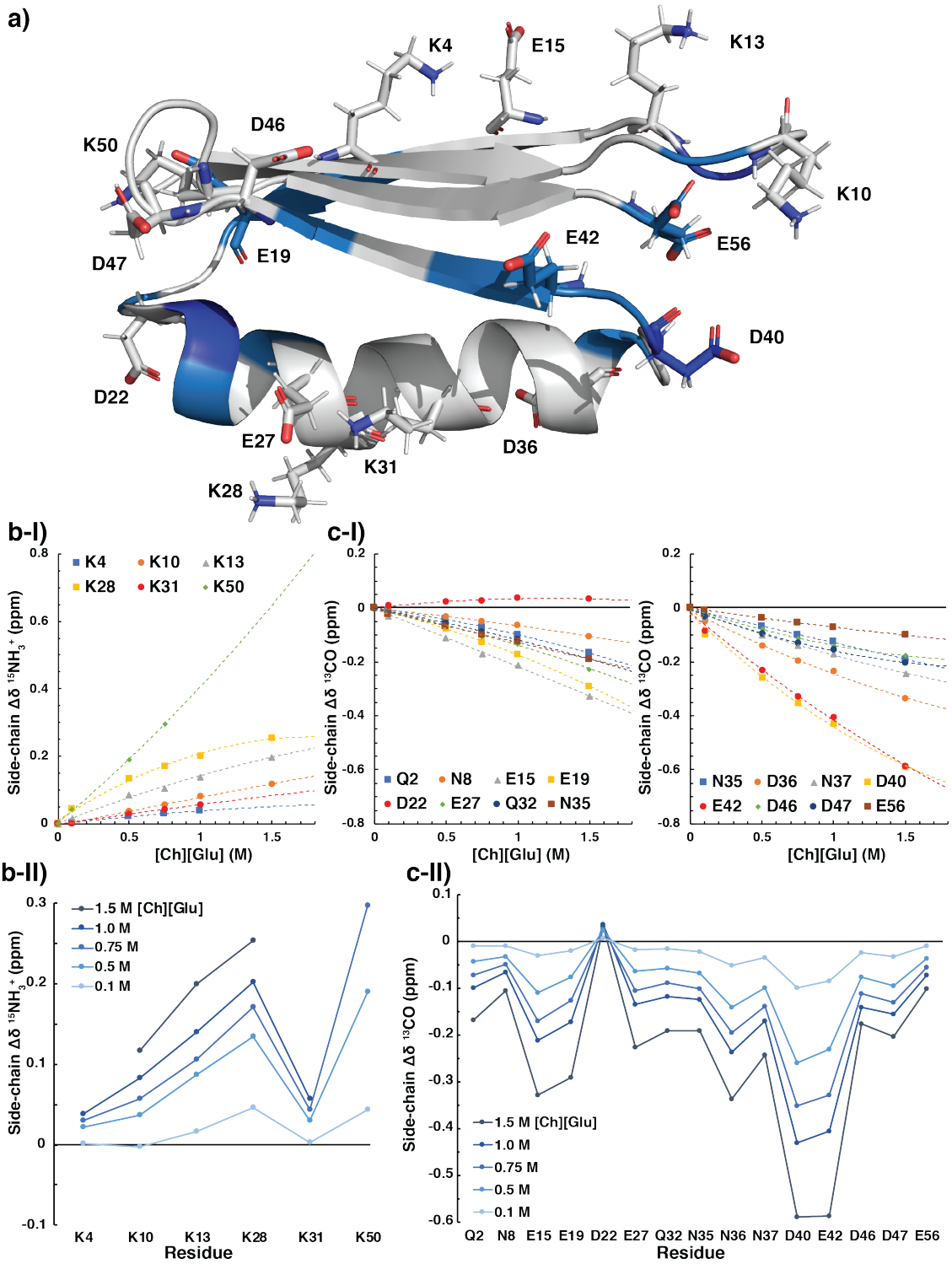



**Figure 2.9: Side chains CSPs as a function of [Ch][Glu] concentration.**

**a**) Structure of GB1 (PDB: 2JSV[21]) showing all lysine side chains and all aspartate and glutamate side chains (in sticks). The affected residues in the backbone are mapped at blue. **b**) Difference of $^{15}N$ chemical shifts of $NH_3^+$ groups ($\Delta\delta\ ^{15}NH_3^+$) of side chains for [Ch][Glu] titration, based on the $^1H^{\varepsilon}$-$^{15}N^{\zeta}$ H2(C)N spectra, as I) a plot of $\Delta\delta\ ^{15}NH_3^+$ as function of [Ch][Glu] and as II) ) overlay of $\Delta\delta\ ^{15}NH_3^+$ at each IL concentration. **c**) Difference of $^{13}C$ chemical shifts of carbonyl/carboxyl groups ($\Delta\delta\ ^{13}CO$) of side chains for [Ch][Glu] titration, based on $^1H^{\beta/\gamma}$-$^{13}CO^{\gamma/\delta}$ H(C)CO spectra, as a I) a plot of $\Delta\delta\ ^{13}CO$ as function of [Ch][Glu] and as II) ) overlay

of $\Delta\delta$ $^{13}CO$ at each IL concentration. The dashed lines in I) are only indicative (polynomial with two variables were found to better fit the experimental data).

Although some peaks in the H2(C)N spectrum disappeared at high IL concentrations (K4, K31, K50), the magnitude of the $\Delta\delta^{15}N^{\zeta}$ (to downfield) indicates that the $NH_3^+$ groups more affected are those that are found to be more spatially close to the affected backbone segments (Fig. 2.9a): K50 and in lesser extent K28, K13 and K10. Interestingly, with exception of K50, the Lys side chains less affected are K4 and K31 that are more prone to form salt bridges with carboxylate side chains and so less available to interact with ions (K4-E15, E27-K31 and D47-K50 are consistently present in the crystal structure of GB1[61]). Nonetheless, only the intrahelical E27-K31 interaction should be exist in the solution[62] which partially explains its absence of deviation in $\delta^{15}N^{\zeta}$. The general deviations to upfield in the $^{13}CO^{\gamma/\delta}$ chemical shifts from carbonyl/carboxyl groups of Asx/Glx (Fig. 2.9c) clearly indicate the carboxylate group of D40 and E42 are significantly affected by the presence of [Ch][Glu]. Since interactions at the protein surface could be stronger than those with the backbone[14], the perturbation in the amide group of D40/E42 residues probably derives from interaction of its anionic side chain with the $[Ch]^+$ cation.

Since the appearance of NMR resonances are sensitive to chemical exchange across a wide range of timescales[63], during a protein-ligand titration experiment, it depends on the rate of exchange, $k_{ex}$, between free and bound states relative to the frequency difference, $\Delta\omega$. If exchange is fast ($k_{ex} \gg \Delta\omega$), then a single resonance will be observed with a population-weighted average chemical shift, while if exchange is slow ($k_{ex} \ll \Delta\omega$), then two signals will be observed at the chemical shifts of the free and bound states, weighted according to the bound population. If the exchange rate is comparable to the frequency difference ($k_{ex} \approx \Delta\omega$), extensive line broadening is observed (intermediate exchange). Even in the case of an extreme fast exchange limit (associated with very weak interactions), where NMR titration data is typically analysed in terms of CSPs alone (as discussed above, Fig. 2.8), analysis of peak broadening is indicative of intermediate-to-slow exchange and/or slow tumbling[59].

Lineshape analysis of the peaks[64] in the [Ch][Glu]-GB1 NMR spectra (Fig. 2.7), by means of peak height intensity, shows a global [Ch][Glu] viscosity-dependence with some residues more affected than others (Fig. 2.10). Residues located in unstructured/loop regions connecting β1-β2 and β3-β4, respectively, are those with a significant decrease in intensity (more than expected due to viscosity effects which lead to slow tumbling). Amide groups in β2 and the $\alpha$-helix's N-terminal show a pattern of alternate signal reduction ratios. While A23 backbone amide peak is significantly affected, D22 and A24 are not. Indeed, the backbone amide hydrogens more affected are the most accessible to exchange with solvent[23]. Their intensities are also sensible to minimal pH changes (with small additions of sodium hydroxide), as recently demonstrated by Faustino et al[65]. Analysis of the $^1H^{\beta/\gamma}$-$^{13}CO^{\gamma/\delta}$ (carbonyl or carboxyl) peak intensity of Asx/Glx (Fig. 2.10c) indicates a significant intensity reduction of D22 side chain with the addition of [Ch][Glu] (despite the insignificant $^{13}CO$ chemical shift perturbation, Fig. 2.9c). This is closely linked with the nature of the interaction between $Ch^+$/ $Glu^-$ ions and A23, as the preceding residue is not affected by chemical shift but only by decrease of intensity in its side chain (i.e., local dynamics are affected). For D40

and E42, while carboxyl groups are the significantly perturbed in δ $^{13}$CO, are nearly unaffected in terms of intensity. Analysis of lysine's side chain peaks yielded no significant information.

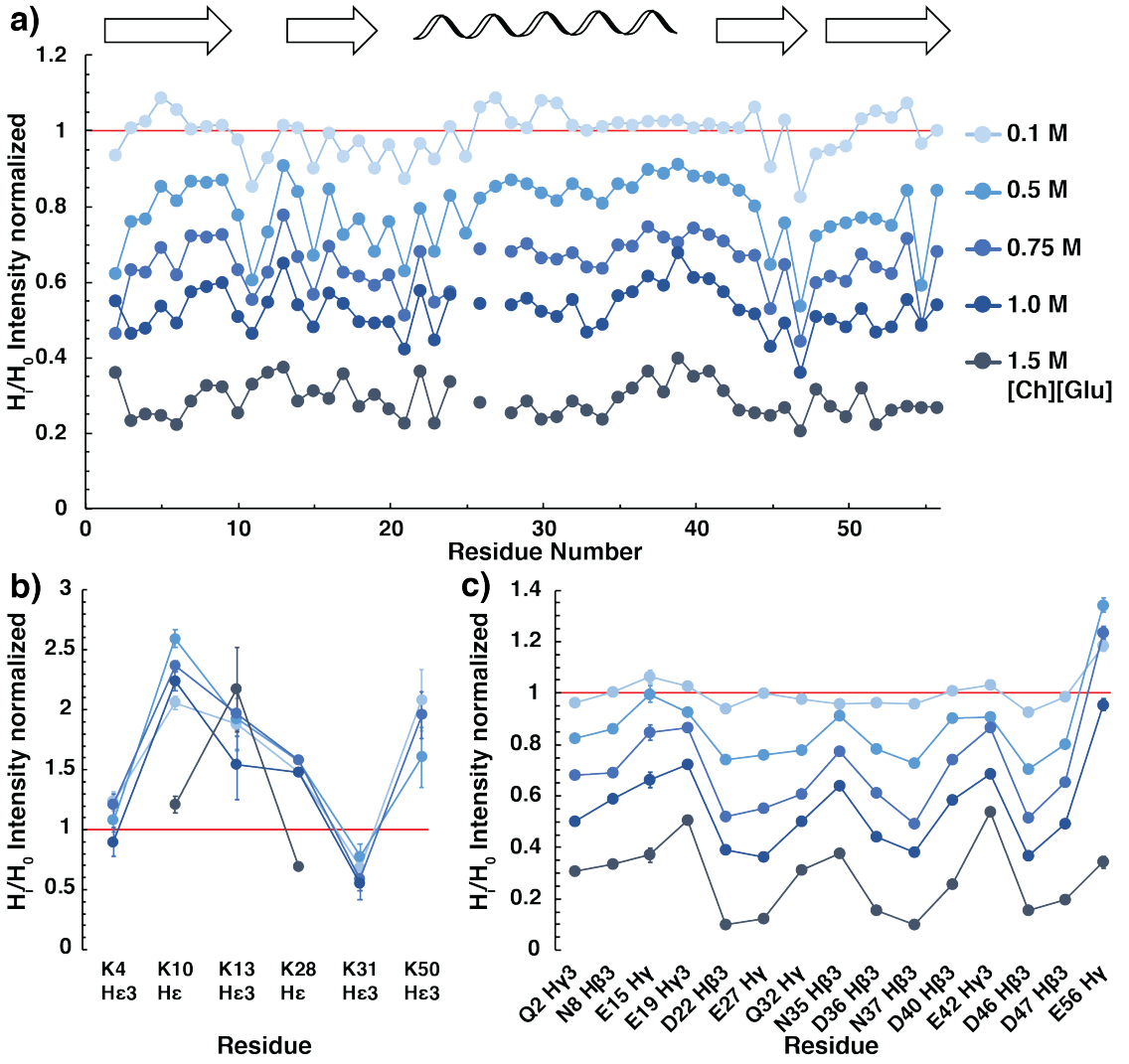



**Figure 2.10. [Ch][Glu]-GB1 interactions concerning peak intensity.**

**a**) Overlay of the height intensity ratio of GB1 in the presence of increasing [Ch][Glu] concentrations (shades of blue). The height intensity ratios were calculated against the peak height intensity in water. T25 and E27 residues are not represented above 1.0 M [Ch][Glu] since they became super-positioned. Above the plot it is depicted the secondary structure of the protein. **b**) Overlay of the height intensity ratio from $NH_3^+$ groups of side chains for [Ch][Glu] titration, based on the $^1H^{\varepsilon}$-$^{15}N^{\zeta}$ H2(C)N spectra. For those with two peaks due to non-equivalent protons, only one of them are shown. **c**) Overlay of the height intensity ratio from carbonyl/carboxyl groups of side chains for [Ch][Glu] titration, based on $^1H^{\beta/\gamma}$-$^{13}CO^{\gamma/\delta}$ H(C)CO spectra. Peak heights were extracted from the spectra. The data were normalized according to $H_i/H_0$ where "0" denotes the reference – water, a value of 1 (red line) or 0 indicates no change or disappearance of the peak (height) intensity, respectively.

To elucidate the species responsible for the shifts of the most affected residues in [Ch][Glu] CSP profile (Fig. 2.8), I repeated the experiments in the presence of [Ch]Cl and Na[Glu] salts (Fig. 2.11). The plots of $\Delta\delta_{comb}$ for each backbone amide versus concentration for each of these conditions can be found in Fig. A1 of Appendix A.

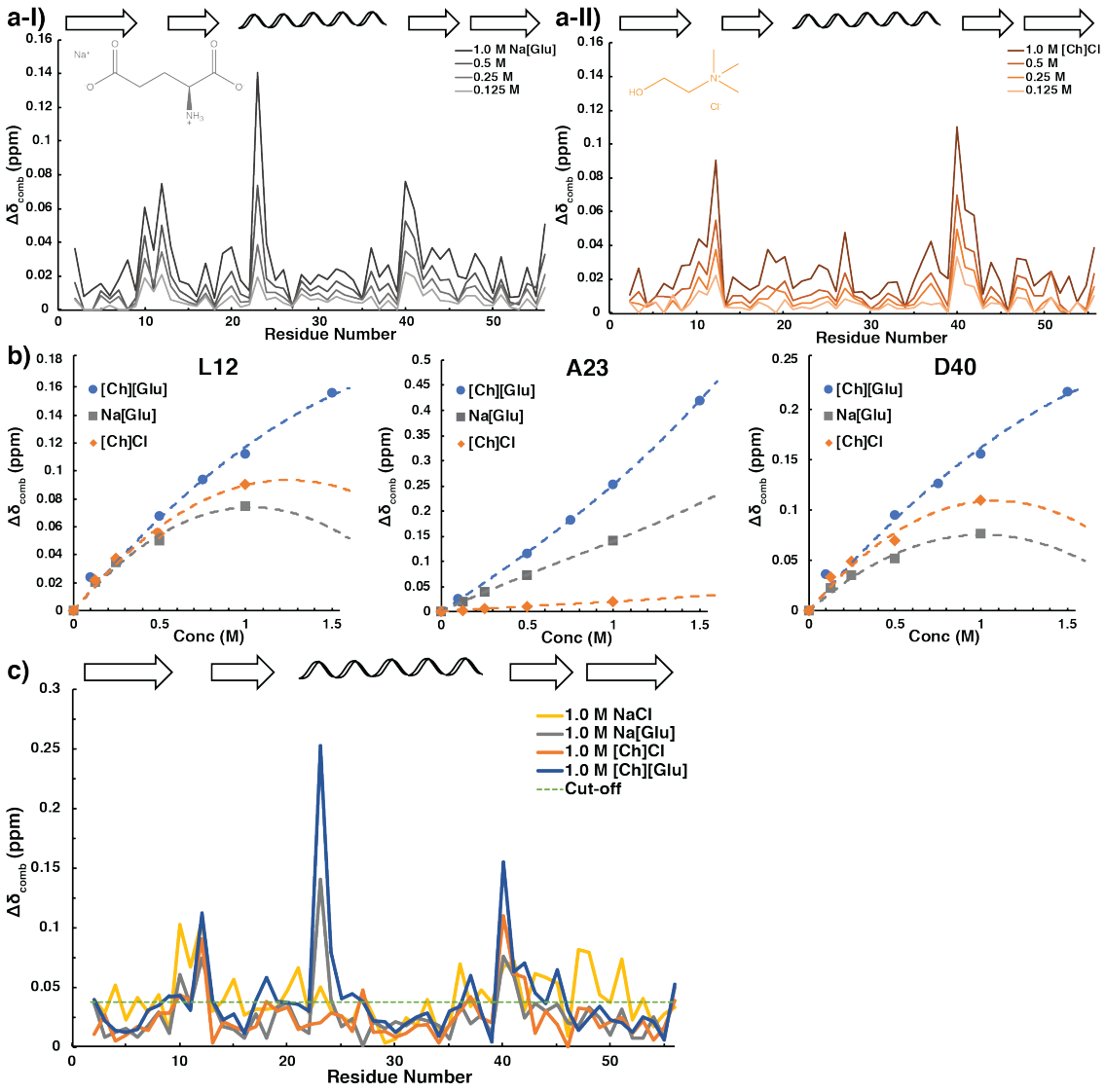



**Figure 2.11. Effect of [Ch][Glu] and related salts on backbone amides of GB1.**

**a)** Overlay of the combined chemical shift of GB1 in the presence of increasing **I)** Na[Glu]) or **II)** [Ch]Cl concentrations. **b)** Plot of $\Delta\delta_{comb}$ as function of [Ch][Glu], Na[Glu] and [Ch]Cl concentration for Leu12, Ala23 and Asp40. The dashed lines are only indicative (polynomial with two variables were found to better fit the experimental data). **c)** Overlay of the combined chemical shift of GB1 in the presence of 1.0 M [Ch][Glu] (blue), Na[Glu] (gray), [Ch]Cl (orange) and NaCl (yellow). The green dashed line represents the threshold for 1.0 M [Ch][Glu]. Above each plot it is depicted the secondary structure of the protein. The combined chemical shifts were calculated against the chemical shifts in water.

I observe that the chemical shift of L12 is perturbed in all conditions independent of the nature of the salt (Fig. 2.11b). Thus, the observed CSP is merely a consequence of the increased ionic strength. For A23 however, I observe that the residue is only affected in the presence of $[Glu]^-$ (i.e., [Ch]Cl and NaCl do not affect this residue), with the strongest effect being observed in the presence of [Ch][Glu] (Fig. 2.11c). The fact that $[Glu]^-$ affects the observed chemical shift of A23 while $[Cl]^-$ does not, indicates that it is not simply an electrostatic effect. Because high concentrations of [Ch][Glu] and Na[Glu] increase the thermal stability of GB1 (Fig. 2.4a and Table 2.1), these observations highlight the importance of stabilising $[Glu]^-$/A23 interactions on the overall GB1 stability. For D40, while all added compounds seem to affect its shift, it is clear that [Ch][Glu] and [Ch]Cl are the ones originating the most significant shifts. As such, despite the evident ionic strength dependence of D40[29], it appears that the observed effects (not only electrostatics but also hydrogen bonding) are a consequence of the presence of the cation $[Ch]^+$. These effects on D40 residue should contribute for the overall destabilisation of GB1 since the increase of [Ch]Cl salt concentration led to a linear decrease on the $T_m$ of GB1 (Fig. 2.4a).

Although the combined chemical shift (Fig. 2.8) in the backbone amide is a good indication on whether a given residue in involved in binding or not, it does not hint at the nature of the interaction[59]. In this sense, amide proton ($H_N$) chemical shifts alone have long been correlated with protein secondary structure and hydrogen bonding[37,66,67] and shown to be sensitive to their environment (e.g., pH, temperature, H-bonds, or ring currents). Overall, in the presence of [Ch][Glu], I observe that the residues in the N-terminal $\alpha$-helix shift downfield which is consistent with a general deshielding of the amide groups, typically attributed to stronger H-bonding, while those in the β-sheet more or less alternate between up and downfield shifts (Fig. 2.12). A closer analysis of the observed variation in the $\delta\ ^1H_N$, specifically at those that are more affected by the presence of the IL, shows that while amide protons of A23 and D40 shift downfield, for L12 the shifts are upfield.

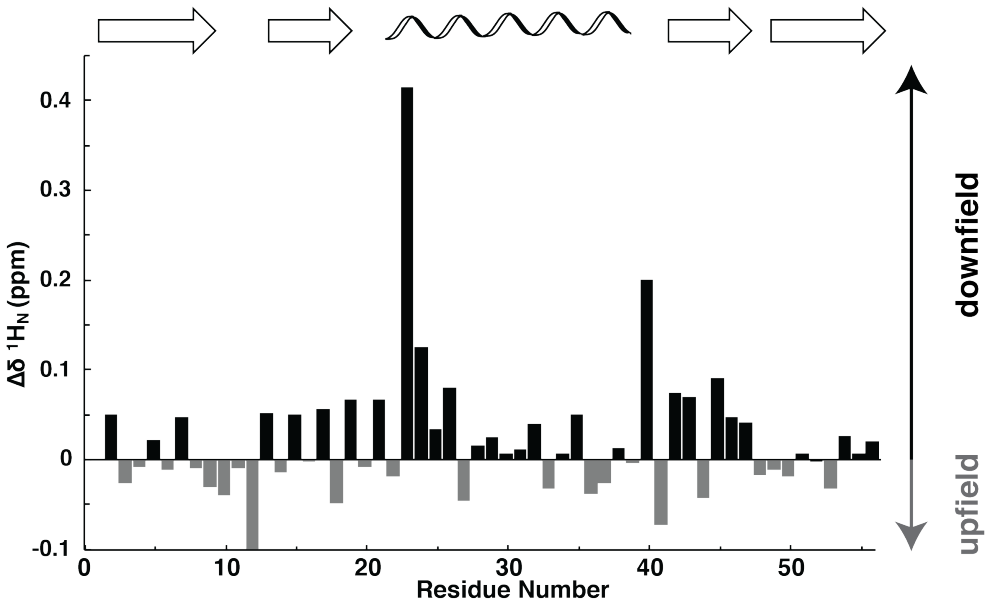



**Figure 2.12. Amide proton chemical shift differences of GB1 with [Ch][Glu].**

Plot of the $^{1}H_N$ chemical shift variations on GB1, in the presence of 1.5 M [Ch][Glu], $\Delta\delta\ ^{1}H_N = \delta\ ^{1}H_N$ (IL) - $\delta\ ^{1}H_N$ (water). Positive and negative values represent downfield and upfield shifts (deshielding and shielding effects), respectively.

Of note are the downfield shifts of the $H_N$ of A23 and D40. These indicate strong a change in the electrostatic environment around these groups and are of particular relevance due to their charged nature (see Fig. 2.13a). Since the amide group of A23 is facing the solvent, the observed downfield shift could indicate that the [Ch][Glu] is interfering in possible contacts from the amide group with the solvent. In fact, when analysing a 3D crystal structure of GB1 (PDB ID: 2QMT[68]), I see that the amide group of A23 is making an H-bond with a solvent molecule and it is also very close to a phosphate ion ($PO_4^{3-}$). So, it is possible that the $Glu^-$ anion and/or ion-pair interferes with this interaction by replacing the existent water molecule and establishing a new and stronger/ short H-bond with the amide group of A23 (Fig. 2.13b).

At the working pH, the negatively charged protein should enhance the pairing of the small choline cation with acidic, negatively charged residues. D40 is the ideal target[69] as it is the only acidic residue in a surface-exposed loop, and it lacks intramolecular side chain hydrogen bonds[19]. Therefore, an electrostatic interaction between its anionic side chain and the $[Ch]^+$ cation (and/ or ion-pair) is plausible with the observed deshielding effects in its amide group. For L12, the presence of positively charged ions (e.g., $[Ch]^+$ and $Na^+$) in this region would withdraw electron density from its amide group, making it more acidic and causing the observed $H_N$ upfield shift.

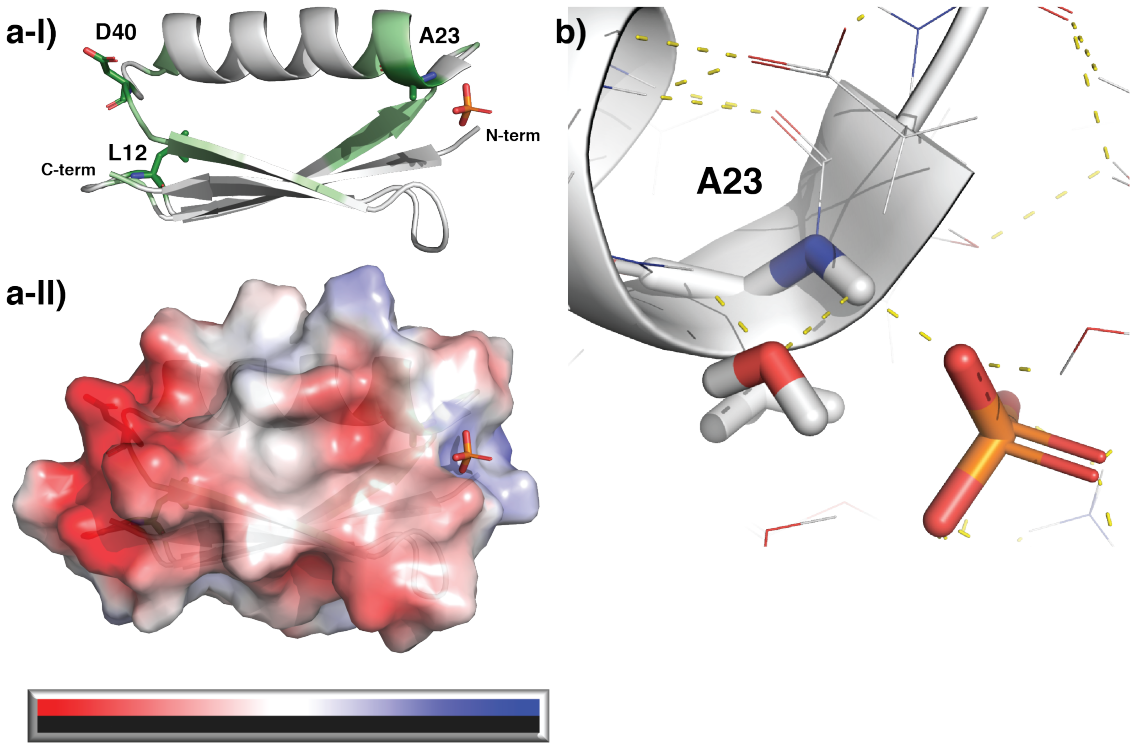



**Figure 2.13. [Ch][Glu] accumulates in the surface charged areas of GB1.**

**a)** Representation of the affected residues (green) onto the 3D X-ray crystal structure of GB1 corresponding to the PDB code: 2QMT[68]. The most affect residues (L12, A23 and D40) are represented as sticks. The phosphate anion ($PO_4^{3-}$) is also displayed. **b)** Electrostatic potential (+/- 3) onto the molecular surface of GB1 calculated using the APBS (Adaptive Poisson–Boltzmann Solver) electrostatics plugin[70] within PyMOL 2.4.2. **c**) H-bonds involving the amide proton of A23 of GB1, including an intermolecular H-bond with the phosphate anion, are shown in the 3D crystal structure.

Temperature dependence

The amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$), which correspond to the slope of a plot of $^1H^N$ upfield shift against increasing temperature, can be used as indicators of intra-protein hydrogen bonding[71,72]. For each residue of GB1 at 1.5 M [Ch][Glu], the plot $^1H^N$ chemical shift versus temperature (283 to 323 K) can be found in the following Fig. 2.14. In general, the hydrogen bond is weakened with temperature increase due to a larger thermal motion, leading to an upfield shift of the amide proton resonance[71].

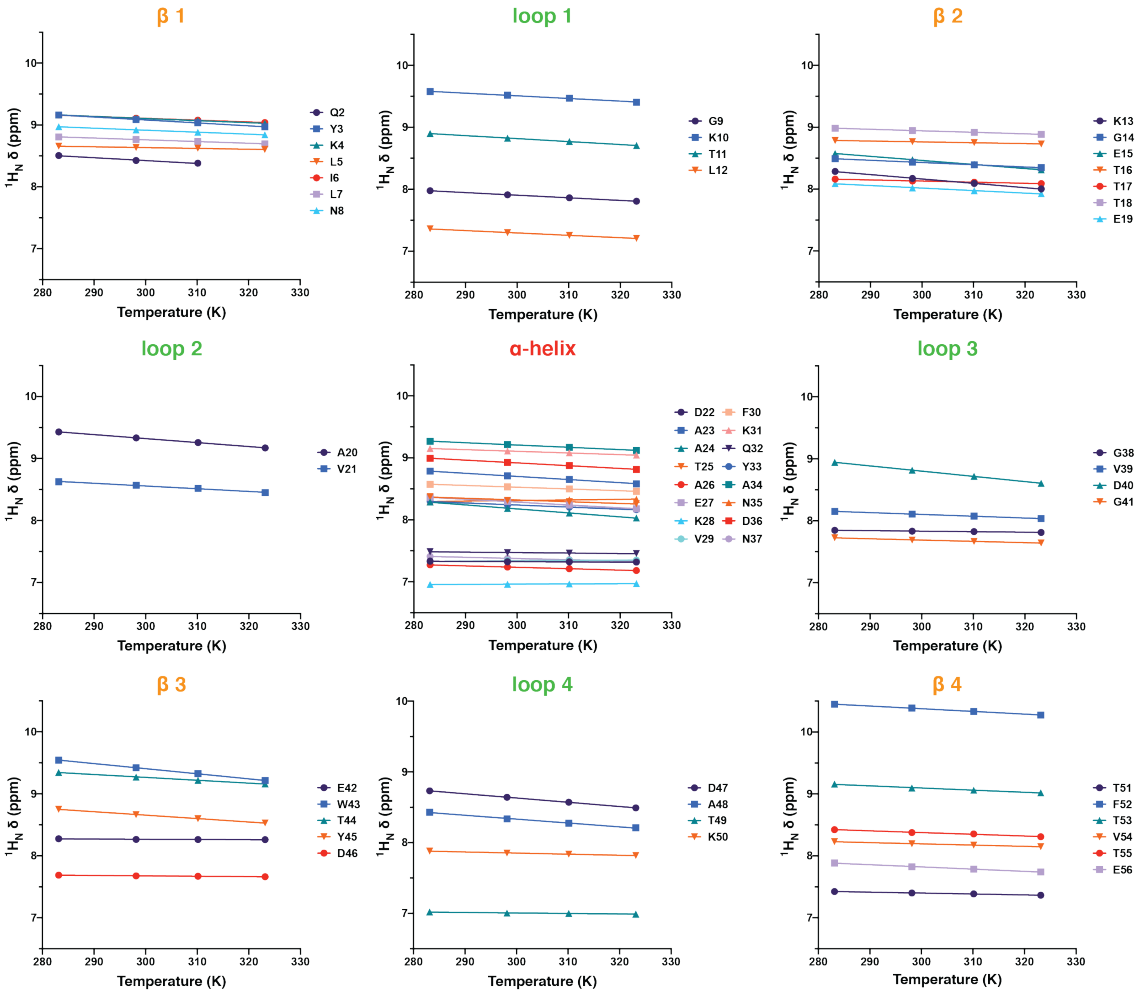



**Figure 2.14. Plot of $^1H_N$ chemical shift versus temperature for each residue of GB1 in the presence of [Ch][Glu].**

A linear least-squares fit (lines) of $^1H_N$ chemical shifts at 4 temperatures (283.2, 298.2, 310.2 and 323.2 K) were determined for each backbone amide at 1.5 M [Ch][Glu]. For each residue, the δ $^1H_N$ was extracted from the peak in the 2D $^1H$-$^{15}N$ HSQC measured at a given temperature. The slope (δ $^1H_N$ versus temperature)

corresponds to the amide proton temperature coefficient ($\Delta\sigma_{HN}/\Delta T$) and its standard deviation represents uncertainty. Each plot corresponds to the group of residues that are part of that secondary structure.

H-bond weakening is usually more pronounced for intermolecular interactions, such as between a protein and the surrounding water molecules. Therefore, while $^1H_N$ temperature coefficients in proteins (such as the well-folded GB1) are mainly determined by local melting/ fractional loss of the structure[73], and not by the strength of hydrogen bonds to the amide, chemical shifts of protons involved in intermolecular H-bonds show stronger temperature dependence (larger coefficient, < -7 ppb/K) than those of protons involved in intramolecular hydrogen bonds (smaller/less negative coefficient > -4.6 ppb/K) (Fig. 2.15).

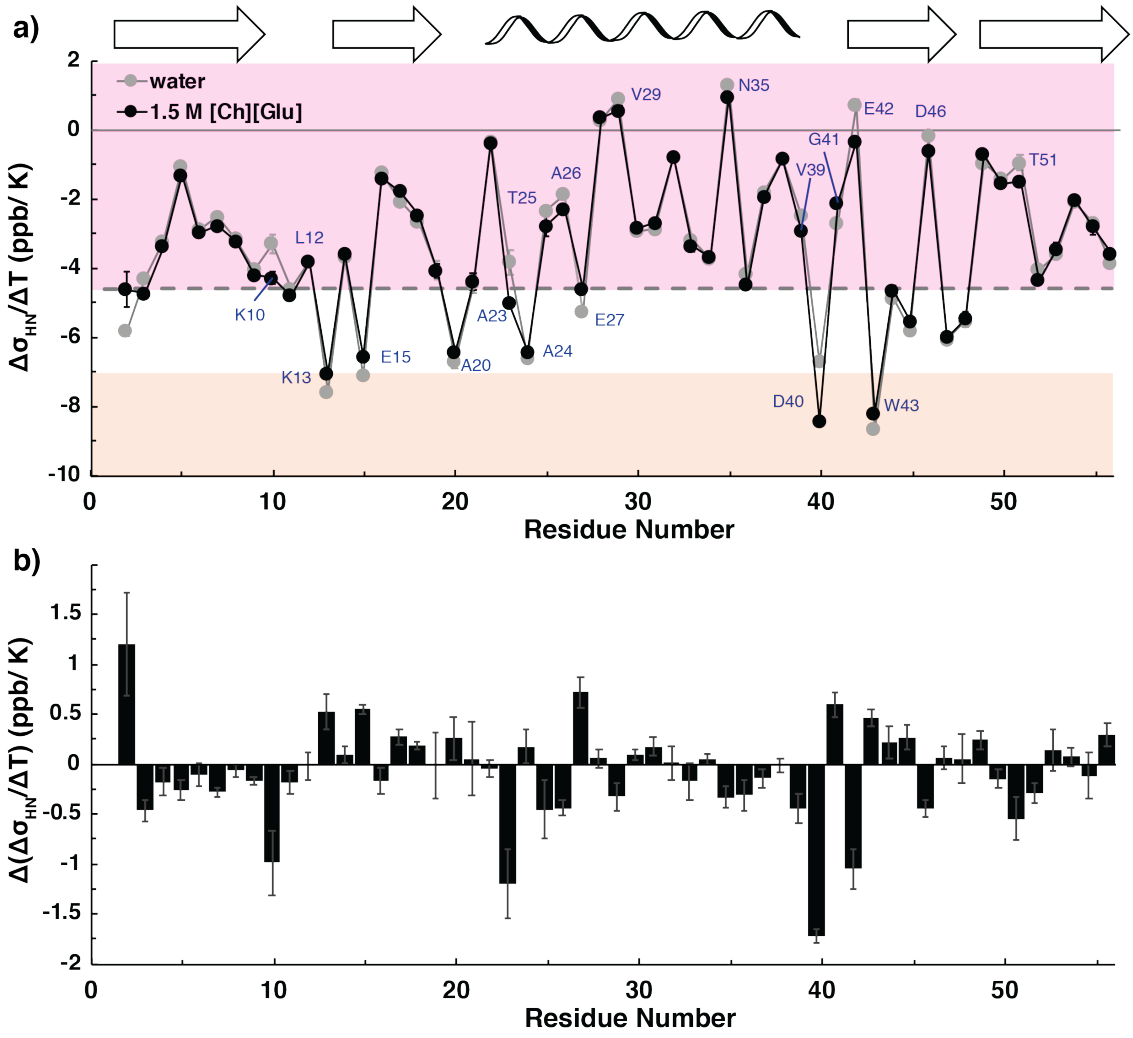



**Figure 2.15. Amide proton temperature coefficients for GB1 in the presence of [Ch][Glu].**

**a)** Amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) for the GB1 in water (grey) and 1.5 M [Ch][Glu] (black). The horizontal dashed line corresponds $\Delta\sigma_{HN}/\Delta T = -4.6$ ppb/K. Bars ending in the pink box have a ≥ 85% probability of participating in an intramolecular hydrogen bond. Bars ending in an orange box have a ≤ 20% probability of participating in an intramolecular hydrogen bond. Values were determined using a linear least-squares fit of $^1H_N$ chemical shifts at 4 temperatures: 283.2, 298.2, 310.2 and 323.2 K. Uncertainties

represent the slope standard deviation. Above the plot it is depicted the secondary structure of the protein. **b)** difference between the amide proton temperature coefficients in [Ch][Glu] and in water [$\Delta(\Delta\sigma_{HN}/\Delta T)$].

In general, I see that the presence of 1.5 M [Ch][Glu] does not significantly alter the overall temperature coefficients of GB1 (average of ~ -3.1 ppb/K in water vs ~ -3.2 ppb/K in IL), meaning that the structure of the protein and H-bonding network remains fairly constant. Regular secondary structure elements correlate well with first-order temperature coefficients greater than −4.6 ppb/K (pink area in Fig. 2.15a), whereas temperature coefficients below −4.6 ppb/K preferentially occur for amino acid residues located in loop regions (Fig. 2.15a). As shown by the relatively high temperature coefficients, both in water and in [Ch][Glu], most of GB1 amide groups are involved in intramolecular hydrogen bonds: there are a total of 45 H-bonds[19], of which 41 are backbone-backbone, 3 side chain-backbone (D22-T25, N37-Y37, D46-A48) and 1 side chain-side chain (Y3-Y45). Residues K13, E15 and W43, although they are part of the regular β-sheet, their amide protons face the solvent, thus justifying the low temperature coefficients. The same is true for A24.

A closer analysis of the data, specifically focusing on the residues whose chemical shift perturbation was the largest (i.e., L12, A23 and D40) shows that:

(1) The temperature coefficient of L12 is not altered by the presence of [Ch][Glu]. This is a good indication that its amide H-bond network remains identical, and it shows that the observed upfield shift (Fig. 2.12) should be caused by indirect interactions with the IL. In the case of a direct IL-protein interaction, it is expected a variation in the local $\Delta\sigma_{HN}/\Delta T$ value. (2) For A23 however, I see that the temperature coefficient in IL is lower than that in water, which indicates a weakening of (short-lived) intermolecular H-bonds (e.g., those with the solvent). This is in good agreement with the observed downfield shift of the HN resonance (Fig. 2.12) and the interpretation made above. Considering these observations and the fact that only the $Glu^-$ anion affects A23 (Fig. 2.11), it is safe to assume that the anion somehow acts at the level of the helix N-terminal, and due to its size, it displaces the existent (and defined) solvent molecules, and possibly establishes a new H-bond with the amide group of A23. (3) For D40, an intermolecular H-bond between its amide group and the sidechain of the C-terminal E56 is observed in the crystal structures of GB1[74,75]. However, looking at the relatively low temperature coefficients obtained both in water and in IL, it seems that the observed H-bond may be, at best, transient, and only detected due to crystal packing[73]. With this in mind, I see that the behaviour presented by this amide proton is very similar to that observed for A23. The difference is that, in this case, the ion responsible for the observed weakening of the H-bond potential is the cation ($Ch^+$) as explained above. (4) Interaction of $[Ch]^+$ with the sidechain of E56 could also account for the relatively accentuated decrease in the temperature coefficient observed for K10. In the crystal structure, the amide group of K10 makes an H-bond with the carbonyl group of E56[29,73] (both residues are well represented in the observed temperature coefficients, Fig. 2.15). Due to the presence of $Ch^+$ cation, this H-bond may be broken, or at least weakened, in agreement with the decrease in the K10 temperature coefficient.

In summary, I have observed/concluded that:

(1) In terms of thermal unfolding variation ($\Delta T_m$), while [Ch][Glu] and Na[Glu] affect GB1 stability in a similar way (the protein is stabilised above 1.5 with a minimum of destabilisation is found around 0.75 M, [Ch]Cl destabilises GB1. (2) Backbone amide's CSP of the well-folded state of GB1 along the titration indicate that [Ch][Glu] interacts weakly (CSP increases linearly with concentration) but preferably with flexible, charged and surface-exposed regions (L12, A23 and D40). While the $\delta\ ^1H_N$ for amide protons of A23 and D40 are shifted downfield (i.e., deshielding effect due to a strong electrostatic change), for Leu12 it is shifted upfield. (3) Comparison of CSP profile for IL and related salts show that the $Glu^-$ anion is responsible for the perturbation in the environment of A23 amide. On the other hand, the $Ch^+$ cation affects the D40 amide group through electrostatic interaction with its carboxylic side chain, as demonstrated by the variation of $^{13}C$ chemical shifts of carbonyl/carboxyl end groups. L12 amide proton is affected by the presence of any positively charged ion. (4) The minor changes observed for $^1H_N$ temperature coefficients with the addition of 1.5 M [Ch][Glu] are a good indicative that the structure of GB1 and its H-bonding network remain fairly constant. A small decrease in the temperature coefficient was found for K10, A23 and D40 residues suggesting a weakening of the intermolecular H-bonds involving these residues. A preferential accumulation of $Ch^+$ or $Glu^-$ ions in these regions at the protein surface could contribute for the weakening of those H-bonds.

As noted in the introductory section of this chapter, previous studies into the conformational dynamics of GB1 (un)folding[25] have reported two melting hot spots (associated with thermal unfolding) formed around T11, L12 and K13 and the N-terminus of the helix at A24 and T25. These melting hot spots are formed prior to GB1's complete unfolding and probably reflect a wide distribution of conformational substates within the native structural ensemble. Taken together the results obtained hint at a possible role of these residues in the modulation of GB1 folding and stability in the presence of [Ch][Glu], in particular the importance of the N-terminal helix hot spot for the increase of GB1 stability. The possible mechanism for [Ch][Glu] effects on GB1 stability, in terms of protein structure, solvent accessibility and dynamics, will be analysed and discussed in the following sections.

[Ch][Glu] and [Bmim][dca]: divergent effects on GB1

To better understand the effects of aqueous-IL on protein structure and stability, I studied the backbone structural perturbations by $^1H$-$^{15}N$ HSQC NMR with 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca]). The combination of imidazolium-based cation with larger and longer hydrophobic alkyl chain and a weakly hydrated chaotropic $[dca]^-$ anion as noted in the Hofmeister series as strong destabiliser results in a strong denaturant IL[3].

In [Bmim][dca] aqueous solutions, the most affected residues of GB1 (when compared to water) are nonpolar and charged polar: Y3, L5 and L7 in β1-strand; L12, K13, G14, T16 to A20 in β2; T25, A26, E27, V29, F30, K31, Q32, N35 in the $\alpha$-helix; G38 in loop3, E42 and T44 in β3 and T53, V54 in β4. The CSP is maximum for G14, T18, A20, V29, Q32, and T44, therefore $\alpha$-helix and β2 strand are the most affected segments (Fig. 2.16).

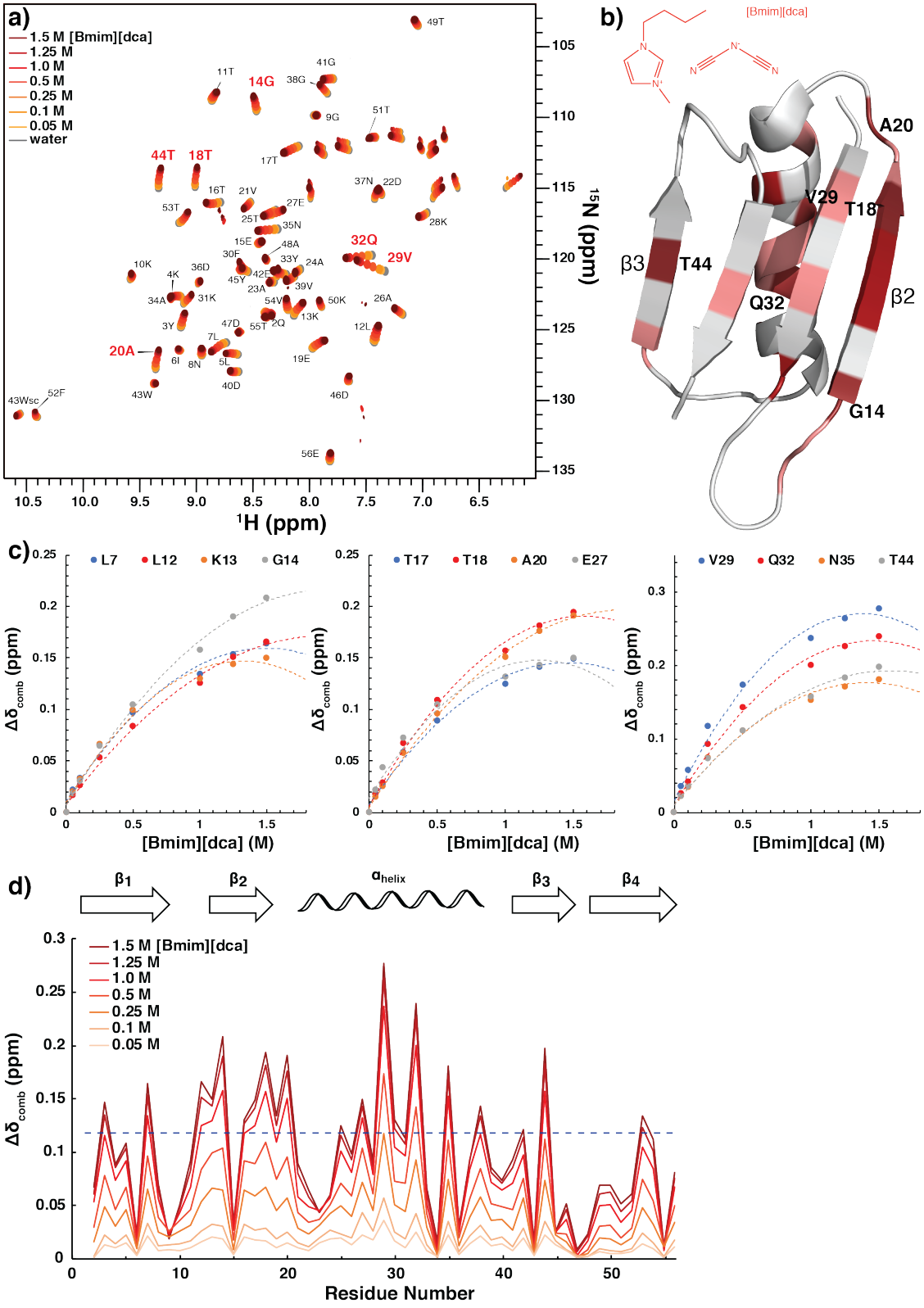



**Figure 2.16. [Bmim][dca]-GB1 interactions.**

**a**) Overlay of 2D $^1$H-$^{15}$N HSQC spectra of GB1 with increasing amounts of [Bmim][dca]. The labels show the assignments for all residues. **b**) (top) [1-butyl-3-methylimidazolium][dicyanamide] structure, (bottom) the affected residues are mapped onto the respective 3D structure (PDB: 2JSV[21]). **c)** Plot of $\Delta\delta_{comb}$ as function of

[Bmim][dca] concentration for the most affected residues. The dashed lines are only indicative (polynomial with two variables were found to better fit the experimental data). **d**) Overlay of the combined chemical shift of GB1 in the presence of increasing [Bmim][dca] concentrations. The threshold that distinguishes the affected from non-affected residues is represented by the blue dashed line for 1.5 M [Bmim][dca]. Above the plot it is depicted the secondary structure of the protein. The combined chemical shifts were calculated against the chemical shifts in water. Data were acquired for 0.25 mM $^{15}$N GB1 with increasing amounts of [Bmim][dca] (0, 0.05 M, 0.1 M, 0.25 M, 0.5 M, 1.0 M, 1.25 M, 1.5 M, from gray to shades of red - light red towards dark red) in 90% $H_2O$/ 10% $D_2O$, 0.1% $NaN_3$, pH 7.1 ± 0.1 with 50 μM DSS as internal reference at 298.2 K, 600.13 MHz.

The CSP values display a typical binding isotherm behaviour with saturation around ~ 0.5 M (Fig. 2.16c), allowing the calculation of apparent equilibrium dissociation constant ($K_D$)[59]. The obtained $K_D$ values are in the range of 2.8 M (L7) < 0.7 M (V29) < 0.4 M (T16) and indicate that these perturbations are caused by weak binding interactions as previously reported for other proteins, such as the negatively charged Im7 small protein[6].

It has been shown before that [Bmim][dca] IL strongly denatures proteins due to direct hydrophobic interactions[3,4,6,56]. Furthermore, analysis of GB1's 3D structure (Fig. 2.16b) shows that the affected residues are located essential in its hydrophobic core (L5, L7, F52 and V54 residues are completely buried[19,76]). The affected residues are also important to hold the intramolecular hydrogen bonding: Q32-V29-T25 in the helix, Y3-T18 in β1-β2 strand, and T44-T53 in β3-β4 strand. Because V29 side chain is constituted by a isopropyl group and its backbone amide is highly exposed to the solvent, a maximum CSP from a direct hydrophobic interaction is expected. Although the anion dictates the overall stability, most of these interactions are mediated by the cation $[Bmim]^+$ and/or ion-pair. This observation is in agreement with Warner et al[77]. They investigated the [Bmim][Br] effect on GB1 by high-resolution magic-angle spinning (HR-MAS) and obtained similar results as those reported here, with the strongest interaction in the $\alpha$-helical region. At 3.5 M concentration the authors[77] observed the coexistence of the folded and unfolded states of GB1. To determine if [Bmim][dca] has a similar effect, $^1$H-$^{15}$N HSQC spectra were collected above 1.5 M [Bmim][dca] showing a complete GB1 unfolding attained at 3.5 M of concentration (in Fig. 2.17 one can observe the peaks of the unfolded protein arising above 1.5 M, in slow exchange on the NMR time scale). This indicates, at least for GB1, that the denaturing power of [Bmim][dca] is greater than [Bmim][Br] (in agreement with the Hofmeister's series for the anions) and common denaturing agents (e.g., urea and guanidinium chloride [GdmCl])[78].

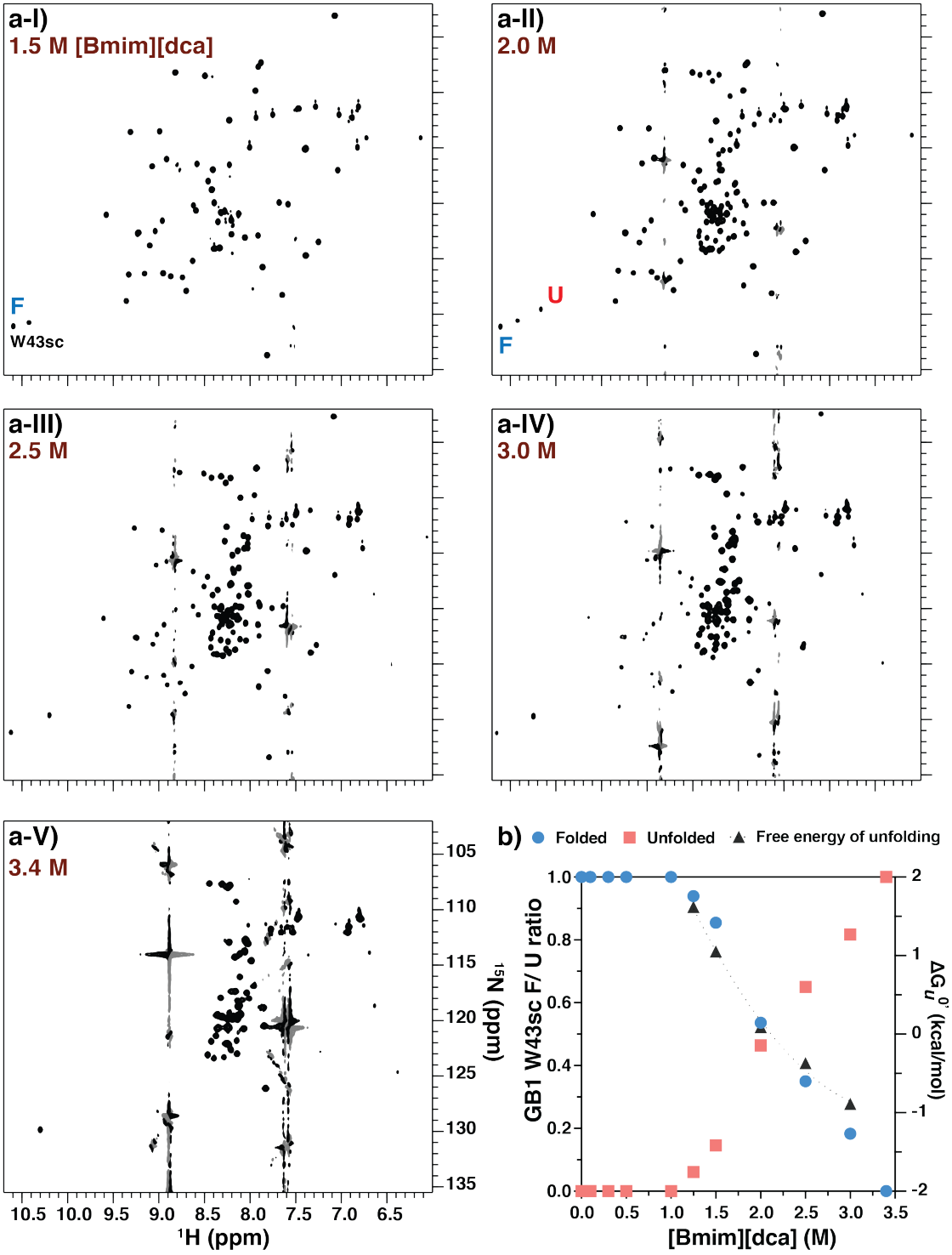



**Figure 2.17. Denaturation of GB1 with [Bmim][dca].**

**a**) 2D $^{1}$H-$^{15}$N spectra of GB1 in the presence of different [Bmim][dca] concentration (1.5, 2.0, 2.5, 3.0, and 3.4 M). The label F and U indicates the folded or unfolded state of W43 indole side chain peak. **b**) Populations and Gibbs free energy of unfolding ($\Delta G_u^0$) were calculated as function of [Bmim][dca] concentration. Based on each $^{1}$H-$^{15}$N HSQC spectrum, peak volume intensities of the exposed indole side chain of W43 were used to calculate the populations of folded state ($p_f$, blue) and unfolded ensemble ($p_u$, red). For each titration point, the Gibbs free energy of unfolding ($\Delta G_u^{0'}$, black) was calculated assuming a two-state unfolding model.

In general, the two distinct patterns observed for [Ch][Glu] and [Bmim][dca] ILs, Fig. 2.18, seem to be directly linked to a stabilisation and destabilisation mechanism, respectively. First, [Ch][Glu] interacts weakly with solvent accessible and charged residues. Whilst the hydroxylated cation accumulates in the negatively charged areas of protein surface and stimulates hydrogen bonding interactions (e.g., $[Ch]^+$ interaction with D40 carboxyl side chain), the strongly hydrated anion accumulates around the N-terminus of the $\alpha$-helix (D22-A24), an recognised (un)folding hotspot[25] which could be responsible by the observed improvement of GB1 stabilisation. The hydrophobic core of GB1 is not perturbed at all up to concentrations of 1.5 M opening the possibility of an entropic stabilisation mechanism, mainly derived from excluded volume effects[79]. Second, the strong destabiliser [Bmim][dca] interacts with nonpolar and charged polar residues that compose the hydrophobic cluster, strongly affecting the $\alpha$-helical region and the β2 strand and promoting the overall unfolding of GB1. The hydrophobic imidazolium-based cation weakly interacts with these hydrophobic segments and, in combination with the weakly hydrated $[dca]^-$ anion (high ion pair strength), the protein unfolds by exposing the hydrophobic surface to the water.

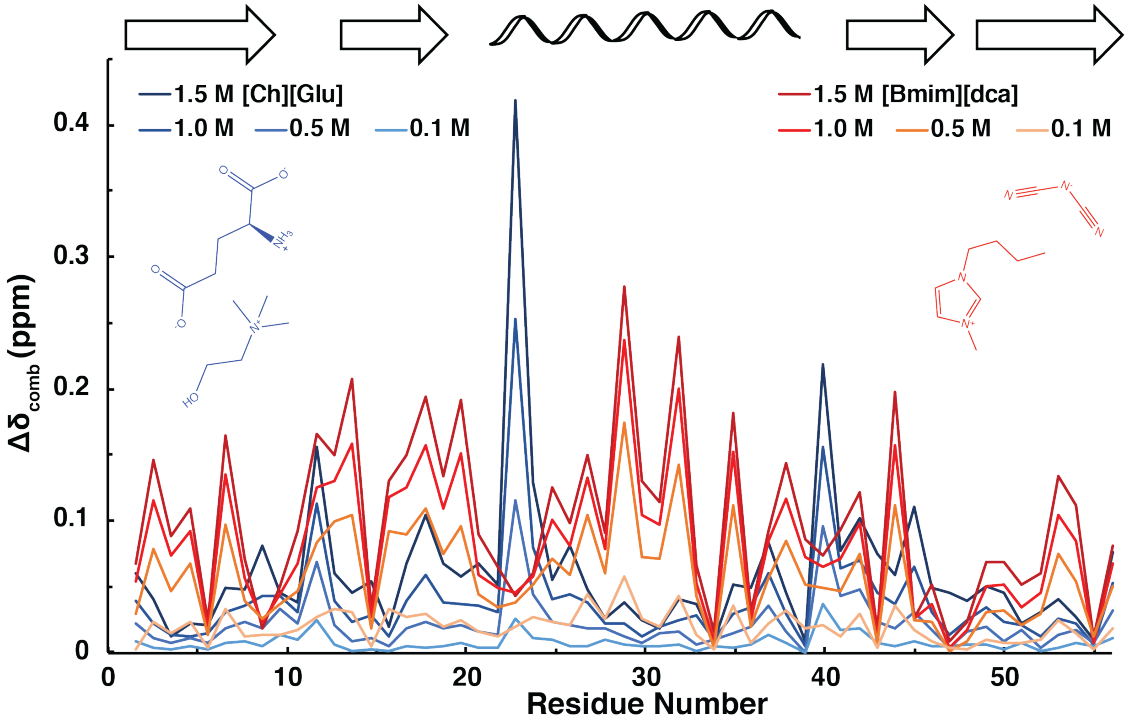



**Figure 2.18. [Ch][Glu] and [Bmim][dca] prompts distinct perturbations on GB1.**

Overlay of the combined chemical shift of GB1 in the presence of increasing [Ch][Glu] (blue) or [Bmim][dca] concentrations (red).

### Structure perturbations

Although the presence of [Ch][Glu] IL is not expected to induce a significant structural perturbation on GB1 (based on the rather small CSP, Fig. 2.8), [Bmim][dca] leads to a complete denaturation of the protein, forcing the transition from a well-folded state towards an unfolded ensemble at 3.5 M (Fig. 2.17 and 2.19). To study the effects of [Bmim][dca] at the secondary structure level I used conventional triple-resonance NMR experiments (3D HNCO, HN(CA)CO, HN(CO)CACB and HNCACB) and assigned the backbone and aliphatic $H^N$, $N^H$, C$\alpha$, C$\beta$ and CO chemical shifts for GB1 in water and in the presence of 1.5 M [Ch][Glu] ($F_{[Ch][Glu]}$) or 3.5 M [Bmim][dca] (completely denatured, $U_{[Bmim][dca]}$) at 298.2 K and pH = 7. See Material and Methods for details and Table A2 of Appendix A for the backbone chemical shifts of GB1 in these conditions. While in water and 1.5 M [Ch][Glu] the observed chemical shifts are very similar and indicative of a well-folded protein (Fig. 2.7), in 3.5 M [Bmim][dca] (Fig. 2.19) the narrow amide $^1$H chemical shifts dispersion in the $^1$H–$^{15}$N HSQC is indicative of an unfolded/ unstructured protein, i.e., an ensemble of rapidly exchanging polypeptide conformations. Yet even if this ensemble is best described by a random coil model, the presence of residual structure should be considered[80].

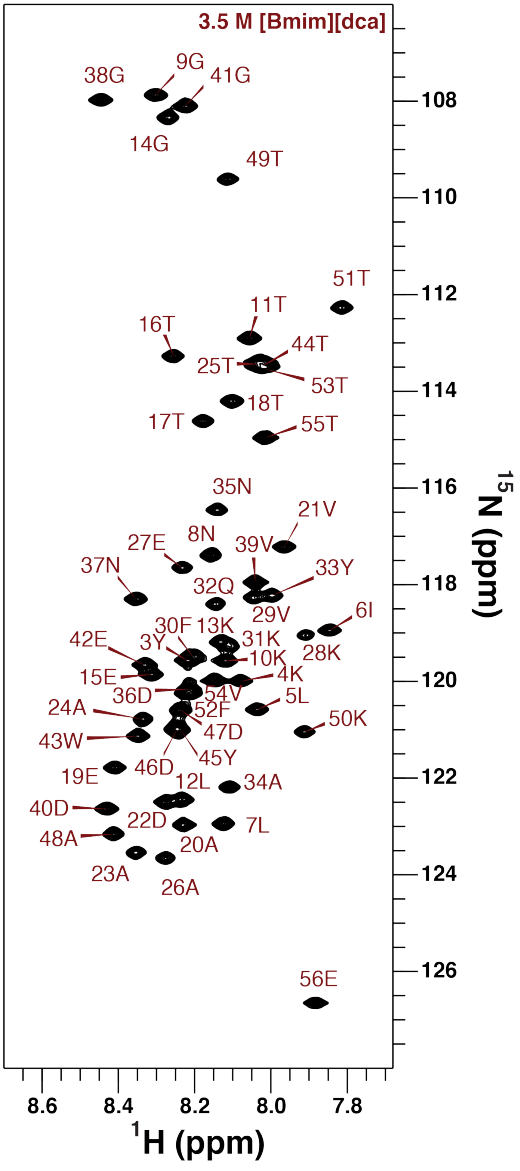



**Figure 2.19. Amide NH backbone assignment of unfolded GB1 ensemble in [Bmim][dca].**

$^1$H-$^{15}$N HSQC Spectrum of 1.2 mM $^{15}$N-$^{13}$C GB1 at 3.5 M [Bmim][dca], 90% $H_2O$/10% $D_2O$, 0.1% $NaN_3$ and 100 $\mu$M DSS – fully unfolded (stabilised) state. Red label indicates amide NH peaks assignments.

The carbon nuclei ($^{13}$C$\alpha$, $^{13}$C$\beta$, and $^{13}$CO) are highly sensitive to backbone $\phi/\psi$ torsion angles and the deviations of these $^{13}$C chemical shifts from their random coil values ($\Delta\delta = \delta_{observed} - \delta_{coil}$) are commonly used as a measure of secondary structure[81]. As a general rule, while $^{13}$C$\alpha$ shifts experience a downfield shift of ~2.5 ppm in helices and upfield shift of ~2.0 ppm in β-sheets, $^{13}$C$\beta$ shifts assume a near random coil values in helices and are shifted downfield ~2.5 ppm in β-sheets. $^{13}$CO shifts are affected in a similar way as C$\alpha$, however their use is less consistent as these are more sensitive to local sequence and may be subject to misreferencing. Using the SSP program[82], the C$\alpha$ and C$\beta$ chemical shifts were combined into a single residue-specific secondary structure propensity (SSP) score, in which, the contribution of different

chemical shifts are weighted by their sensitivity to α- and β-structure. It has been shown that differences in the SSP scores are much clearer than in the $\Delta\delta C^{\alpha}$- $\Delta\delta C^{\beta}$ plots[82]. Figure 2.20 shows the predicted SSP scores for the $F_{[water]}$, $F_{[Ch][Glu]}$ and $U_{[Bmim][dca]}$ states of GB1. Similar SSP values are observed for GB1 in water and 1.5 M [Ch][Glu] suggesting an identical secondary structure (Fig. 2.20a). Only small changes on the SSP scores are observed for the α-helix and β-strands (SSP decreases for β1 and β3 and it increases for β4), which may be considered insignificant. However, for the $U_{[Bmim][dca]}$ state, the SSP scores indicate the collapse of the β-structure, as well as a significant reduction in the propensity for the α-helical motif (Fig. 2.20b). Nevertheless, while this reduction is significant, I still observe SSP scores above 0.4 for some residues, in particular for the segment between T25 and F30, indicative of a residual α-helical structure. This observation contrasts with those in urea (7.4 M urea, 298.2 K, pH = 2), where it was observed that the unfolded structure is closer to a random coil, with no evidence for the existence of residual native or non-native ordered structure[78].

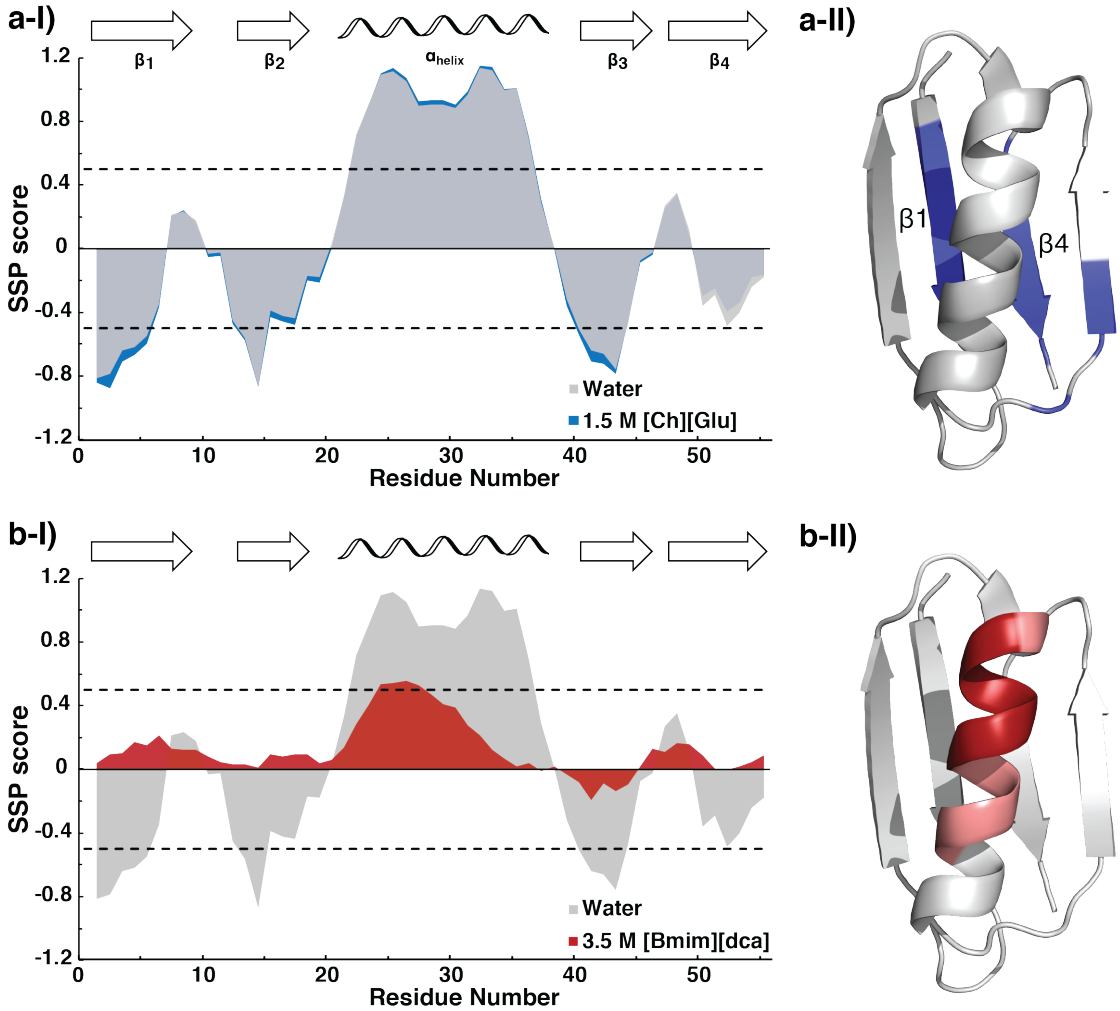



**Figure 2.20. Secondary structure propensity (SSP) of GB1.**

Comparison of the SSP scores for GB1 for **a**) [Ch][Glu] and **c)** [Bmim][dca] conditions. **b)** 3D structure of GB1 (PDB: 2JSV[21]) highlighting the affected regions (in blue, $\Delta SSP = SSP_{[Ch][Glu]}-SSP_{[water]} > 0.05$ or $< -0.05$). **d)** The SSP scores obtained from [Bmim][dca]-unfolded state are highlighted in the 3D structure of GB1 (in

shades of red, $SSP_{[Bmim][dca]} > 0.2$). An unfolded GB1 structural ensemble is not available. The SSP scores were calculated on the basis of Cα and Cβ chemical shifts, and a score at a given residue of 1 or -1 reflects fully formed α- or β-structure, respectively, while a score of 0.5 or -0.5 (horizontal dashed lines) indicates that 50% of the conformers in the state ensemble are in α- or β-structure, respectively, at that position[82]. All the data were acquired at 298.2 K, pH = 7 in water (grey) and in the presence of 1.5 M [Ch][Glu] (blue) or 3.5 M [Bmim][dca] (red). Above each plot it is depicted the secondary structure of the protein according to the 3D structures.

Our data for unfolded GB1 indicates that an unfolded ensemble in denaturing IL and in denaturing agents, such as urea or [GdmCl] salt, may be significantly different. Considering that i) there is no such thing as an unfolded structure, but instead an ensemble of unfolded structures, ii) the different unfolded structures may be involved in physiological roles (e.g., drkN SH3 as explored in the next chapter) and iii) their study is almost always hindered by their transient nature, the data presented here, both in terms of methodology and the application of ILs, shows great potential to study these elusive protein conformations and may have a significant impact on our understanding of the molecular determinants underlying protein folding and unfolding.

**Solvent accessibility/ stability**

The dual behaviour observed for GB1 overall stability when increasing [Ch][Glu] IL concentration (see Fig. 2.4 in the section of ILs and salts effects on protein stability, with a shallow minima found at ~ 0.75 M), is not explained by the linearity of amide's CSP observed with increasing IL concentration ([Ch][Glu]-GB1 interactions section, Fig. 2.8), neither by the amide proton temperature coefficients (temperature dependence section, Fig. 2.15), or the SSP evaluation of GB1 at 1.5 M [Ch][Glu] (structure perturbations, Fig. 2.20). As discussed in the introduction about enthalpic-entropic fingerprint of cosolutes (ILs – protein interactions section, Fig. 1.5), [Ch][Glu] as other hydrophilic salts and osmolytes[10] should populates states where both the enthalpic and entropic contributions are negative. If the entropic stabilisation is dominant compared to the enthalpic destabilisation (negative $\Delta\Delta H_u$ and $T\Delta\Delta S_u$ but $\Delta\Delta H_u < T\Delta\Delta S_u$), the protein is stabilised ($\Delta\Delta G_u = \Delta\Delta H_u - T\Delta\Delta S_u > 0$). This mechanism is consistent with the absence of direct chemical interactions between GB1 and [Ch][Glu] and with the fact that the N-*terminus* of the α-helix is affected but without a significant structural change. However, note that the more destabilising enthalpic contribution, the larger the protein destabilisation ($\Delta\Delta H_u > T\Delta\Delta S_u$ and $\Delta\Delta G_u < 0$), as the one found for [Ch][Glu] concentrations around 0.75 M.

The protein stability behaviour might be related with a modification of the water structure around the protein[83] undetected with CSP or SSP experiments due to a different timescale. To tackle this, I measured equilibrium protein stability by acquiring NMR-detected amide $H^N/D^N$ exchange experiments[84]. Experiments were performed for uniformly labelled $^{15}N$ GB1 in the absence and presence of 0.1, 0.75, 1.0, and 1.5 M of [Ch][Glu] at two temperatures, 298.2 K and 310.2 K. Amide $H^N/D^N$ exchange, and further experiments were not performed for [Bmim][dca] IL since the observed direct chemical interactions gave a reasonable explanation for the GB1 denaturation. In contrast to [Ch][Glu], but similar to other

hydrophobic cosolutes[10] that interact unfavourably with the protein backbone, [Bmim][dca] destabilises GB1 entropically and the enthalpic contribution is even stabilising.

Briefly, the observed exchange rate of backbone amide protons for deuterons, $k_{obs}$, is quantified by measuring the decrease in signal from individual amide protons in a series of $^{1}H$-$^{15}N$ HSQC spectra over time (see Material and Methods). GB1, as most globular proteins, undergoes EX2 kinetics[23], as reviewed in the last section of the introduction – Biomolecular NMR spectroscopy: methods for timescale dynamics above seconds to hours / days). Therefore, $k_{ex}{}^{obs}$ values can be converted to the Gibbs free energy required to open the protein and expose individual backbone amide protons to solvent, $\Delta G_{HX}^{0'}$ [85,86]:

$$\Delta G_{HX}^{0'} = -RTln\left(\frac{k_{ex}{}^{obs}}{k_{int}}\right), \qquad [2.1]$$

where $R$ is the gas constant, $T$ is the absolute temperature, and $k_{int}$ is the intrinsic rate in an unstructured peptide[87].

Profiles for representative residues (e.g., T16, T18, A20, K28, A34, D36, V39, K50, E56), accurately followed for all conditions ($D_2O$, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu]) at both temperatures (298.2 K and 310.2 K), are shown in Fig. 2.21. Values of $k_{ex}{}^{obs}$ are tabulated in Table A3 of Appendix A. In the 10 studied conditions, there are merely 5 cross-peaks from backbone amide (Q2, T11, E15, V21, D47) that appear to exchange too fast to be accurately quantified. For the other residues that are neighbour of these 5 (K10, L12, K13, T17, A23, A24, D40, G41, W43, Y45, A48), $k_{ex}{}^{obs}$ values were only obtained under dilute conditions ($\leq$ 0.1 M) and at the lowest temperature. Under these conditions, as a consequence of the low temperature, a series of backbone amide protons exchange too slowly or do not appear to exchange at all (e.g., K4 to I6; F30, K31; T51 to V54). Their quantification is enabled by an increase in temperature.

The plots of the $\Delta G_{HX}$ along the protein sequence, as determined at a residue level using Eq. 2.1 for water and different aqueous-[Ch][Glu] solutions, are shown in Fig. 2.22 (the values and their uncertainties can be found in Table A4 of Appendix A). See further details in the section of Material and Methods: free energies of exchange. Note that the residues more protected from exchange show a higher value of $\Delta G_{HX}$[89]. As expected, the exchange data correlates well with the secondary structure of GB1, i.e., the rapidly exchanging residues are more likely to have greater solvent exposure as they are found in loops, the outer strands of the four-stranded sheet (β2 and β3), and the ends of the helix[19,23,90]. In a similar way, the exchange also reflects the hydrogen-bond network (see Fig. 2.23 in the next landscape page), since H-bond breakage must occur for exchange to take place, the residues involved in H-bond present a higher $\Delta G_{HX}$. The data acquired at 310.2 K (circles, Fig. 2.22) was found to be more reliable in terms of fitting analysis (Fig. 2.21, and Table A3, A4), in particular for the secondary structure elements as the solvent exchange is slower here (i.e., H-bonded NHs). Yet, the exchange data from 298.2 K measurements (squares, Fig. 2.22) supports the latter results and adds information for those residues located in more solvent-exposed areas. Furthermore, the similarity of the patterns of H/D exchange along the primary structure for GB1 in water and in the presence of IL support the conclusion that the tertiary structure is unchanged.

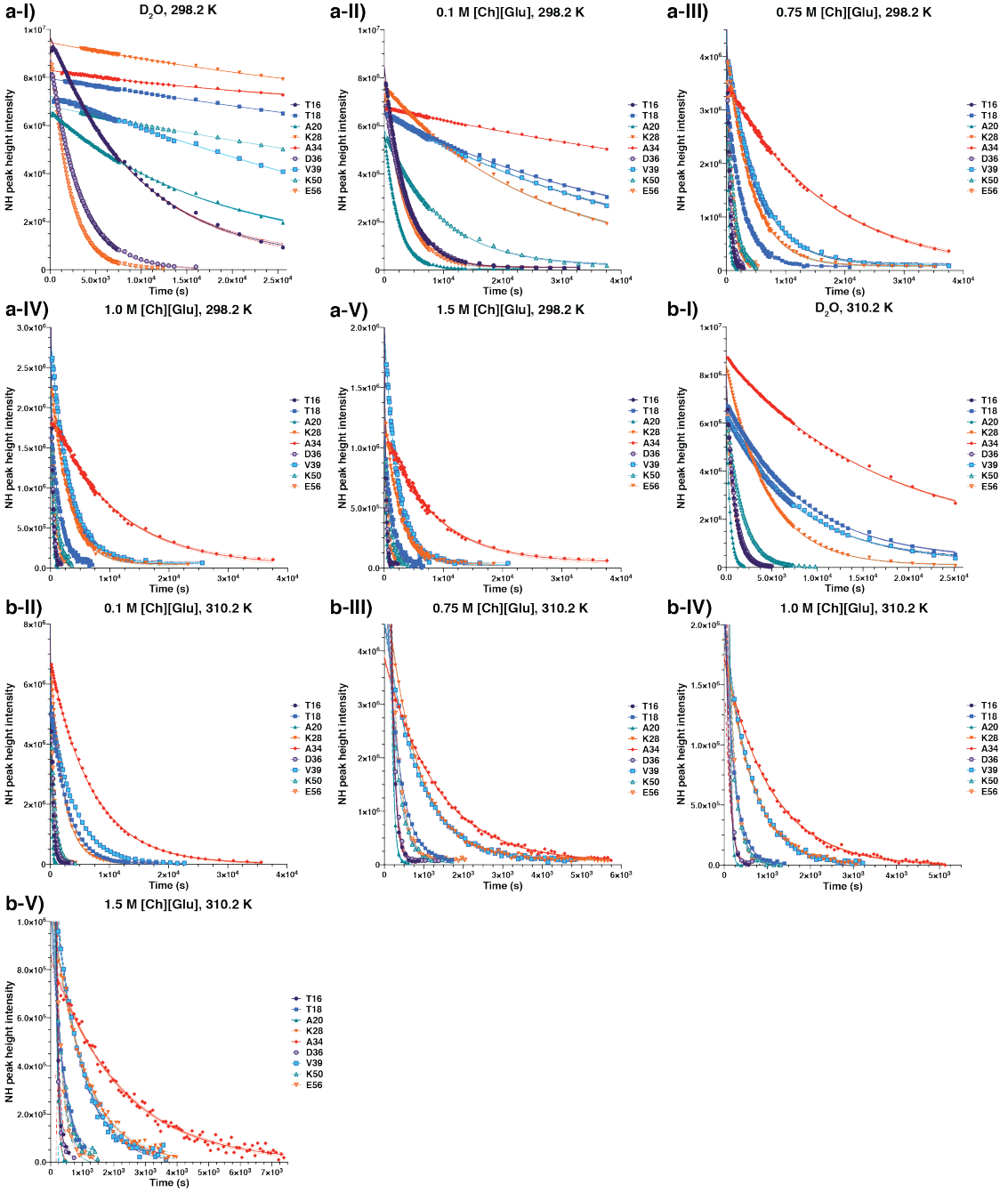



**Figure 2.21. Backbone amide $H^N/D^N$ decay profiles from representative residues of GB1.**

The decay of different cross-peaks with corresponding best fits for T16, T18, A20, K28, A34, D36, V39, K50, E56 residues in **a**) 298.2 K and **b**) 310.2 K for **I-V**), absence of IL (only $D_2O$), 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentration, respectively. For each amide proton, cross-peak height intensity was analysed and extracted using PINT[64,88], and normalized over the number of scans of each $^1H$-$^{15}N$ HSQC spectrum. In general, the NH peak height intensity alongside exchange time follow a single exponential function. The exchange time was defined as the period from protein dissolution in $D_2O$ solution (or perdeuterated [Ch][Glu] solution) to the midpoint of the estimate taken to acquire each spectrum.

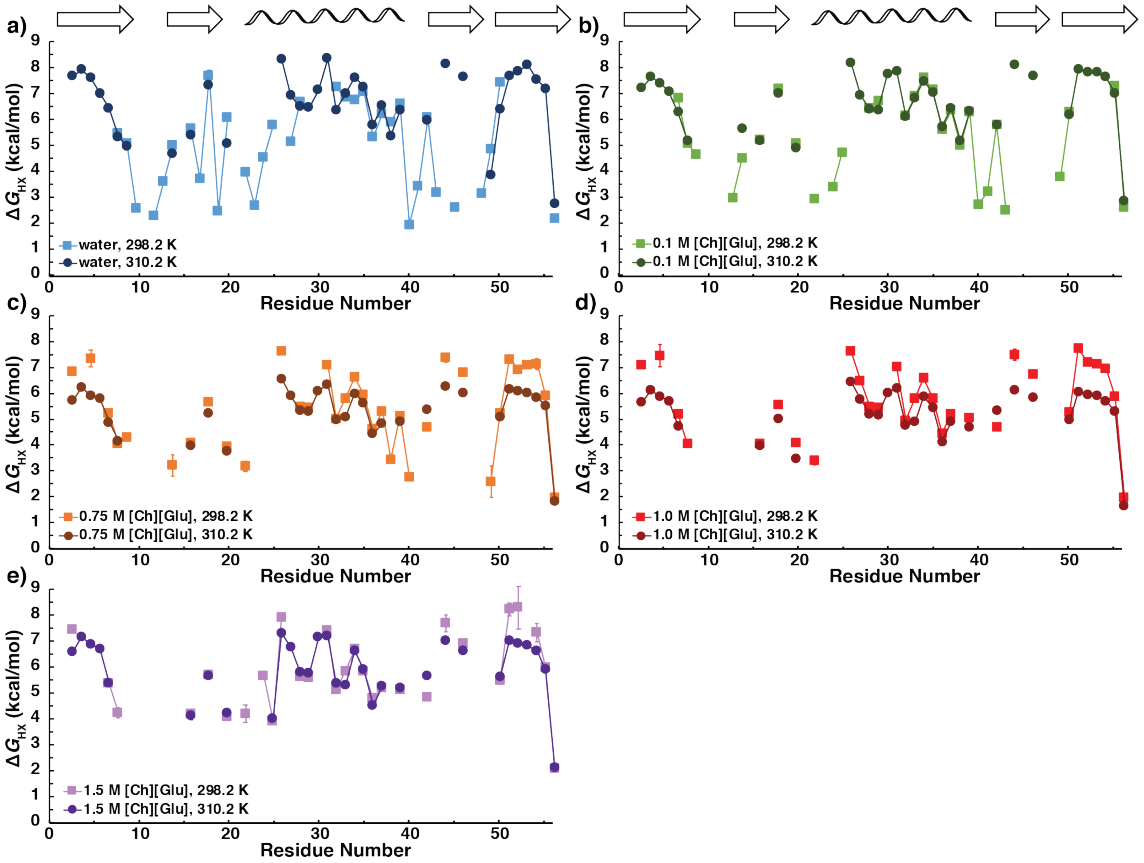



**Figure 2.22. Free energy of exchange ($\Delta G_{HX}$) for residues in GB1.**

Plot of $\Delta G_{HX}$ versus residue number for GB1 in **a**) water (blue), and **b**) to **e**) for increasing amounts of [Ch][Glu], that is, 0.1 M (green), 0.75 M (orange), 1.0 M (red) and 1.5 M (purple), respectively. The majority of slow exchanging NHs were measured at 310.2 K (circles and dark colour), and fast exchanging NHs were measured at 298.2 K (squares and light colour). $\Delta G_{HX}$ values were calculated accordingly to Eq. 2.1 and the uncertainties were propagated from $k_{ex}$ error. For clarity, separate plots are shown for each condition. Secondary structure elements are illustrated on the top of the figure. All the data was recorded at pH meter reading ($pH_{read}$) ≈ 6.8, which corresponds to a pH corrected ($pD_{corr}$) of ≈ 7.2, due to deuterium effects[91].

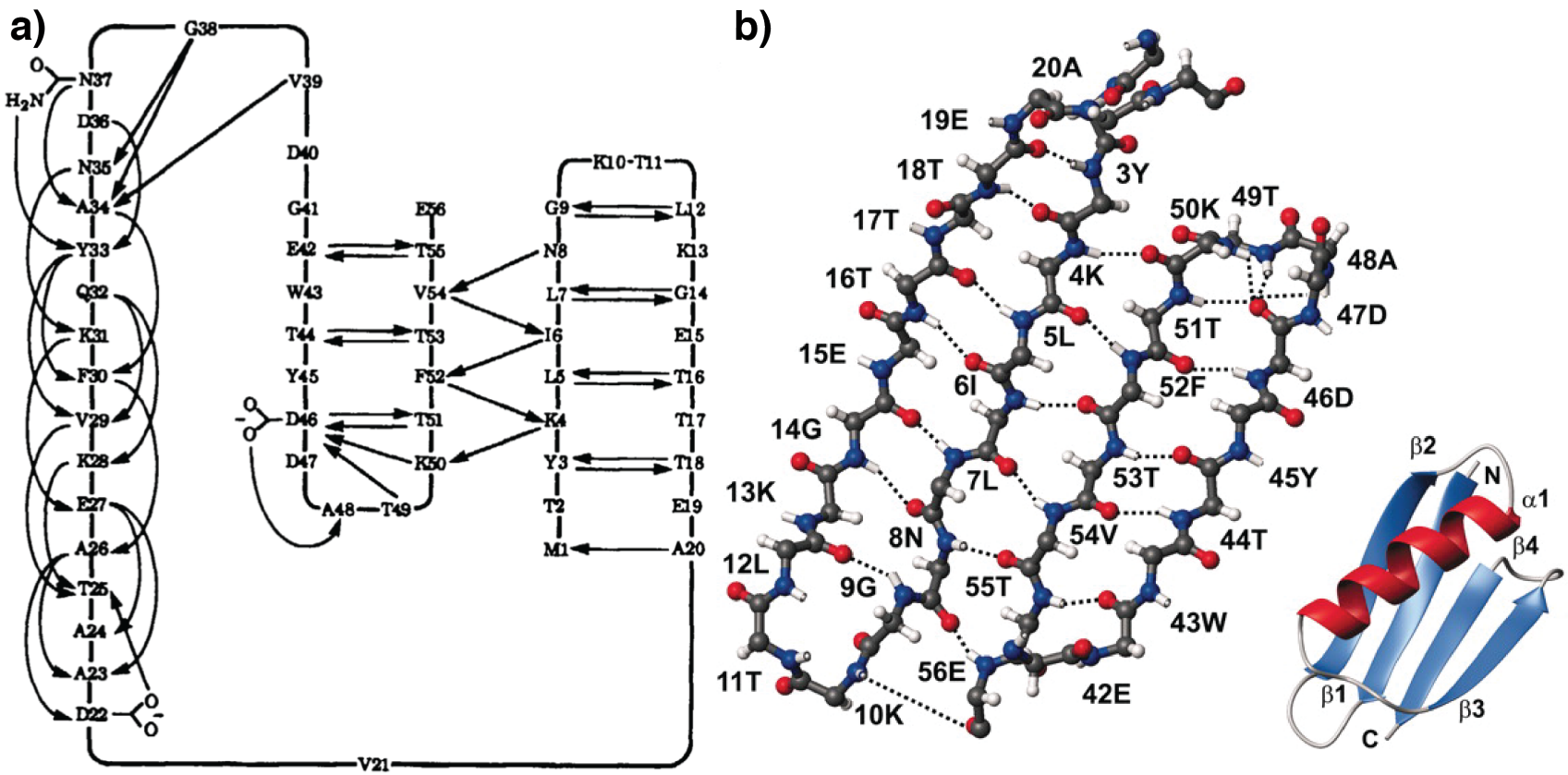



**Figure 2.23. Hydrogen-bond network of GB1.**

**a**) Diagrammatic representation of H-bonding. The involvement of nearly all residues in regular secondary structure (~ 95 %) ensures a large number of stabilising hydrogen bonds. There are a total of 45 hydrogen bonds, of which 41 are backbone-backbone, 3 are side-chain-backbone and 1 side chain-side chain. Over 60% of the backbone donor and acceptor groups (38 backbone amides and 36 backbone carbonyls) participate in H-bonding. The criteria used for defining a hydrogen bond are that the distance between the acceptor and donor heavy atoms is ≤ 3.3 Å and that the angle between the acceptor heavy atom, the hydrogen, and the donor heavy atom is > 120º. Adapted from Gronenborn et al[19]. **b**) Ribbon representation of GB1 structure (right) and the H-bond network (left) underlying the β-sheet structure are depicted in the same orientation. The α-helix and side-chain moieties are removed for clarity. Adapted from Selenko et al[92].

In order to better visualise the effect of [Ch][Glu] concentration on GB1 solvent accessibility, for each temperature, the $\Delta G_{HX}$ values along the protein sequence for the different conditions were plotted together (Fig. 2.24).

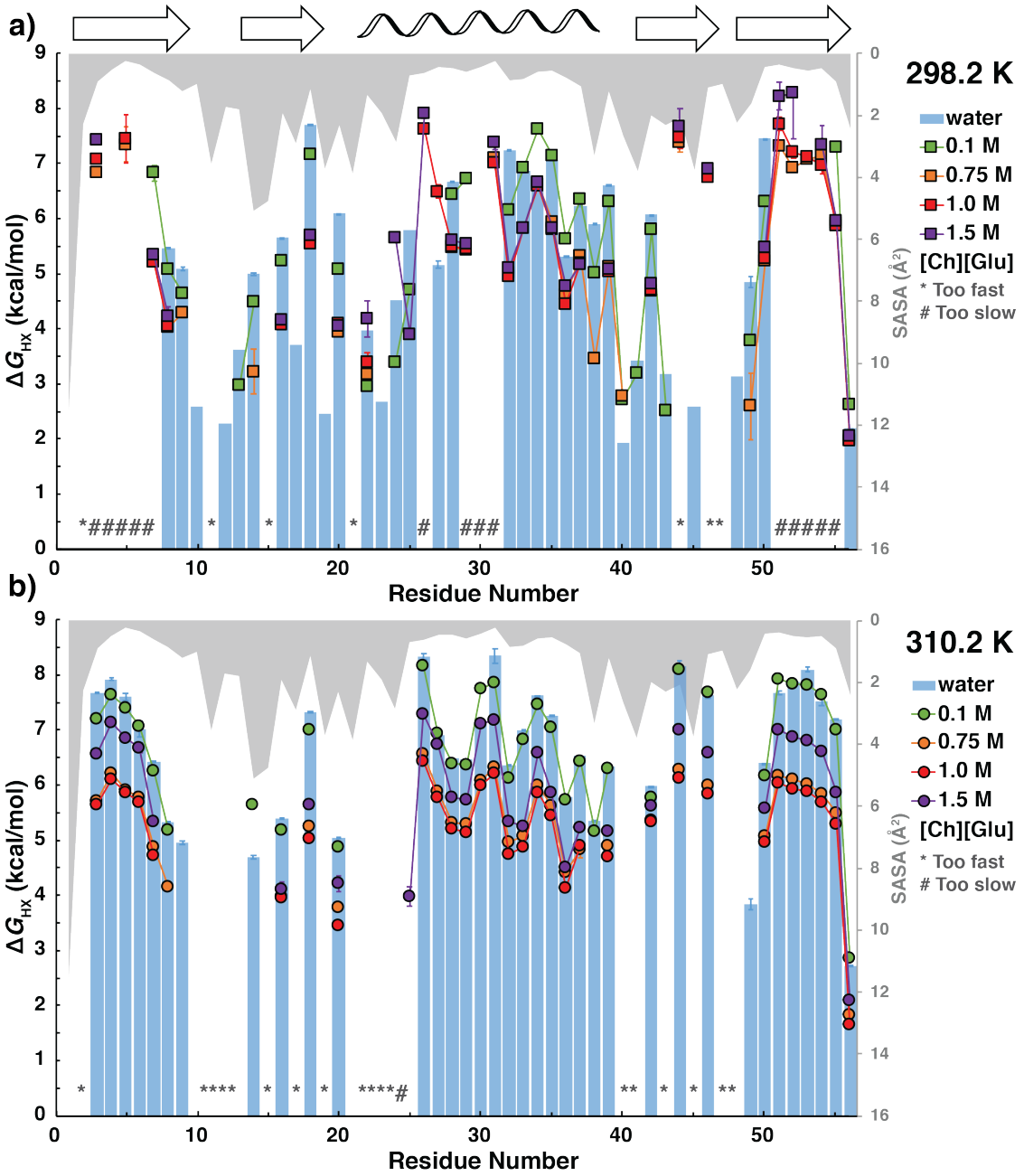



**Figure 2.24. $\Delta G_{HX}$ of GB1's residue-level superposition for [Ch][Glu] titration.**

$\Delta G_{HX}$ for residues in GB1 (left y axis) at **a**) 298.2 K and **b**) 310.2 K, in the absence of IL/ water (blue bars), and 0.1 M (green), 0.75 M (orange), 1.0 M (red) and 1.5 M (purple) of [Ch][Glu] concentration. For the condition in the absence of IL, residues that decay too slowly for accurate measurement are labelled with a hash (#) and those that decay too rapidly with an asterisk (*). The solvent-accessible surface area (SASA) for each backbone nitrogen atom (right y axis, in gray) was computed using the Parameter Optimized Surfaces (POPS) program[93] with the crystal structure of GB1 (PDB code: 2QMT[68], without ions and molecules of

water). For clarity, plots are shown for each temperature with the secondary structure of the protein depicted on the top.

Because rapidly exchanging backbone NHs are more likely to have greater solvent exposure, I observe that the $\Delta G_{HX}$ has an approximate inverse relationship with the pattern of solvent-accessible surface area[26] (SASA is defined as a probe sphere of a particular radius rolling over the protein surface[93]). Indeed, the average SASA for the residues that decay too slowly, or too quickly, to be quantifiable at both temperatures are $0.5 \pm 0.2$ Å$^2$, or $3 \pm 1$ Å$^2$, respectively, where the uncertainties are the sample standard deviation. In general, I find that addition of [Ch][Glu] decreases $\Delta G_{HX}$ values in a non-monotonic behaviour (with a shallow minimum at 0.75 and 1.0 M of concentration), in parallel to the behaviour of GB1's $T_m$ (as noted in Fig. 2.4). The values of $\Delta G_{HX}$ in water are similar to those in 0.1 M of [Ch][Glu] and followed by 1.5 M (Fig. 2.24). Note that a positive value of $\Delta\Delta G_{HX} = \Delta G_{HX,\ cosolute} - \Delta G_{HX,\ water}$ is associated with the residue being more protected from solvent exchange when compared to the reference aqueous solution (Fig. 2.25). Deviations from this general trend at 298.2 K are residues D22 and A24, which become more protected ($\Delta\Delta G_{HX} > 0$, $\Delta G_{HX,\ cosolute} > \Delta G_{HX,\ water}$) in 1.5 M [Ch][Glu] (when compared to water), as well as D40 in the presence of 0.75 M of concentration. Interestingly, the D40K surface mutation on GB1 is 10-fold more destabilising in cells than in buffer[69] because it increases non-specific and attractive quinary interactions with the cytoplasmic milieu. These observations agree with their changes in chemical shift and amide proton coefficients (Fig. 2.8 and Fig. 2.15) and not only reveals a local stabilisation of the protein (by means of the D22-A24 (un)folding hotspot[25]) but also an surface residue (D40) as important as core residues to modulate stability.

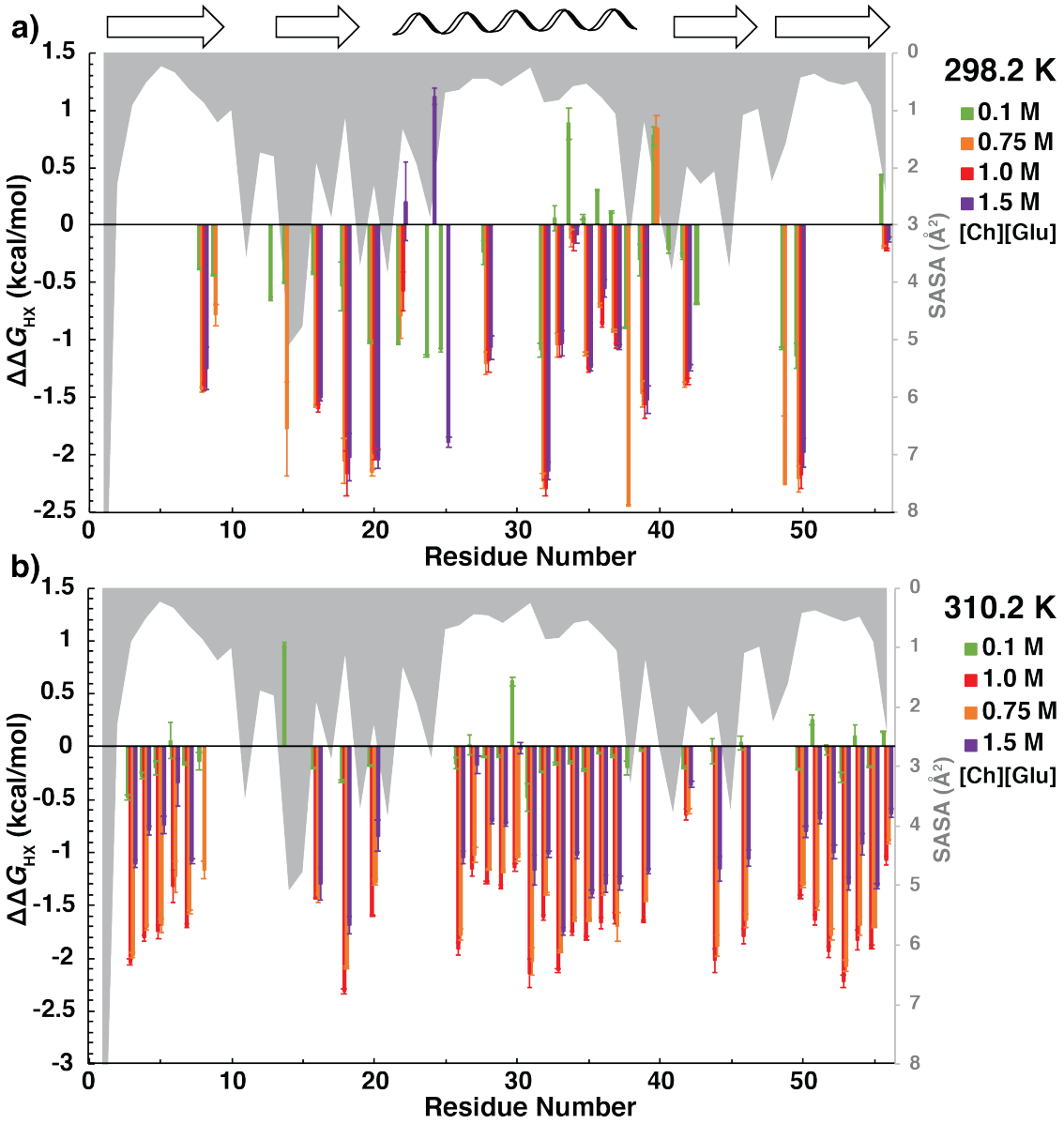



**Figure 2.25. $\Delta\Delta G_{HX}$ of GB1's residue-level superposition for [Ch][Glu] titration.**

Change in $\Delta G_{HX}$ ($\Delta\Delta G_{HX} = \Delta G_{HX,cosolute}$-$\Delta G_{HX,water}$, whereas the reference is the $\Delta G_{HX}$ determined in the absence of IL / water) for residues in GB1 (left y axis) at **a**) 298.2 K and **b**) 310.2 K for 0.1 M (green), 0.75 M (orange), 1.0 M (red) and 1.5 M (purple) of [Ch][Glu] concentration. The solvent-accessible surface area (SASA) for each backbone nitrogen atom (right y axis, in gray) was computed using the Parameter Optimized Surfaces (POPS) program[93] with the crystal structure of GB1 (PDB code: 2QMT[68], without ions and molecules of water). For clarity, a separate plot is shown for each temperature with the secondary structure of the protein depicted on the top.

Global stability from $N^1H/N^2H$ exchange measurements

A global unfolding mechanism can be estimated from the exchange rates of selected backbone amides. Because they are strongly involved in H-bonded regions of secondary structure, their exchange happens only in the fully unfolded state, yielding a maximum value of opening free energy ($\Delta G_{HX}$)[94]. In GB1, such residues are located in β1 (Y3, K4), α1

(A26, K28, V29, K31, A34), β3 (T44, D46) and β4-strand (T51, F52, T53). Hence, the $\Delta G_{HX}$ value approximates the free energy of unfolding $\Delta G_u$ (extrapolated from calorimetry) providing information on global stability. This approximation was validated by Orban et al[23] and, more recently, by Monteith and Pielak[26]. Since, using Differential Scanning Calorimetry (DSC), the authors could not obtain thermodynamic data for the (wild type) GB1 due to its high thermal stability ($T_m$ = 79.0 ± 0.2 ºC, in buffer ≈ 0.01 M sodium phosphate, 0.15 M NaCl, $pH_{read}$ ≈ 7.2), they used the destabilised I6L variant of GB1 ($T_m$ = 75.6 ± 0.2 ºC, in buffer 0.15 M NaCl) and obtained an excellent agreement between the average of $\Delta G_{HX}^{0'}$ (of those 12 residues) and $\Delta G_u^{0'}$ at 310.2 K ($\Delta G_{HX,I6L}^{0'}$ = 6.3 ± 0.3 ≈ $\Delta G_{u,I6L}^{0'}$ = 6.30 ± 0.01 kcal/mol). The approximation has proven to be valid to the wild-type form by comparing $\Delta\Delta G$ values ($\Delta\Delta G_{HX,WT\text{-}I6L}^{0'}$ = -0.9 ± 0.3 ≈ $\Delta\Delta G_{u,WT\text{-}I6L}^{0'}$ = -0.68 ± 0.06 kcal/mol, where $\Delta G_{HX,WT}$ = 7.1 ± 0.5 kcal/mol).

Here, exchange rates from backbone amide of these 12 residues could be accurately quantified only at 310.2 K (Table A3 and Table A4 of Appendix A). Assuming that the residues that exchange via global unfolding are the same in water and in the presence of [Ch][Glu], their $\Delta G_{HX}$ average was used as an approximation for the global free energy of unfolding $\Delta G_u$ for the analysis that follows. For clarity, the corresponding $\Delta G_{HX}$ are reproduced in Table 2.2.

**Table 2.2. $\Delta G_{HX}$ values (kcal/mol) for GB1's selected residues in water and [Ch][Glu].**

| | | | [Ch][Glu] (M) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | water | | 0.1 | | 0.75 | | 1.0 | | 1.5 | |
| | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE |
| **Y3** | **7.69** | 0.02 | **7.21** | 0.005 | **5.719** | 0.009 | **5.65** | 0.01 | **6.58** | 0.02 |
| **K4** | **7.93** | 0.02 | **7.65** | 0.01 | **6.221** | 0.007 | **6.12** | 0.01 | **7.14** | 0.02 |
| **A26** | **8.34** | 0.05 | **8.187** | 0.006 | **6.570** | 0.004 | **6.434** | 0.007 | **7.30** | 0.01 |
| **K28** | **6.494** | 0.002 | **6.400** | 0.005 | **5.334** | 0.005 | **5.21** | 0.01 | **5.79** | 0.02 |
| **V29** | **6.471** | 0.006 | **6.382** | 0.004 | **5.296** | 0.005 | **5.146** | 0.007 | **5.74** | 0.01 |
| **K31** | **8.4** | 0.1 | **7.876** | 0.009 | **6.33** | 0.01 | **6.215** | 0.009 | **7.19** | 0.02 |
| **A34** | **7.631** | 0.008 | **7.479** | 0.004 | **5.997** | 0.008 | **5.871** | 0.009 | **6.60** | 0.02 |
| **T44** | **8.2** | 0.1 | **8.12** | 0.01 | **6.282** | 0.006 | **6.14** | 0.01 | **7.01** | 0.02 |
| **D46** | **7.646** | 0.057 | **7.684** | 0.007 | **6.005** | 0.008 | **5.85** | 0.01 | **6.59** | 0.02 |
| **T51** | **7.68** | 0.04 | **7.932** | 0.006 | **6.180** | 0.006 | **6.042** | 0.009 | **7.01** | 0.01 |
| **F52** | **7.88** | 0.04 | **7.852** | 0.008 | **6.11** | 0.01 | **5.94** | 0.02 | **6.88** | 0.03 |
| **T53** | **8.11** | 0.04 | **7.820** | 0.006 | **6.04** | 0.01 | **5.90** | 0.01 | **6.81** | 0.02 |
| **Mean** | **7.7** | 0.6 | **7.5** | 0.6 | **6.0** | 0.4 | **5.9** | 0.4 | **6.7** | 0.5 |

$\Delta G_{HX}$ values for the 12 residues that approximate an exchange by global unfolding. The standard error (SE) was obtained from propagation of the error of $k_{ex}$. In the bottom is indicated the average of $\Delta G_{HX}$ values and their standard deviation of the mean.

While their average corresponds to the free energy associated with global unfolding, lower values of $\Delta G_{HX}$, in the secondary structure, might reflect a destabilisation of the native

state (or stabilisation of U-state) leading to local fluctuations, thus allowing NH exchange from a small fraction of partially unfolded forms (including two-state folding)[95–97]. For clarity, the plots of $\Delta G_{HX}$ *versus* protein residue (at 310.2 K) are reproduced in Fig. 2.26. The average of the values obtained for those global unfolders in water ($\Delta G_{HX,water}^{0}$ = 7.7 ± 0.6 kcal/mol, $pH_{read} \approx 6.8$) are consistent with those previously published for buffer ($\Delta G_{HX,buffer}^{0}$ = 7.0 ± 0.1 kcal/mol)[26], considering that the sodium phosphate buffer with 0.15 M NaCl destabilises GB1. In [Ch][Glu], on average, the determined $\Delta G_{HX}$ values decrease, as compared with water. This decrease is associated with an overall destabilisation of the protein (most of the residues are less protected from solvent exchange in IL conditions than in water). Despite this decrease, which depends on the IL concentration (Fig. 2.26), the $\Delta G_{HX}$ profile or the pattern of H/D exchange is similar. Thus, I can reasonably assume that the structures of the native state are the same in water and in IL-aqueous solutions.

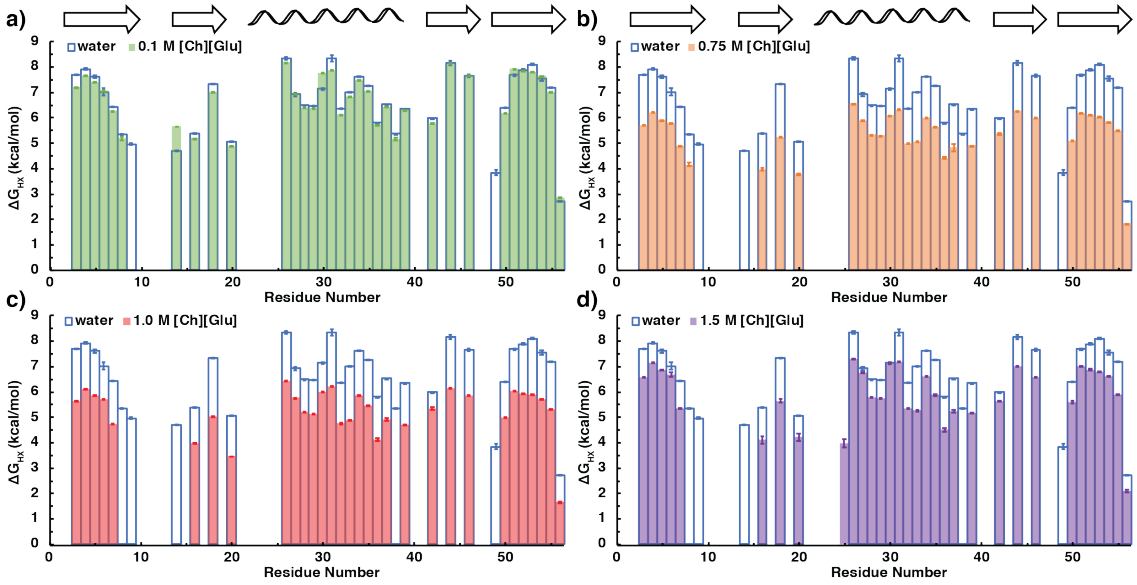



**Figure 2.26. $\Delta G_{HX}$ of GB1's residue-level in the presence of [Ch][Glu].**

$\Delta G_{HX}$ for GB1's residues in the absence of IL/ water (blue bars), and in the presence of **a)** 0.1 M (green), **b)** 0.75 M (orange), **c)** 1.0 M (red) and **d)** 1.5 M (purple) of [Ch][Glu]. $\Delta G_{HX}$ values were calculated accordingly to the equation 1 and the uncertainties were propagated from $k_{ex}$ error. For clarity, separate plots are shown for each condition and secondary structure elements are illustrated on the top of the figure. All the data was recorded at 310.2 K and $pD_{corr} \approx 7.3$.

Additionally, when trying to determine the free energy of denaturation for GB1 in water and in the presence of [Ch][Glu] using DSC, not only the high thermal stability of GB1 ($T_m$ = 83.3 ± 0.3 °C, water) but also the high concentration of [Ch][Glu] (> 0.1 M) contributes to inadequate post transitional baselines, preventing this determination. Still, despite the limited DSC data (Table 2.3), $\Delta G_u$ values could be obtained for water and 0.1 M [Ch][Glu]. These are, however, lower than those determined by H/D exchange, but still within the standard deviation of the mean $\Delta G_{HX}$ for the 12 residues (Table 2.2 and 2.3). The $T_m$ values

derived here are also in good agreement with those obtained by fluorescence experiments (as discussed in the section of ILs and salts effects on protein stability).

**Table 2.3. Thermodynamic parameters of GB1 determined by DSC and NMR.**

| [Ch][Glu] (M) | $T_m$ (°C) | $\Delta H_u^{0'}$ (kcal/mol) | $\Delta G_u^{0'}$ (kcal/mol) | $\Delta G_{HX}^{0'}$ (kcal/mol) |
|---|---|---|---|---|
| **0.0** | 83.3 ± 0.3 | 70 ± 2 | 6.9 ± 0.6 | 7.7 ± 0.6 |
| **0.1** | 80.5 ± 0.3 | 69.7 ± 0.7 | 6.6 ± 0.5 | 7.5 ± 0.6 |
| **0.75** | - | - | - | 6.0 ± 0.4 |
| **1.0** | 77.4 ± 0.2 | n. d. | n. d. | 5.9 ± 0.4 |
| **1.5** | - | - | - | 6.7 ± 0.5 |
| **2.0** | 83.0 ± 0.3 | n. d. | n. d. | - |
| **Buffer 0.1 M Sodium Phosphate** | 74.28 ± 0.08 | 76 ± 2 | 6.7 ± 0.6 | - |

Experimental $T_m$ and $\Delta H_u$ obtained directly with DSC. $\Delta G_u$ values were extrapolated to 310.2 K as described by Becktel and Schellman[98] and assuming a temperature-independent $\Delta C_p$ of 0.69 ± 0.07 kcal/mol K (2.9 ± 0.3 kJ/mol K)[20]. The uncertainties of experimental parameters correspond to the standard deviation of the mean from three trials, and the $\Delta G_u$ uncertainties from propagation of error analysis. n. d. Not determined, due to inadequate post transitional baselines. The average of $\Delta G_{HX}$ values, determined by NMR, and their standard deviation of the mean are shown in the right. Details can be found in Material and Methods section.

Figure 2.27a shows the relative change in the stability, $\Delta\Delta G_u^{0'} \approx \Delta\Delta G_{HX}^{0'} = \Delta G_{HX,cosolute}^{0'} - \Delta G_{HX,water}^{0'}$ for the selected representative residues. As the average $\Delta G_{HX}$ approximates $\Delta G_u$, constant values of $\Delta\Delta G_{HX}$ across the different residues might be expected. However, variations are observed e.g., K28, V29, K31 and T53 values often differ from the mean by more than 1 standard deviation. Besides the inherent uncertainty in $k_{int}$, as estimated from model and unstructured peptides[87], subtle differences in their protection factors ($Pf = k_{int}/k_{ex}$), possibly affected by $Ch^+$, $Glu^-$ ions and/or ion-pair, might also contribute for such variations. Because these effects cannot be separated, I focus my analysis on the average values of $\Delta\Delta G_{HX}$. The plot of $\Delta\Delta G_{HX}$ *versus* [Ch][Glu] concentration is a shallow parabola with a minimum between 0.75 and 1.0 M (Fig. 2.27b, whereby a stability order can be observed: pure water ≥ 0.1 M > 1.5 M > 0.75 ≥ 1.0 M). The stability changes for 0.1 M [Ch][Glu] is also consistent with those found by DSC data ($\Delta\Delta G_{HX,\ 0.1\ M\ [Ch][Glu]} = -0.1 \pm 0.2 \approx \Delta\Delta G_{u,\ ,\ 0.1\ M\ [Ch][Glu]}^{0'} = -0.3 \pm 0.8$ kcal/mol).

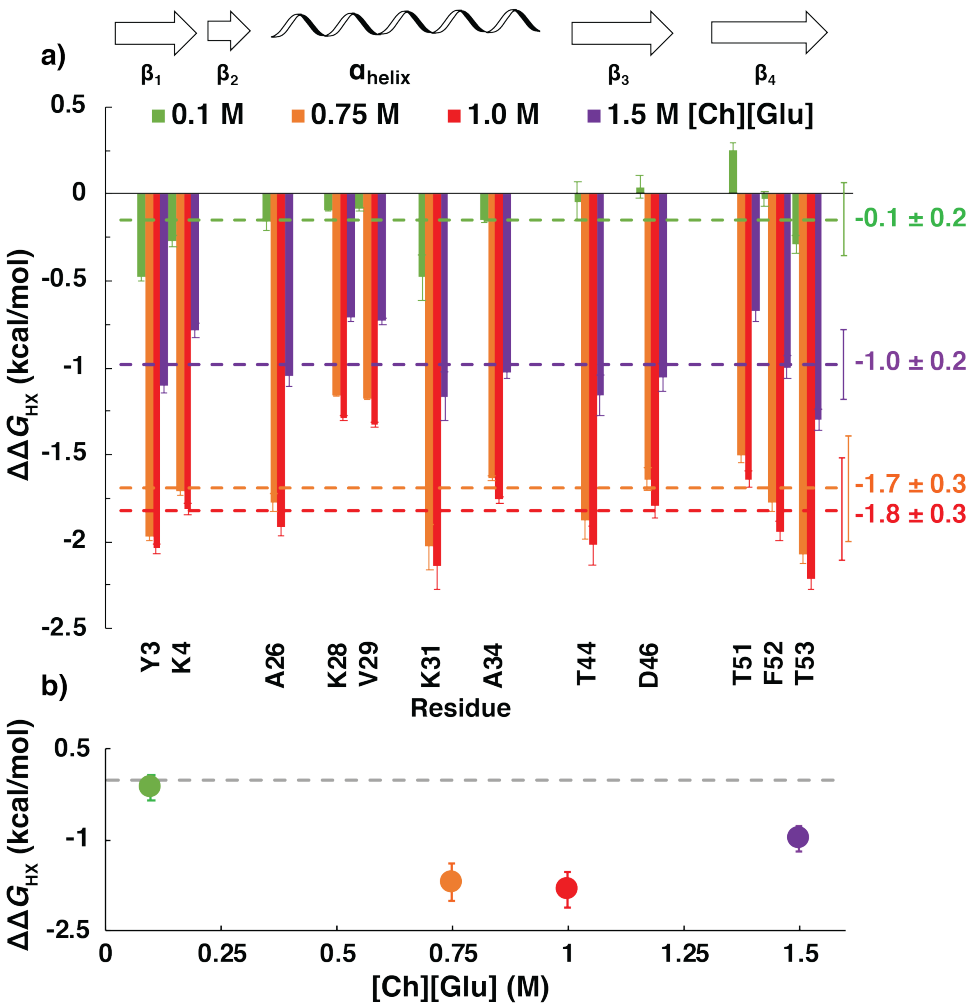



**Figure 2.27. [Ch][Glu]-induced changes of GB1 stability.**

**a**) $\Delta\Delta G_{HX}^{0'}$ ($\Delta G_{HX,cosolute}^{0'}$ - $\Delta G_{HX,water}^{0'}$) of selected residues caused by addition of [Ch][Glu] along protein sequence. Error bars represent the uncertainty propagated from $k_{ex}$ fitting error. Dashed lines and associated error bars are the average $\Delta\Delta G_{HX}^{0'}$ values and their standard deviation of the mean. **b**) Plot of average $\Delta\Delta G_{HX}^{0'}$ values and their SD along [Ch][Glu] concentration. Negative values indicate decreased stability.

The stability data, summarized in the following Table 2.4, shows a linear relationship between changes in the folding free energy ($\Delta\Delta G_{HX}$) and $\Delta T_m$ obtained through fluorescence measurements (i.e., a shallow minimum of $\Delta G_{HX}$ and $T_m$ near 0.75 and 1.0 M [Ch][Glu] are observed for the two experiments). This could indicate that the observed stability behaviour in increasing amounts of [Ch][Glu] is mainly driven by an entropic mechanism. While protein destabilisation is associated with the residues being less protected from exchange with the solvent at 0.75 - 1.0 M of IL (maximum of preferential hydration), the protein is more stable when the residues are more protected from exchange with the solvent as in the presence of [Ch][Glu] at high concentrations (resembling a mechanism of preferential accumulation). These ideas will be further discussed in the summary and outlook chapter.

**Table 2.4: [Ch][Glu]-induced changes of $\Delta G_{HX}^{0'}$ and $\Delta T_m$.**

| | $N^1H/N^2H$ exchange [(a)] | | Intrinsic fluorescence [(b)] | |
|---|---|---|---|---|
| **[Ch][Glu] (M)** | **$\Delta G_{HX}^{0'}$ (kcal/mol)** | **$\Delta\Delta G_{HX}^{0'}$ (kcal/mol)** | **$T_m$ (°C)** | **$\Delta T_m$ (K)** |
| **0.0** | 7.7 ± 0.6 | 0 | 80 ± 1 | 0 |
| **0.1** | 7.5 ± 0.6 | -0.1 ± 0.2 | 77.6 ± 0.2 | -3 ± 1 |
| **0.75** | 6.0 ± 0.4 | -1.7 ± 0.3 | 73.5 ± 0.2 | -7 ± 1 |
| **1.0** | 5.9 ± 0.4 | -1.8 ± 0.3 | 74.6 ± 0.2 | -6 ± 1 |
| **1.5** | 6.7 ± 0.5 | -1.0 ± 0.2 | 79.1 ± 0.4 | -1 ± 1 |

a) Average 12-residues $\Delta G_{HX}^{0'} \approx \Delta G_u^{0'}$ values and their excess changes with [Ch][Glu] using as reference water at 310.2 K (37 ºC). Uncertainties represent the standard deviation from the average. **b**) $T_m$ and their changes with [Ch][Glu] using as reference $T_m$ of GB1 in water. The values and uncertainties were calculated as discussed above.

In addition, $T_m$ values were also determined using NMR, exclusively using side chain methyl groups as probes for global protein stability[99]. The peak intensity of the non-exposed methyl groups of A26, A34 ($^{13}C\beta$) and V54 ($^{13}C\gamma a$) were followed in deuterated samples ($D_2O$, 0.1 M, 0.25 M, 0.5 M, 1.5 M [Ch][Glu] and buffer 0.1 M sodium phosphate) at increasing temperatures (Fig. 2.28 and Table 2.5). Since the change in methyl side chain intensity could not be strictly due to unfolding, the reported values are less consistent than those obtained by fluorescence/ DSC. However, they are in line with GB1 being less stable in 0.5 M [Ch][Glu] than in water or in 1.5 M of IL ($\Delta T_{m,\ 0.5M\ [Ch][Glu]\text{-}water}$ = -4.0 ± 0.8 and $\Delta T_{m,\ 1.5M\ [Ch][Glu]\text{-}water}$ = -3.9 ± 0.7) and with the observed destabilisation caused by the buffer ($\Delta T_{m,\ buffer\text{-}water}$ = -6.0 ± 0.7).

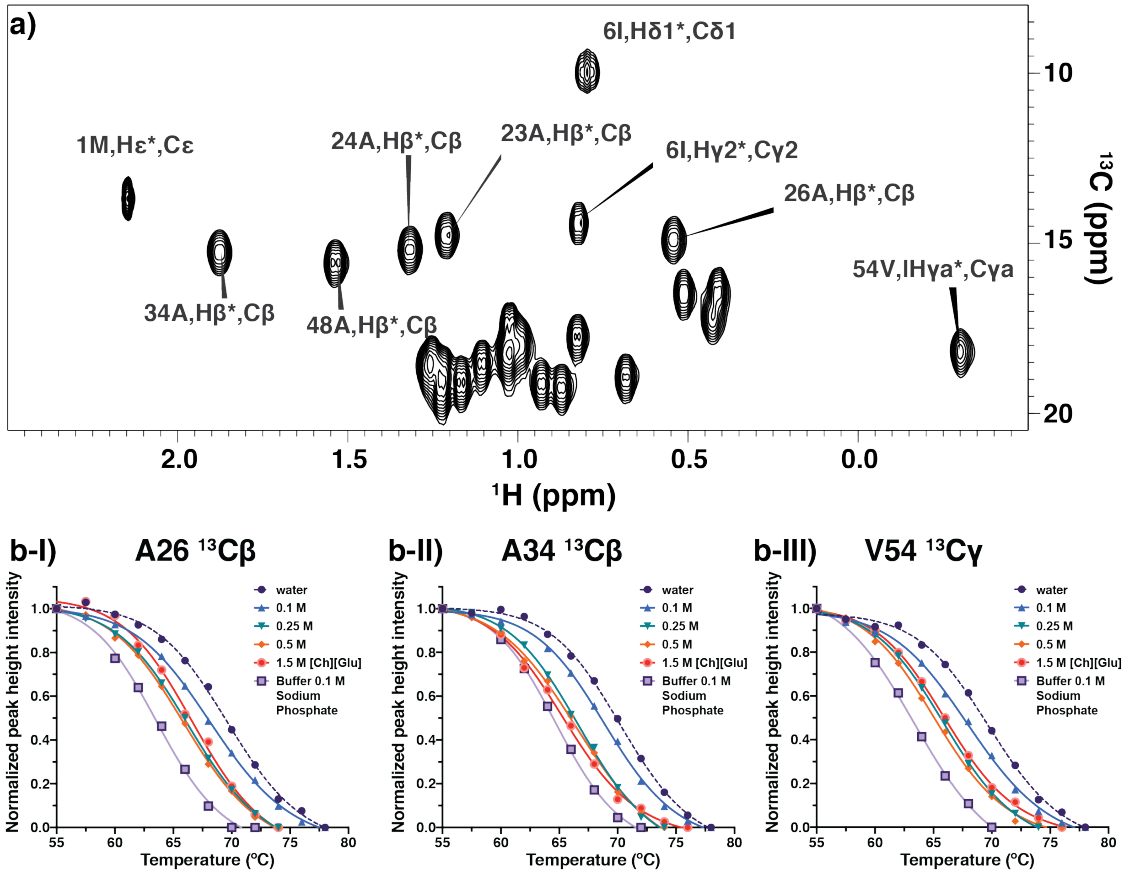



**Figure 2.28. GB1 stability through their non-exposed methyl's.**

**a**) Selected region from 2D $^{1}$H,$^{13}$C HSQC spectrum of 0.3 mM $^{15}$N$^{13}$C GB1 in $D_2O$, 298.2 K, 400 MHz. The label indicates side chain methyl group assignment as based on previous publications[99,100] and on backbone chemical shifts determined above. For the different conditions (water, 0.1, 0.25, 0.5, 1.5 M [Ch][Glu] and buffer 0.1 M sodium phosphate), peak heigh intensity was extracted from each spectrum acquired at a given temperature. **b**) Normalized peak heigh intensity of non-exposed side chain methyl groups A26 and A34 $^{13}$Cβ, and V54 $^{13}$Cγa (I, II, and II, respectively), from 55ºC to 78ºC, are shown in $D_2O$ and in the presence of increasing amounts of [Ch][Glu]. The buffer condition is also indicated. $T_m$ values were determined by fitting the intensity decay to a sigmoidal function using Prism 8.

**Table 2.5. Change in $T_m$ of GB1 determined by methyl peak intensity.**

| [Ch][Glu] (M) | $T_m$ (°C) | $\Delta T_m$ (°C) |
|---|---|---|
| **0.0** | 70.1 ± 0.3 | 0 |
| **0.1** | 68.6 ± 0.4 | -1.5 ± 0.5 |
| **0.25** | 66.5 ± 0.4 | -3.6 ± 0.5 |
| **0.5** | 66.1 ± 0.7 | -4.0 ± 0.5 |
| **1.5** | 66.2 ± 0.7 | -3.9 ± 0.7 |
| **Buffer 0.1 M Sodium Phosphate** | 64.1 ± 0.7 | -6.0 ± 0.7 |

$T_m$ values are the average of those obtained by monitoring A26, A34 and V54 methyl peak height. The uncertainties are the standard deviation of the mean. $\Delta T_m = T_{m,\,\text{cosolute}} - T_{m,\,\text{water}}$ and their uncertainties were propagated from the fitting error for each condition [$\sigma = \sqrt{(\delta_{Tm,\text{cosolute}})^2 + (\delta_{Tm,\text{water}})^2}$]. Details can be found in Material and Methods section.

CLEANEX-PM and fast exchanging residues

In order to obtain information on some of the most affected regions by [Ch][Glu] (as observed by chemical shift changes), in particular for the solvent exposed L12, A23 and D40 residues, for which exchange data is absent or limited, I used a water-saturation transfer experiment, CLEANEX-PM (phase-modulated clean chemical exchange)[101] to measure the rates of the faster exchanging residues. In contrast to traditional amide $^1$H exchange by 2D HSQC that allows to determine exchange rates on the order of $k_{ex} \approx 10^{-3}$ to $10^{-8}$ s$^{-1}$, (CLEANEX-PM)-FHSQC does not require fully deuterated solutions and can measure much faster exchange rates i.e., ≈ 1 - 100 s$^{-1}$ [102]. In this experiment, designed specifically to study solvent-exposed residues, the water resonance is selectively excited and allowed to exchange over a variable mixing period ($\tau_m$) with the amide proton spins. At the end of the mixing period a fast HSQC sequence is used for detection, and the amide resonances peak intensities are monitored for different $\tau_m$ on a per-residue basis. Using short mixing times (< 50 ms), only the fast-exchanging amides will receive enough magnetization coming from water, thus, only they will be visible. Here, the exchange rate ($k_{ex}$) is approximated to the initial time-dependent slope[101] (lines in $V/V_0$ vs $\tau_{mix}$ plots - Fig. 2.29, and details can be found in Material and Methods section).

Using CLEANEX-PM in the different conditions (water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu]), 12 residues could be monitored at 310.2 K: Q2, K10, T11, L12, E15, T17, E19, V21, A33, Y45, D47 and A48. Of these, only Q2, T11, E15, V21 and D47 could be followed also at 298.2 K dataset. Values of $k_{ex}$ and corresponding $\Delta G_{CLEANEX}$ values (determined in a similar way as $\Delta G_{HX}$ – Eq. 2.1), are tabulated in Table A5 and A6 of Appendix A, respectively. The exchange peaks with slower rates ($k_{ex} < 1$ s$^{-1}$) have diminished intensities (or not appear at all) which difficult their analysis, therefore the data measured at 310.2 K is more reliable than the one acquired at lower temperature.

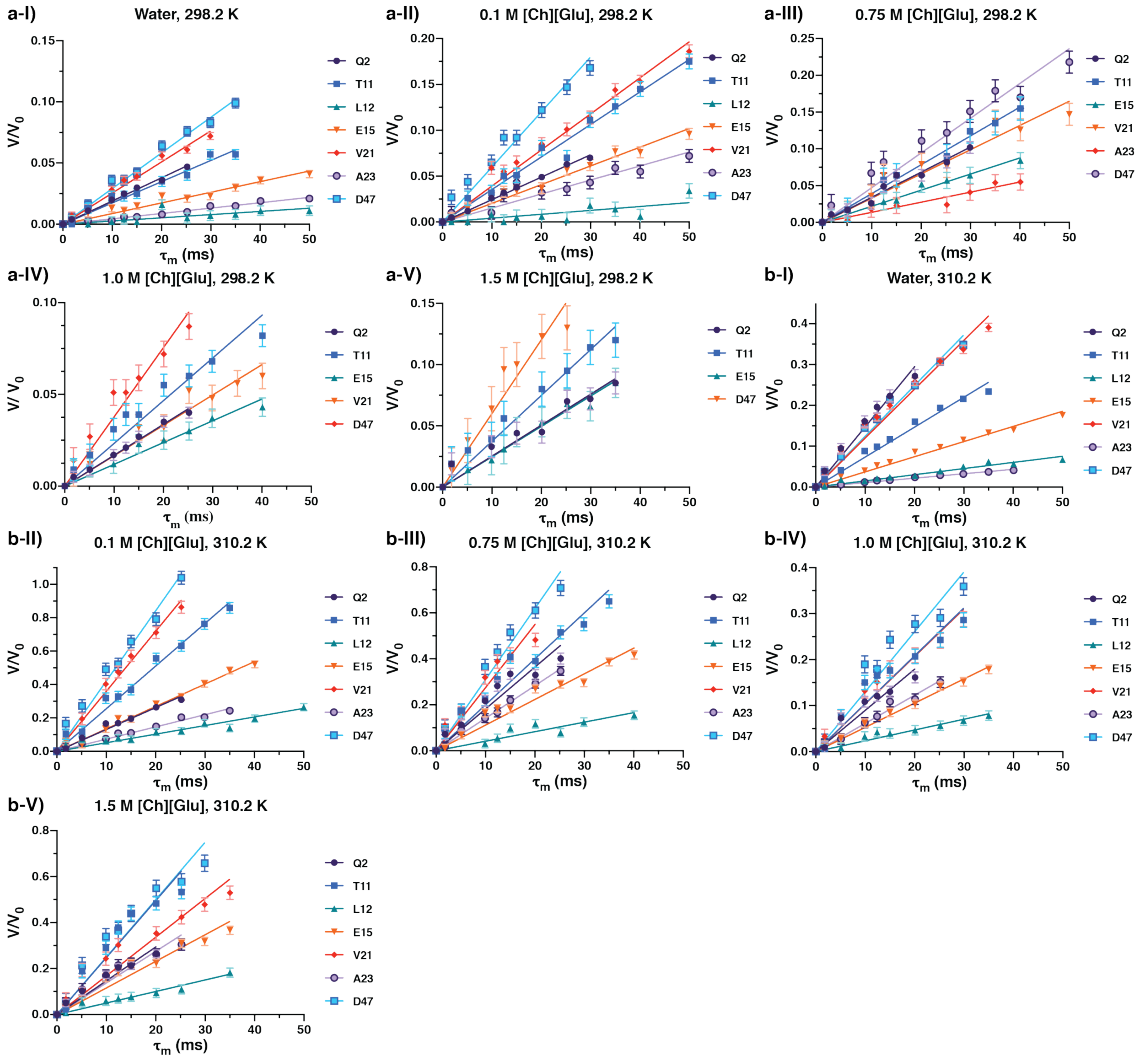



**Figure 2.29. Examples of initial slope analysis applied to fast exchanging residues in GB1.**

Initial slope analysis applied to different cross-peaks with corresponding linear fits for Q2, T11, L12, E15, V21, A23, D47 residues in **a**) 298.2 K and **b**) 310.2 K for **I-V**), absence of IL/ water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentration, respectively. For a few conditions at 298.2 K: L12, V21, A23, are not represented because they are not accurately followed. For each amide proton, cross-peak volume intensity (V) was analysed and extracted from (CLEANEX-PM)-FHSQC at different mixing times using PINT[64,88], and normalized with a reference peak volume ($V_0$) measured from FHSQC spectrum. At short mixing time, the initial slope corresponds to $k_{ex}$. These points were plotted without the correction for water saturation factors and the errors were obtained from propagation of peak volume error ($[\sigma = (V/V_0)\sqrt{((\delta_{V0}/V_0)^2 + (\delta_V/V)^2)}]$.

In general, unprotected surface hydrogens on unstructured segments do exchange rapidly, at their expected unprotected rates with protection factor ≈ 1 (log *Pf* ≈ 0). Small differences in these exchange rates (retarding or accelerating) might be explained by local electrostatic effects[102]. If $k_{ex} > k_{int}$, i.e., faster than predicted in ref.[87], *Pf* < 1 and $\Delta G < 0$. In the absence of IL, for the 12 NHs monitored by CLEANEX (Fig. 2.30), I observed that not only amide hydrogens on dynamically unstructured protein segments in contact with solvent (T11, L12, V21, D47) but also unprotected/ solvent-exposed amide hydrogens placed on β2

strand (E15 and E19) do exchange at or faster than their expected free peptide rate ($\Delta G < \sim 0$). The other NHs (Q2, K10, T17, A23, Y45, and A48) have slightly slower rates than the expected by $k_{int}$ ($\Delta G_{CLEANEX,\ water,\ 310.2\ K} \approx 1.1 \pm 0.5$ kcal/mol). Of the unstructured segments, only the loop from G38 to G41 is less solvent-exposed as there are no cross-peaks detected in CLEANEX-PM spectra. This agrees with the view that the amide group of D40 is not so solvent-exposed and it could form a transient and intermolecular H-bond with the side chain of the C-terminal E56[73–75]. The addition of [Ch][Glu] decreases $\Delta G_{CLEANEX}$ of the residues under study (Fig. 2.30), further increasing the exchange with the solvent, in the following order: pure water < 1.0 M ≤ 1.5 M < 0.75 < 0.1 M. Despite the change for most of the backbone amides being constant, the exchange rate of Q2 is unaffected and NHs from A23 and Y45 show a transition from slower-to-faster (compared to the random coil expectation). A23 is the most affected residue with 0.1 or 0.75 M of [Ch][Glu] (Fig. 2.30, $\Delta\Delta G_{CLEANEX} = -1.70 \pm 0.02$) where the decrease of the corresponding protection factor from 4 to 0.25 suggests an accumulation of water molecules, deprotecting this residue. However, above 0.75 M IL, the $\Delta G_{CLEANEX}$ determined for A23 NH, and for other residues, approximates the value of that in water (under 1.0 and/or 1.5 M [Ch][Glu] at 298.2 K, most of NHs become outside of the limit – too slow to be detected). This is associated to a general protection of backbone NHs from solvent exchange and indicative of preferential accumulation of [Ch][Glu] ions at the protein surface. In the case of A23, this could be assigned to a displacement of defined water molecules by $Glu^-$ anion in this region (as discussed above in other sections).

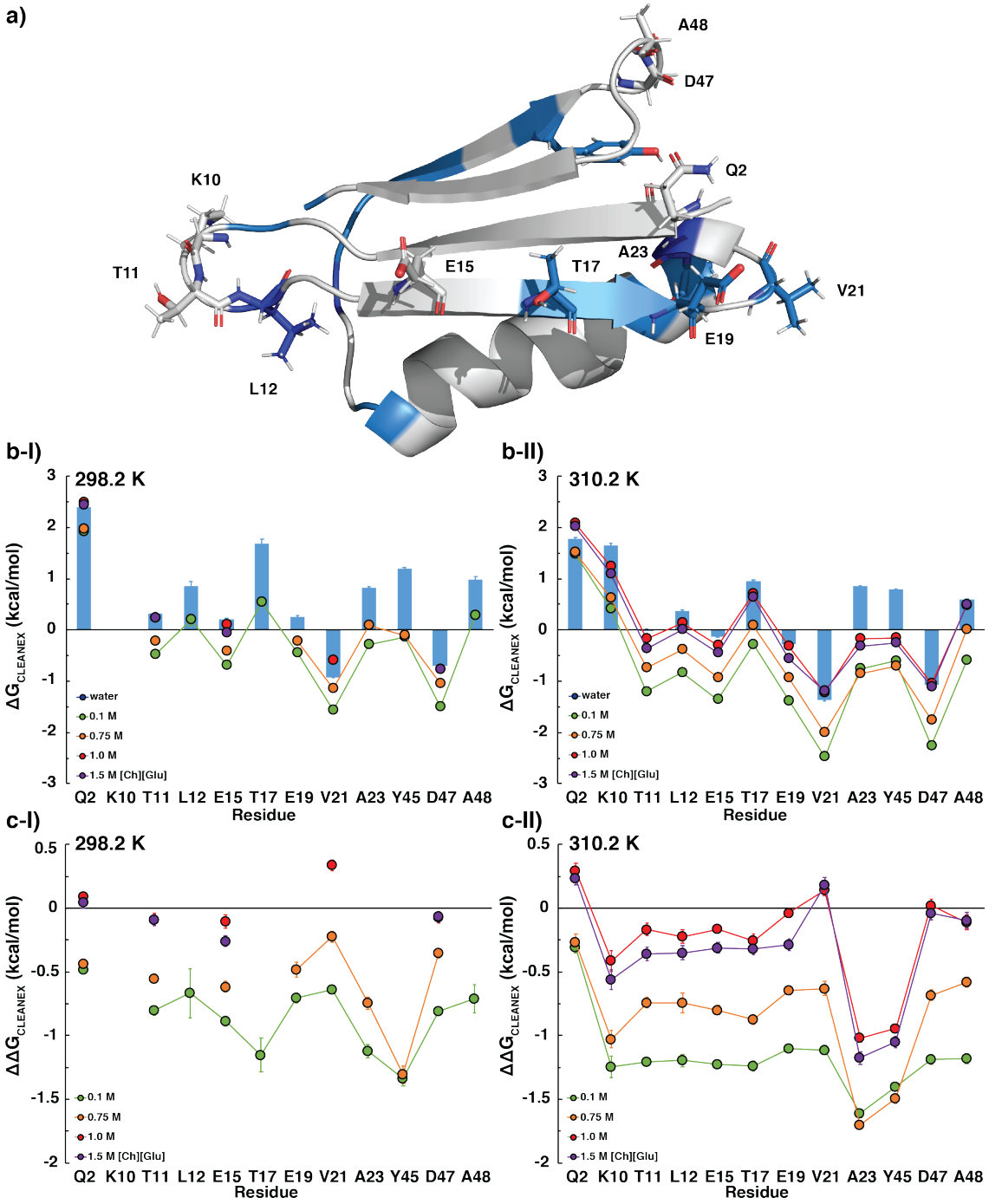



**Figure 2.30. Changes of $\Delta G_{CLEANEX}$ for GB1's fast exchanging residues with addition of [Ch][Glu].**

**a**) Structure of GB1 (PDB: 2JSV[21]) showing all fast exchanging residues followed by CLEANEX (in sticks). The affected residues in the backbone as determined by CSP are mapped at blue. **b)** $\Delta G_{CLEANEX}$ for fast exchanging residues in GB1 at **I**) 298.2 K and **II**) 310.2 K, in water (blue bars), and in 0.1 M (green), 0.75 M (orange), 1.0 M (red) and 1.5 M (purple) of [Ch][Glu]. **c**) Change in $\Delta G_{CLEANEX}$ ($\Delta\Delta G_{CLEANEX} = \Delta G_{CLEANEX,cosolute} - \Delta G_{CLEANEX,water}$, whereas the reference is the $\Delta G_{CLEANEX}$ determined in the absence of IL/ water) for **I**) 298.2 K and **II**) 310.2 K datasets.

$^{15}$N-resolved NOESY and ROESY

As analysed in the introduction of this thesis, the traditional explanation that ions influence bulk and/ or local water structure for water-mediated interactions, although appealing, it is not well supported by experimental evidence[13,103,104]. Nonetheless, I have seen that ions interact directly with some specific protein sites in the presence of water molecules[105]. To study the role of water in the GB1:[Ch][Glu] interactions and in the origin of the exchange rates differences above described, site-resolved measurement of protein-water interactions were conducted via intermolecular dipolar magnetization exchange between water and protein protons[106]. Dipolar exchange and chemical exchange can be distinguished by comparison of both the laboratory-frame nuclear Overhauser effect (NOE) and the rotating-frame NOE (ROE). In the slow tumbling limit, dipolar exchange is identified by the opposite phase of the NOE/ROE cross-peaks at the water resonance, while hydrogen exchange gives rise to ROE peaks of identical phase to the diagonal. However, NMR studies of protein hydration in aqueous solutions have been questioned due to potential contaminations from labile protein hydrogens and long-range dipolar couplings with hydrogens of bulk solvent[107]. These complications could be overcome, for example, by performing the experiments with the protein in a reverse micelle[108,109] (not done here).

Despite the difficulties to interpret NOE/ ROE signals, $^{15}$N-resolved 3D NOESY-HSQC and ROESY-HSQC were acquired for $^{2}$H,$^{15}$N,$^{13}$C-GB1 in the presence and absence of 1 M [Ch][Glu] (at 298.2 K and pH 7.2). The presence of "real"/ direct NOE between the solvating water and amide hydrogens of the protein is assuming that 40-ms of NOESY mixing time here used is within the linear build-up regime and corresponds to a maximum NOE-detection distance of approximately 4.3 Å, as calibrated by Nucci et al[109] for ubiquitin. This should avoid ambiguous cross-peaks involving hydrogen exchange of labile side chain hydroxyl or amine hydrogens (followed by subsequent intramolecular NOE) because these are normally outside of the distance limit established. In addition, the use of a highly deuterated protein minimizes the existence of different H$\alpha$ with similar chemical shifts to the water (4.70 ppm).

The amide hydrogens that showed a cross peak to water with sufficient resolution were qualitatively compared in NOESY and ROESY $^{1}$H-$^{15}$N planes at the water $^{1}$H resonance (Table 2.6 in the following page). As expected, the cross-peaks detected are few and weak. Most of them are highly solvent-exposed and in fast hydrogen exchange with water, in agreement with their detection in CLEANEX experiments.

**Table 2.6. NOE/ ROE cross-peaks to water $^{1}$H resonance interpreted qualitatively on GB1 in water and in 1 M [Ch][Glu].**

| | water | | 1 M [Ch][Glu] | | | water | | 1 M [Ch][Glu] | |
|---|---|---|---|---|---|---|---|---|---|
| | NOE | ROE | NOE | ROE | | NOE | ROE | NOE | ROE |
| Q2 | + | - | + | - | F30 | N/A | N/A | N/A | N/A |
| Y3 | N/A | N/A | N/A | N/A | K31 | N/A | N/A | N/A | N/A |
| K4 | N/A | N/A | N/A | N/A | Q32 | N/A | + | N/A | + |
| L5 | N/A | N/A | N/A | N/A | Y33 | N/A | N/A | N/A | N/A |
| I6 | N/A | N/A | N/A | N/A | A34 | N/A | N/A | N/A | N/A |
| L7 | N/A | + | N/A | N/A | N35 | N/A | N/A | N/A | N/A |
| N8 | N/A | N/A | N/A | N/A | D36 | N/A | N/A | N/A | N/A |
| G9 | N/A | N/A | N/A | N/A | N37 | N/A | N/A | N/A | N/A |
| K10 | + | + | + | N/A | G38 | N/A | N/A | N/A | N/A |
| T11 | + | - | + | - | V39 | N/A | N/A | N/A | N/A |
| L12 | + | - | + | N/A | D40 | N/A | N/A | N/A | N/A |
| K13 | N/A | N/A | + | N/A | G41 | N/A | N/A | + | N/A |
| G14 | N/A | N/A | N/A | N/A | E42 | N/A | N/A | N/A | N/A |
| E15 | + | - | + | - | W43 | N/A | N/A | N/A | N/A |
| T16 | N/A | + | N/A | N/A | T44 | + | + | + | + |
| T17 | + | - | + | + | Y45 | + | - | + | - |
| T18 | N/A | + | N/A | N/A | D46 | N/A | + | + | + |
| E19 | + | - | + | - | D47 | + | - | + | - |
| A20 | N/A | + | + | + | A48 | + | - | + | - |
| V21 | + | - | + | - | T49 | + | + | + | + |
| D22 | + | + | + | + | K50 | + | + | + | N/A |
| A23 | + | - | + | - | T51 | N/A | + | + | + |
| A24 | N/A | N/A | N/A | N/A | F52 | N/A | N/A | N/A | N/A |
| T25 | + | + | + | + | T53 | N/A | + | + | + |
| A26 | N/A | N/A | N/A | N/A | V54 | N/A | N/A | N/A | + |
| E27 | N/A | N/A | + | + | T55 | N/A | N/A | N/A | N/A |
| K28 | N/A | N/A | N/A | N/A | E56 | N/A | + | N/A | + |
| V29 | N/A | N/A | N/A | N/A | W43sc | N/A | N/A | + | N/A |

Exchange

NOE

"Real" NOE

Dipolar exchange (green) is identified by the opposite phase of the cross-peaks (positive NOE, negative ROE, when diagonal is positive), whereas direct hydrogen exchange with solvent (yellow) gives rise to peaks of identical phase (positive in both spectra). NOE cross-peaks are denoted in red. N/A, not available.

The amides protons with direct NOE contact in both water and IL conditions were observed for residues D22, T25, T44, and T49. Additional direct NOEs were detected upon addition of IL, for T17 (this was in chemical exchange in water), A20 (no detected NOE but positive ROE in water), E27, D46, and T53. All of these residues participate in hydrogen bonds. Interestingly, the most affected backbone NH in terms of chemical shift by IL, A23, is here attributed to a signal of chemical exchange and the preceding residue, D22 to a direct NOE with water. V21 is in chemical exchange and A20 is in direct NOE. Our qualitative data, taking in account that a few direct NOEs could be false positives, suggests that A23 and their vicinity are in close proximity to water molecules, and those can be favourably displaced by a larger and strongly hydrated anion, such as $Glu^-$ (as discussed above for the presence of a phosphate anion in the crystal structure[68]). In agreement with this hypothesis and using a similar approach, Clore and Gronenborn[110] identified a tightly bound water with A20 in the solution structure of GB1 (298.2 K and pH 4.3). This long-lived water molecule (> 1 ns) participates in the hydrogen-bonded pair of the amide of A20 with the carbonyl of M1 (likely the amide proton is involved in a 3-center hydrogen bond) and may contribute for the β-sheet stability at the N-terminus. The authors also indicate other bound water that participates in the Y33(NH)-V29(CO) hydrogen bond pair, but this was not detected in our experiments.

**Backbone relaxation**

The effects of [Ch][Glu] on the dynamics of GB1 were investigated by backbone $^{15}N$ spin longitudinal relaxation ($R_1 = 1/T_1$), transverse ($R_2 = 1/T_2$), and steady state $\{^1H\}$–$^{15}N$ NOE measurements[111] at two temperatures (298.2 and 310.2 K) and two magnetic fields (500.34 and 600.13 MHz). A full list of the relaxation dataset acquired at both temperatures can be found in Table A7 and A8 of Appendix A and the values for all 20 conditions (at both temperatures and field strengths) were plotted as a function of residue number in Fig. 2.31. The mean values are summarised in the Table 2.7. As expected, residues at the N-terminus, turn 1 and G41 in loop 3 show lower-than-average $R_1$, $R_2$ and NOE values (less prominent for $R_1$ data), indicative of more mobile segments (Fig. 2.31a-c). These trends are consistent with several relaxation studies performed on the backbone amide groups of GB1 (in buffer conditions and at different temperatures on two or more magnetic field strengths)[90,112–115].

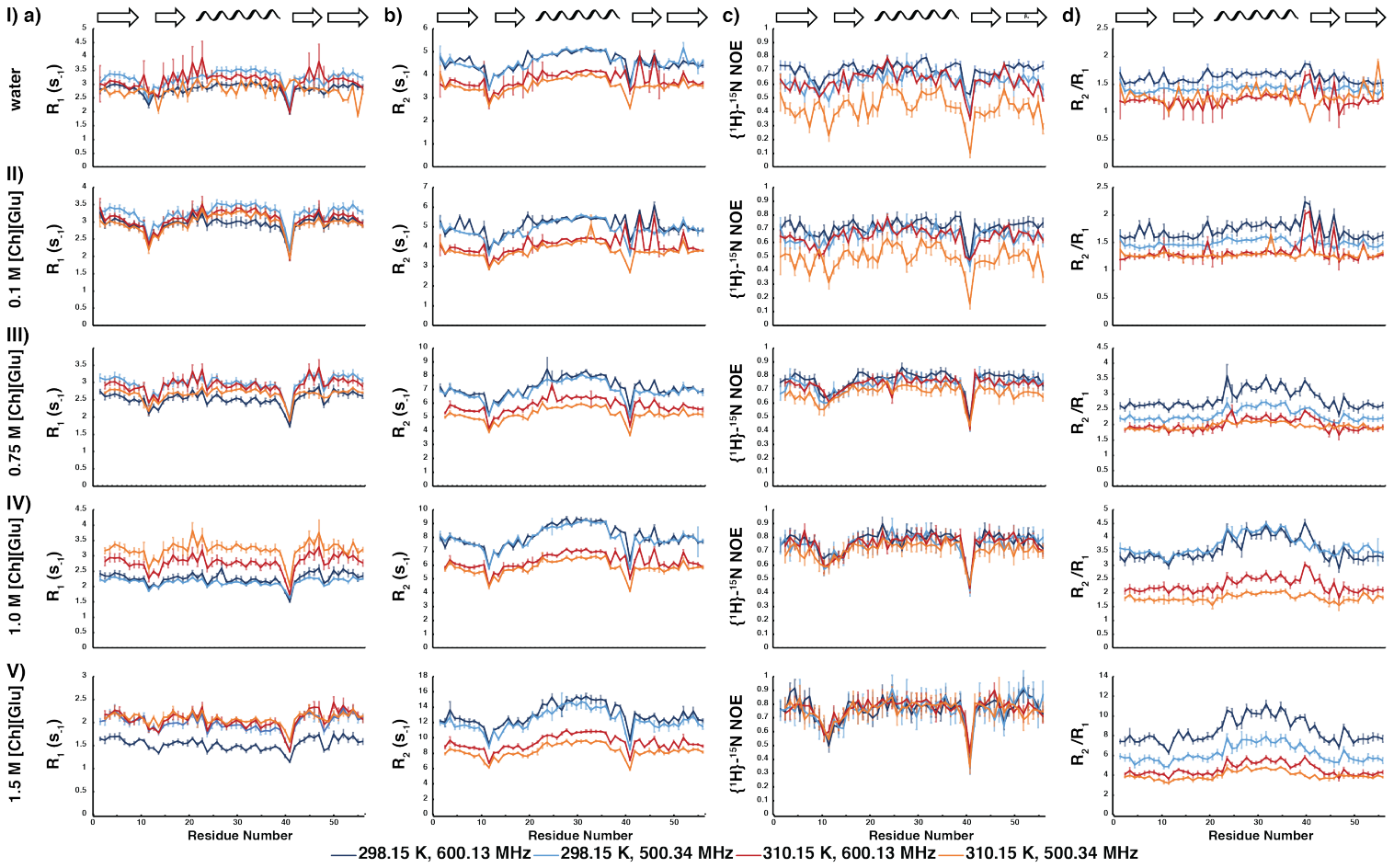



**Figure 2.31. Relaxation rates and $^{1}H$, $^{15}N$ NOE values determined for GB1 with [Ch][Glu].**

**a**) $R_1$ and **b**) $R_2$ relaxation rates, as well as **c**) $^{1}H$, $^{15}N$ heteronuclear NOE and **d**) $R_2/R_1$ values for GB1 in **I**) water, **II**) 0.1 M, **III**) 0.75 M, **IV**) 1.0 M and **V**) 1.5 M of [Ch][Glu] (pH 7.2 ± 0.2) measured at 298.2 K (blue) or 310.2 K (red) of temperature, and 600.13 MHz (dark colour) or 500.34 MHz (light colour) of magnetic field, respectively. The values were plotted as a function of the residue number and the error bars corresponds to the uncertainty in the parameter from the fitting (see material and methods).

Assuming that the contribution from internal fast motions and chemical exchange processes toward relaxation is insignificant, the $R_2/R_1$ ratio is commonly used as first estimation of global rotational correlation time[111,116]. In order to compare all conditions, in this analysis, no exclusion of residues were performed (as those with large amplitude fast internal motions reflected by NOE values lower than 0.65 or exhibiting significant conformational exchange on the $\mu$s/ms timescale[116]. When compared to water, GB1 in [Ch][Glu] shows higher $R_2/R_1$ values resultant from lower $R_1$ values and concomitantly higher $R_2$ values, the {$^1$H}-$^{15}$N NOE values are slightly increased or unaffected (Table 2.7).

**Table 2.7. Average parameters describing global dynamics of GB1 in [Ch][Glu] at two temperatures and two magnetic fields.**

| | 298.15 K | | | | 310.15 K | | | |
|---|---|---|---|---|---|---|---|---|
| | 600.13 MHz | | 500.34 MHz | | 600.13 MHz | | 500.34 MHz | |
| **[ChGlu] (M)** | **$R_1$ ($s^{-1}$)** | SD | **$R_1$ ($s^{-1}$)** | SD | **$R_1$ ($s^{-1}$)** | SD | **$R_1$ ($s^{-1}$)** | SD |
| **0.00** | **2.8** | 0.2 | **3.2** | 0.3 | **3.1** | 0.3 | **2.8** | 0.3 |
| **0.10** | **2.9** | 0.2 | **3.3** | 0.3 | **3.1** | 0.3 | **3.0** | 0.3 |
| **0.75** | **2.5** | 0.2 | **3.0** | 0.2 | **2.9** | 0.3 | **2.7** | 0.2 |
| **1.00** | **2.3** | 0.2 | **2.1** | 0.1 | **2.8** | 0.3 | **3.2** | 0.3 |
| **1.50** | **1.5** | 0.1 | **2.0** | 0.2 | **2.1** | 0.2 | **2.1** | 0.1 |
| **[ChGlu] (M)** | **$R_2$ ($s^{-1}$)** | SD | **$R_2$ ($s^{-1}$)** | SD | **$R_2$ ($s^{-1}$)** | SD | **$R_2$ ($s^{-1}$)** | SD |
| **0.00** | **4.6** | 0.4 | **4.6** | 0.4 | **3.8** | 0.4 | **3.5** | 0.3 |
| **0.10** | **5.1** | 0.4 | **4.9** | 0.4 | **4.1** | 0.5 | **3.8** | 0.4 |
| **0.75** | **7.1** | 0.7 | **6.9** | 0.7 | **5.8** | 0.5 | **5.2** | 0.5 |
| **1.00** | **8.1** | 0.8 | **8.0** | 0.8 | **6.3** | 0.6 | **5.8** | 0.5 |
| **1.50** | **13** | 1 | **12** | 1 | **9** | 1 | **8.3** | 0.8 |
| **[ChGlu] (M)** | **hetNOE** | SD | **hetNOE** | SD | **hetNOE** | SD | **hetNOE** | SD |
| **0.00** | **0.70** | 0.05 | **0.62** | 0.06 | **0.62** | 0.08 | **0.43** | 0.09 |
| **0.10** | **0.71** | 0.05 | **0.65** | 0.06 | **0.65** | 0.06 | **0.49** | 0.08 |
| **0.75** | **0.77** | 0.06 | **0.73** | 0.06 | **0.74** | 0.06 | **0.67** | 0.06 |
| **1.00** | **0.77** | 0.07 | **0.75** | 0.08 | **0.76** | 0.08 | **0.71** | 0.06 |
| **1.50** | **0.8** | 0.1 | **0.8** | 0.1 | **0.77** | 0.07 | **0.75** | 0.08 |
| **[ChGlu] (M)** | **$R_2/R_1$** | SD | **$R_2/R_1$** | SD | **$R_2/R_1$** | SD | **$R_2/R_1$** | SD |
| **0.00** | **1.62** | 0.09 | **1.42** | 0.06 | **1.2** | 0.1 | **1.3** | 0.2 |
| **0.10** | **1.7** | 0.2 | **1.49** | 0.06 | **1.3** | 0.2 | **1.28** | 0.07 |
| **0.75** | **2.8** | 0.3 | **2.3** | 0.2 | **2.0** | 0.2 | **2.0** | 0.1 |
| **1.00** | **3.6** | 0.4 | **3.7** | 0.4 | **2.3** | 0.3 | **1.8** | 0.1 |
| **1.50** | **8** | 1 | **6.2** | 0.9 | **4.6** | 0.6 | **4.0** | 0.4 |

Mean of $R_1$, $R_2$ relaxation rates, hetNOE and $R_2$/ $R_1$ values for [Ch][Glu] titration measured at 298.2 K or 310.15 K, each on a 600.13 MHz or 500.34 MHz spectrometer. SD, standard deviation from the average.

For simplicity in this section, the $R_1$, $R_2$, {$^1$H}–$^{15}$N NOE, and $R_2/R_1$ ratio values of GB1 determined at 298 K and 600 MHz are highlighted in Fig. 2.32. The decrease in $R_1$ and the systematic increase in $R_2$ are expected due to the significant increase of the macro-viscosity ($\eta_{macro}$) of the solution (Table 2.8). Because the {$^1$H}–$^{15}$N NOE values are mostly unaffected,

the differences in viscosity are directly reflected in the increase of $R_2$ values as well as the overall rotation correlation time ($R_2/R_1$ ratio) of GB1. However, this only partially explains the slow molecular tumbling time of GB1 at 1.5 M of [Ch][Glu] (8 ± 1 ns, that corresponds to a molecular weight of ~ 16 kDa[117]) as an increase in the hydrodynamic radius of GB1 was revealed by an independent measurement (discussed in the following section of protein diffusion).

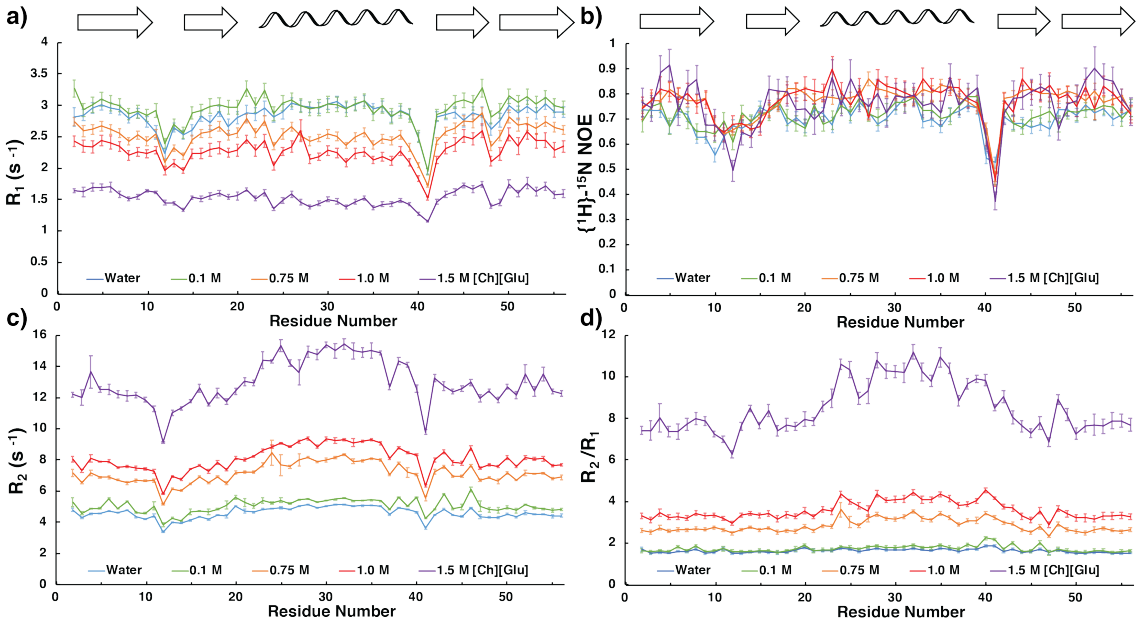



**Figure 2.32. $^{15}$N backbone relaxation data measured for GB1 in water and aqueous-[Ch][Glu] IL for a proton frequency of 600.13 MHz at 298.2 K.**

Relaxation parameters ($R_1$, $R_2$ and hetNOE) and $R_2/R_1$ ratio plotted versus residue number. Data acquired for water (blue), and for aqueous-[Ch][Glu] IL at 0.1 M (green), 0.75 M (orange), 1.0 M (red) and 1.5 M (purple) of concentration (pH 7.2 ± 0.2). The positions of β-strands and the α-helix are indicated schematically at the top.

**Table 2.8: Average parameters describing global dynamics of GB1.**

| | GB1 backbone relaxation [a] | | | | Solution viscosity [b] |
|---|---|---|---|---|---|
| **[Ch][Glu] (M)** | $R_1$ (s$^{-1}$) | $R_2$ (s$^{-1}$) | **HetNOE** | $R_2/R_1$ | $\eta_{macro}$ (cP) |
| **0.0** | 2.8 ± 0.2 | 4.6 ± 0.4 | 0.7 ± 0.1 | 1.62 ± 0.09 | 0.89* |
| **0.1** | 2.9 ± 0.2 | 5.1 ± 0.4 | 0.7 ± 0.1 | 1.7 ± 0.2 | 1.113 ± 0.006 |
| **0.75** | 2.5 ± 0.2 | 7.1 ± 0.7 | 0.8 ± 0.1 | 2.8 ± 0.3 | 1.74 ± 0.01** |
| **1.0** | 2.3 ± 0.2 | 8.1 ± 0.8 | 0.8 ± 0.1 | 3.6 ± 0.4 | 2.04 ± 0.01 |
| **1.5** | 1.5 ± 0.1 | 13 ± 1 | 0.8 ± 0.1 | 8 ± 1 | 3.26 ± 0.06 |

**a**) Mean of $^{15}$N Relaxation data ($R_1$, $R_2$ and HetNOE) for [Ch][Glu] titration (pH 7.2 ± 0.2) acquired at 600.13 MHz and 298.2 K. The uncertainties are the standard deviation from the average. **b**) Measured viscosities for aqueous-[Ch][Glu] solutions. The uncertainties are the standard deviation from the triplicate dataset.

*Viscosity of pure water at 298.2 K[118]. **Viscosity of 0.75 M [Ch][Glu] extrapolated from the data, the dependency of solvent viscosity as a function of IL concentration can be described by an empirical exponential equation (no physical meaning), η (298.2 K, [Ch][Glu]) = 0.946491exp(0.809821[Ch][Glu]).

Interestingly, while at low [Ch][Glu] concentrations there is little difference between the average $R_2/R_1$ ratios for the protein's secondary structure elements, at higher concentrations I observe a clear differentiation between those of the α-helix and those of the remaining residues (including the ones is the β-sheets) (Table 2.9). This is most evident at 1.5 M [Ch][Glu], where the $R_2/R_1$ ratios for the α-helix are ~25% higher than those of the remaining residues. This trend seems to be independent of the temperature or the magnetic field (Fig. 2.33).

**Table 2.9. Average $R_2/R_1$ ratio for secondary structure regions of GB1.**

| | $R_2/R_1$ | | | |
|---|---|---|---|---|
| **[Ch][Glu] (M)** | **β-sheet** | **α-helix** | **Other residues** | **All residues** |
| **0.0** | 1.58 ± 0.07 | **1.67 ± 0.05** | 1.6 ± 0.1 | 1.62 ± 0.09 |
| **0.1** | 1.7 ± 0.1 | **1.78 ± 0.07** | 1.8 ± 0.2 | 1.7 ± 0.2 |
| **0.75** | 2.7 ± 0.1 | **3.2 ± 0.2** | 2.8 ± 0.3 | 2.8 ± 0.3 |
| **1.0** | 3.4 ± 0.2 | **4.0 ± 0.3** | 3.5 ± 0.5 | 3.6 ± 0.4 |
| **1.5** | 7.9 ± 0.6 | **10.0 ± 0.8** | 8 ± 1 | 8 ± 1 |

Average values of the $R_2/R_1$ ratio for each secondary structure segment in water and presence of [Ch][Glu], at 600.13 MHz and 298.2 K. The uncertainties are the standard deviation from the average.

The higher $R_2$ values for the α-helix occur primarily because all its NH groups are aligned relatively closely with the long axis of the molecule, resulting in slower reorientation of these NH vectors compared to most other NH groups in the molecule, as observed by cross-relaxation measurements for GB1 in aqueous-buffer solution at low temperatures[112]. Nevertheless, $R_2$ values can be influenced by chemical or conformational exchange on the microsecond to millisecond timescale, which could increase them due to an exchange broadening ($R_{ex}$) contribution[90].

The $R_2$ profile along protein sequence (where values for the α-helix are evidently higher than those of the remaining residues, particularly in the presence of 1.5 M [Ch][Glu] - Fig. 2.33) are likely a result of both contributions: (1) from orientation of the NH vector along the long axis of the molecule and (2) from exchange broadening. The slower reorientation is resulting from an increase in the macro-viscosity of the solution (at 1.5 M [Ch][Glu] and 298 K, the solution is 3.7 times more viscous), which is similar to a drop in temperature[112]. Further, the intensification of $R_{ex}$ terms in the helical segment are in line with the perturbations already discussed for the α-helix N-*terminus* (changes in chemical shift, temperature coefficient, and exchange rate). According to the model-free analysis of the

relaxation data (as discussed in detail below), the presence of anisotropic diffusion as well as chemical exchange should be true.

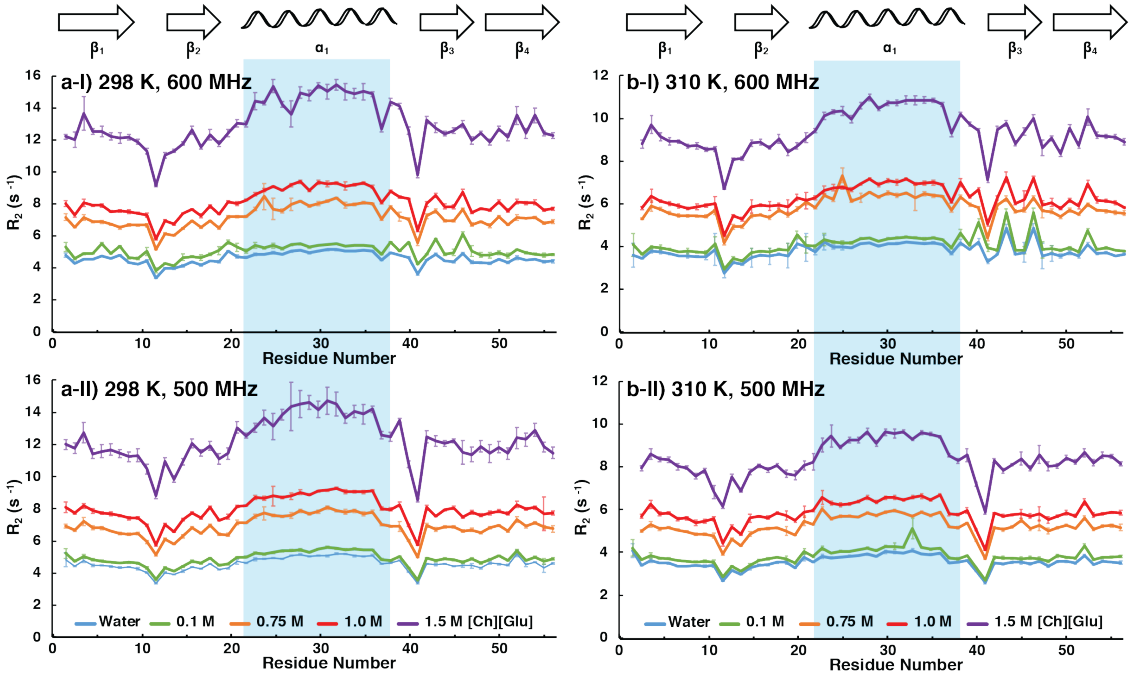



**Figure 2.33: $^{15}$N backbone relaxation $R_2$ data measured for GB1 in water and aqueous-[Ch][Glu] IL.**

$R_2$ relaxation rates determined for water (blue), and aqueous-[Ch][Glu] IL at 0.1 M (green), 0.75 M (orange), 1.0 M (red) and 1.5 M (purple) of concentration, for (**I**) proton frequency of 600.13 MHz at **a**) 298.2 K and **b**) 310.2 K, and for (**II**) 500.34 MHz also at both temperatures. Above each plot it is depicted the secondary structure of the protein.

Protein diffusion

While rotational diffusion or correlation time ($\tau_m$) has been assessed from the $^{15}$N $R_2/R_1$ ratio, an independent measurement of the hydrodynamic radius of the protein is of considerable importance. Diffusion ($D$) is quantified as diffusion coefficients that are dependent on friction factors ($f_T$), related to the molecular size and the viscosity of the solution: $D = k_B T/f_T$, where T is the absolute temperature and $k_B$ the Boltzmann constant. For a spherical molecule of hydrodynamic radius $r_H$ in a solvent of viscosity η, Stokes' law gives $f_T = 6\pi\eta r_H$. Thus, translational diffusion can be described in terms of the Stokes–Einstein equation ($D_t = k_B T/(6\pi\eta r_H)$). In the case of rotational diffusion, the friction is $f_T = 8\pi\eta r_H^3$, and the rotational diffusion constant is obtained by the Stokes-Einstein Debye relation (($D_r = k_B T/(8\pi\eta r_H^3)$).

The translational self-diffusion coefficients were measured for protons from water and $Ch^+$/$Glu^-$ molecules (Fig. 2.34a), as well as for the amide proton envelope of GB1 (Fig. 2.34b), for aqueous-solutions in the absence or presence of [Ch][Glu] (0.1, 0.5, 0.75, 1.0 and 1.5 M).

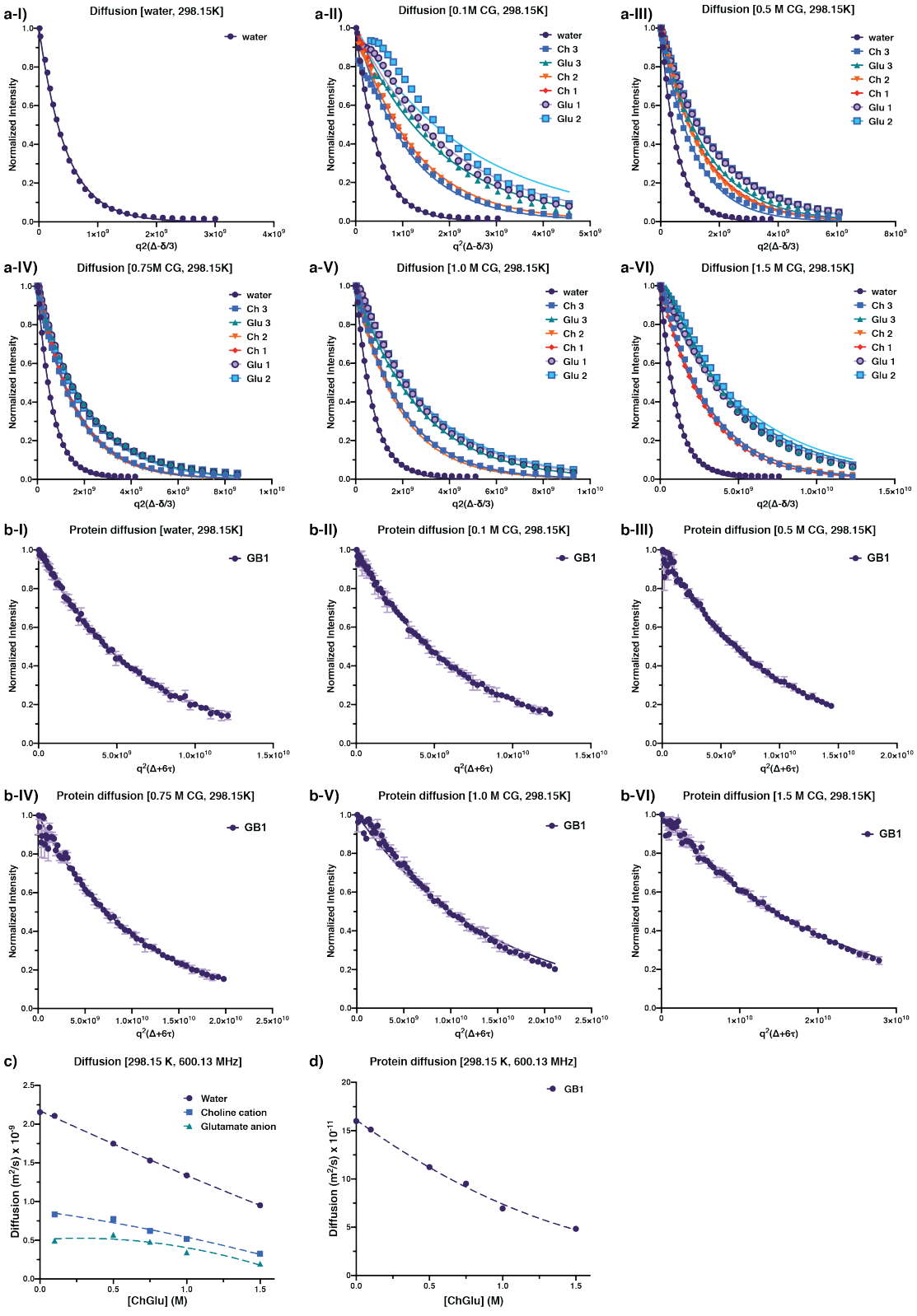



**Figure 2.34. Diffusion constants for water, $Ch^+$/$Glu^-$ ions and GB1 in aqueous-[Ch][Glu] solutions.**

Decay of integrated signals ($I/I_0$) as a function of effective gradient strength for **a**) water and ions protons and **b**) mean (and standard deviation) of 3-5 integrated sections of $^{15}N$ filtered $^{1}H$ amide envelope, in **I**)

water and **II-VI**) 0.1, 0.5, 0.75, 1.0 and 1.5 M [Ch][Glu] solutions. In general, the signal intensity (I) in a diffusion experiment depends on the gradient strength (*G*) and the diffusion coefficient (*D*) of the probe molecule as $I/I0 = \exp(-cDG^2)$, where *c* is a constant that includes the gradient duration and the diffusion delay. The variation in I with the square of the gradient strength was fit to a single exponential decay equation to extract *D*. The values obtained are shown as a function of IL concentration for **c**) water, choline, and glutamate molecules, and for **d**) GB1 protein. The dashed line is intended only to guide the eye and does not have any physical significance.

Standard pulse field gradient (PFG) NMR techniques were used to study the diffusion of water and IL protons, while for protein diffusion, an adapted version of $^{15}$N-edited heteronuclear stimulated echo (X-STE) experiment[119] was employed (see details in Material and Methods). Here, the use of $^{15}$N provides isotope filtration and its long spin-lattice relaxation times allows one to increase the duration of the diffusion delay.

By means of the Stokes-Einstein equation, the obtained $r_H$ value for GB1 in pure water at 298.2 K ($r_{H,GB1,water}$ = 15.33 Å) match well with their prediction ($r_H$ = 15.26 Å = 4.75 $N^{0.29}$ Å, where N is the number of residues in the polypeptide chain, 56)[120]. However, this determination relies on the interpretation of absolute values of diffusion coefficients which are subject to errors from experimental inaccuracies (e.g., exact shape and strength of the gradient pulse). There is also the need for extra experiments to measure the solution viscosity. In order to avoid these problems, particularly those that arise from variations of sample conditions (e.g., viscosity and temperature), some small molecules have been used as internal radius or viscosity standards, such as dioxane[120,121] or tetramethylsilane (TMS)[122]. To circumvent addition of external molecules, the water resonance was used as internal probe, $r_H{}^{ref}{}_{water}$ = 1.12 ± 0.01 Å[123,124]. Changes in GB1's effective hydrodynamic radius ($r_{H,protein}$) upon addition of IL can then be determined from the ratio of its diffusion constant ($D_{protein}$) to that of water ($D^{ref}$), independently of the viscosity,

$$\frac{D_{water}^{ref}}{D_{protein}} = \frac{r_{H,protein}}{r_{H,water}^{ref}}. \qquad [2.2]$$

The results confirmed the increase of the hydrodynamic radius of GB1 with IL concentration (Table 2.10), from $r_{H,protein}$ = 15.1 ± 0.2 Å in the absence of cosolute to $r_{H,protein}$ = 22.0 ± 0.3 Å in the presence of 1.5 M [Ch][Glu]. This indicates that the effect on protein compaction is dependent on IL concentration. Still, the self-diffusion coefficient of water varies considerably along the IL titration due to the differences in viscosity (Fig. 2.34 and Table 2.10).

**Table 2.10. Hydrodynamic radii of GB1 upon addition of [Ch][Glu].**

| [Ch][Glu] (M) | Self-diffusion coefficients, $D$ ($\times 10^{-10}$ m$^2$ s$^{-1}$) | | | |
|---|---|---|---|---|
| | Water | GB1 | $D^{ref}_{water}/D_{protein}$ | $r_{H,protein}$ (Å) |
| **0** | 21.5 ± 0.3 | 1.600 ± 0.007 | 13.5 ± 0.2 | 15.1 ± 0.2 |
| **0.1** | 21.1 ± 0.3 | 1.512 ± 0.009 | 13.9 ± 0.2 | 15.6 ± 0.3 |
| **0.5** | 17.5 ± 0.2 | 1.123 ± 0.009 | 15.6 ± 0.2 | 17.4 ± 0.3 |
| **0.75** | 15.3 ± 0.2 | 0.952 ± 0.008 | 16.1 ± 0.2 | 18.0 ± 0.3 |
| **1.0** | 13.4 ± 0.1 | 0.649 ± 0.006 | 19.3 ± 0.3 | 21.6 ± 0.4 |
| **1.5** | 9.5 ± 0.1 | 0.484 ± 0.003 | 19.7 ± 0.2 | 22.0 ± 0.3 |

Diffusion coefficients of water protons and amide protons of GB1 (by integration of 3-5 sections of the NH envelope) as extracted from diffusion experiments. The uncertainty corresponds to the fitting error. It is assumed that hydrodynamic radius of water is constant, $r_H{}^{ref}{}_{water} = 1.12 \pm 0.01$ Å[123,124], and any changes in the ratio of Eq. 2.2 may be attributed to modifications in the GB1's hydrodynamic radius.

Using the same approach (Eq. 2.2), and assuming $r_H{}^{ref}{}_{water}$ as constant, the hydrodynamic radii of ions were determined (Table 2.11). The values are constant and expected accordingly with its molecular weight ($Ch^+$ ~ 2.9 Å and $Glu^-$ ~ 4.3 Å). Although a decrease of cation/ anion diffusion coefficients along the [Ch][Glu] titration is expected due to the corresponding increase of the macro-viscosity, the change is not linear and reveals an approximation of those coefficients around 0.75 M of IL concentration (i.e., the cation/anion ratio $D_{Ch}/D_{Glu} = 1.29 \pm 0.05$ is the smallest when compared to lower or higher IL concentrations – Table 2.11). This could be indicative of an ion cluster or even a (transient) ion-pair formation[125] with possible repercussions on protein stability (destabilisation observed for GB1 at 0.75 M of IL).

**Table 2.11. Hydrodynamic radii of [Ch][Glu] ions along titration.**

| [Ch][Glu] (M) | Self-diffusion coefficients, $D$ ($\times 10^{-10}$ m$^2$ s$^{-1}$) | | | $D_{Ch}/D_{Glu}$ | $D^{ref}_{water}/D_{Ch}$ | $D^{ref}_{water}/D_{Glu}$ | $r_{H,Ch}$ (Å) | $r_{H,Glu}$ (Å) |
|---|---|---|---|---|---|---|---|---|
| | Water | $[Ch]^+$ | $[Glu]^-$ | | | | | |
| **0** | 21.5 ± 0.3 | - | - | - | - | - | - | - |
| **0.1** | 21.1 ± 0.3 | 8.3 ± 0.2 | 5.0 ± 0.8 | 1.7 ± 0.3 | 2.5 ± 0.2 | 4 ± 1 | 2.8 ± 0.2 | 4.8 ± 0.8 |
| **0.5** | 17.5 ± 0.2 | 8 ± 1 | 5.7 ± 0.6 | 1.4 ± 0.2 | 2.3 ± 0.3 | 3.1 ± 0.3 | 2.5 ± 0.3 | 3.4 ± 0.4 |
| **0.75** | 15.3 ± 0.2 | 6.2 ± 0.2 | 4.8 ± 0.1 | 1.29 ± 0.05 | 2.47 ± 0.07 | 3.18 ± 0.09 | 2.77 ± 0.09 | 3.6 ± 0.1 |
| **1.0** | 13.4 ± 0.1 | 5.2 ± 0.2 | 3.4 ± 0.2 | 1.5 ± 0.1 | 2.6 ± 0.1 | 3.9 ± 0.3 | 2.9 ± 0.1 | 4.3 ± 0.3 |
| **1.5** | 9.5 ± 0.1 | 3.3 ± 0.1 | 2.0 ± 0.1 | 1.6 ± 0.1 | 2.9 ± 0.1 | 4.8 ± 0.3 | 3.3 ± 0.1 | 5.4 ± 0.3 |

Diffusion coefficients of water and $Ch^+$, $Glu^-$ protons as extracted from diffusion experiments. As a rule of thumb, a strong ion-pairing is verified when the ratio $D_{cation}/D_{anion} = 1$[125]. The coefficients obtained for the ions are the average of the corresponding protons (integration of three resonances for each molecule). The

uncertainty in water diffusion corresponds to the fitting error and uncertainties in ions diffusions are the standard deviation of the mean. It is assumed that hydrodynamic radius of water is constant, $r_H^{ref}{}_{water} = 1.12 \pm 0.01$ Å[123,124], and any changes in the ratio of Eq. 2.2 may be attributed to modifications in the hydrodynamic radius of the molecule under study.

Concerning the GB1 diffusion, the relationships above used are commonly applied to spherical proteins in homogeneous solutions, in which it is assumed that the protein is much larger than the cosolute or viscogen added, such as ethylene glycol or glycerol ($r_{H,\text{protein}} \gg r_{H,\text{viscogen}}$). In this case, the protein senses an effective viscosity ($\eta^*$) that is independent of its size, corresponding to the macroscopic solution viscosity, as measured from flow rates using viscometers ($\eta^* = \eta_{\text{macro}}$). By contrast, if $r_{H,\text{protein}} \leq r_{H,\text{viscogen}}$, such as sucrose, polymers or a protein, the effective viscosity is reduced from $\eta_{\text{macro}}$, so-called micro-viscosity (or nano-viscosity) and is dependent upon the exact value of $r_{H,\text{protein}} / r_{H,\text{viscogen}}$. This size dependence of the effective viscosity and effects on translational diffusion is well-documented for several viscogens[124,126,127]. Therefore, high concentrations of macromolecules are expected to cause deviations from the Stokes laws as the macromolecules approach the size of the test protein[128,129].

The increase in effective viscosity ($\eta^*$) resulting from addition of [Ch][Glu] (in analogy to a conventional viscogen), normalized to the viscosity in the absence of IL ($\eta_{0,\text{water}} = 0.89$ cP), as reported for protein GB1 (or another probe), can be extracted from

$$\frac{\eta^*_{IL}}{\eta_{0,water}} = \frac{D_{0,protein,water}}{D_{protein,IL}}, \qquad [2.3]$$

where $D_0$ is the diffusion coefficient of the GB1 in pure water of viscosity $\eta_0$. Although this analysis seems redundant because the size of GB1 is much larger than the ions present in the solution (even if they share the hydrodynamic radius), the comparison of micro- and macro-viscosity at each condition points out extremely weak interactions that affect the protein diffusion (Fig. 2.35). The negative deviation is very small (Fig. 2.35a) but indicates that increased macro-viscosity decreases diffusion less than predicted by the Stokes laws, and the increase in the size of GB1 (Fig. 2.35b) could be interpreted as preferential hydration of their native state (that is, the cosolute is excluded from the immediate surface of the protein molecule). Despite a significative increase in $\eta^*$, similar $r_H \approx 22$ Å have been found for GB1 between 1.0 M and 1.5 M of IL.

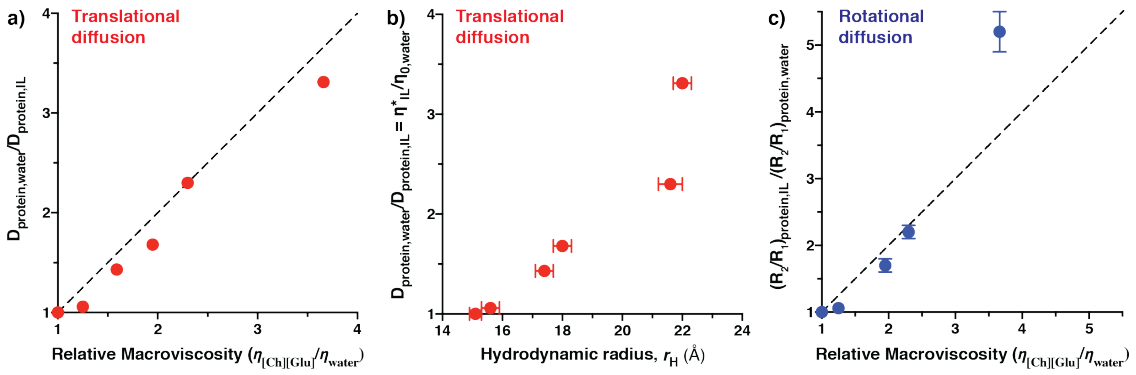



**Figure 2.35. Effective viscosity during translation and rotation of GB1 with [Ch][Glu].**

**a**) Ratio of translational diffusion coefficients of GB1 in water to its diffusion coefficients in [Ch][Glu] (298.2 K, pH 7.2), as function of relative macroviscosity (measured in viscometer). The value in the absence of IL is $\eta_{0,water}$ = 0.89 cP. The dashed line illustrates the unitary slope and origin-intercept expected for Stokes-Einstein law. The uncertainties are smaller than the symbols. **b**) Values of $\eta_{IL}^*/\ \eta_{0,water}$ corresponds to $D_{protein,water}/\ D_{protein,IL}$ and are plotted as a function of $r_H$, as described in the text. All values are based on measurements obtained using GB1 as probe and [Ch][Glu] as a viscogen (0.1, 0.5, 0.75, 1.0 and 1.5 M). Uncertainties were propagated from the standard error in the coefficients. **c**) Ratio of the mean of $R_2/R_1$ in IL to that in water (inversely proportional to rotational diffusion) as function of relative macroviscosity. The uncertainties correspond to the standard deviation of the average. The dashed line illustrates the Stokes-Einstein Debye relationship. For example, in glycerol both rotational and translational diffusion follow the Stokes laws for a 7.4-kDa protein CI2[128]. Points below and above dashed line indicate negative and positive deviations, respectively.

Using an initial estimate of rotational $\tau_m$ based on the average $R_2/R_1$ (Table 2.5) which is inversely proportional to rotational diffusion ($1/D_r$), I find that [Ch][Glu] has a dramatically and different effect on GB1 diffusion for the particular IL concentration of 1.5 M - Fig. 2.35c (below that concentration, the ratios are similar and expected for Stokes laws). For this condition, the rotational motion is severely impeded (more than predicted and more than translational diffusion). This is in line with weak favourable interactions between $Ch^+/Glu^-$ ions and/or ion-pair and GB1 with repercussions in hydration since that should lead to larger effects on rotation compared to translation (because a rotating sphere depends on volume, $r^3$, while a translating sphere depends only on size, r).

Solvent friction/ nanoscale viscosity

The differential effects on $R_1$ and $R_2$ along the sequence of GB1 suggest the existence of different time scales (Fig. 2.31-33), with the slowest observed for the $\alpha$-helix segment. On the basis that time scales of intermediate and slow backbone motions respond very differently to changes in viscosity, resulting in distinct ranges of friction coefficients ($\varepsilon$), Adamski et al[130] propose a general expression that describes these different rotational correlation times as function of both temperature and viscosity,

$$\tau_k(C,T) = \tau'_{k,\infty}(\varepsilon_k \rho(C) + 1) \exp\left(\frac{E_{a,k}}{RT}\right), \qquad [2.4]$$

where $\tau_k$ is the correlation time of motion, $\varepsilon_k$ is a friction coefficient, relative to solvent friction measured at a given viscogen concentration C, $\rho$(C), and $E_{a,k}$ is the activation energy of intermediate and slow motions. $\tau'_{k,\infty}$ is analogous to a timescale and in abstract terms corresponds to the correlation time at infinite dilution and infinite temperature. The dependence of each correlation time on viscosity can be understood by the relationship of effective viscosity (see above) and the dimensions of the protein (probe) and viscogen. The authors testified the robustness of Eq. 2.4 by accurately predictive relaxation rates for an IDP of interest under any given conditions since the intrinsic solvent friction can be estimated from experimental measurements.

In the same [Ch][Glu]-samples used to acquire relaxation data sets at 298 K (as well as protein diffusion and CLEANEX data), the evolution of water proton intensity was monitored as a function of recovery time (Fig. 2.36), and the solvent friction (or nanoscale

viscosity) was estimated by measuring the $^1$H longitudinal relaxation rates of water, $R_{1,water}$ (Fig. 2.37a). In this experiment, it is assumed, at high magnetic fields, that rotation diffusion dominates $R_1$, and relaxation dispersion effects are insignificant (e.g., exchange between free water from the bulk and bound-water molecules or translational diffusion effects at the protein surface)[131,132].

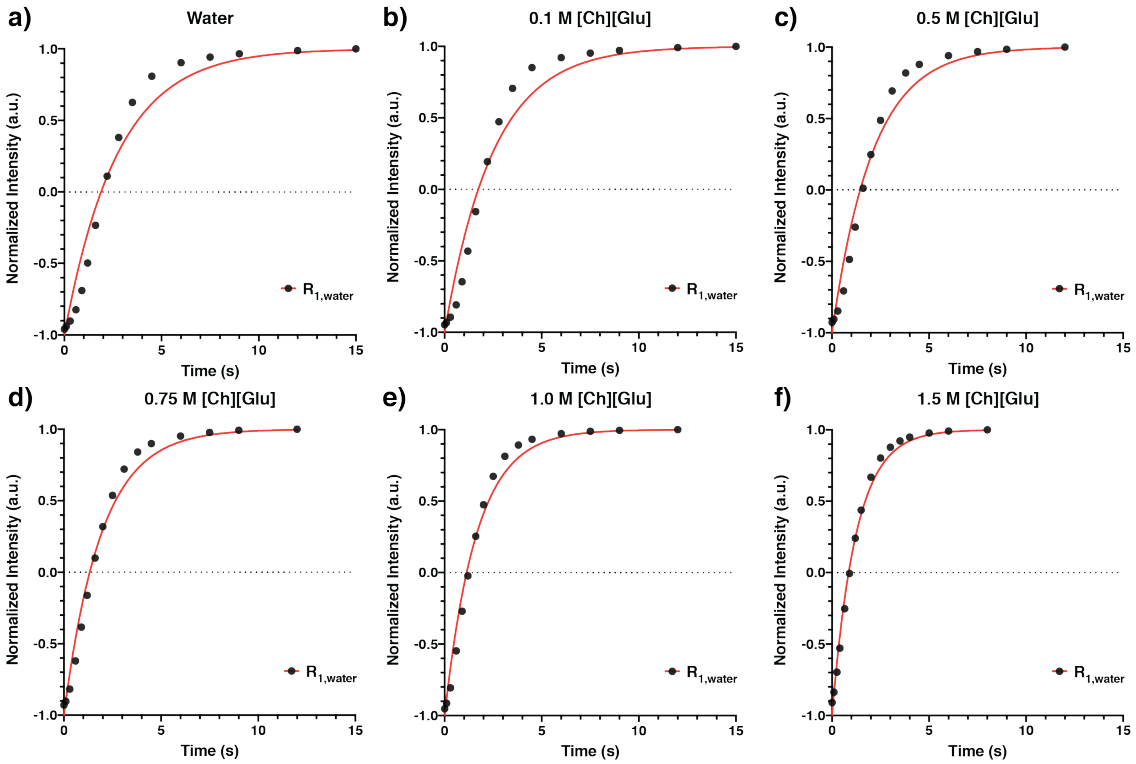



**Figure 2.36. Evolution of water proton intensity during saturation-inversion-recovery.**

Water proton intensity was monitored as a function of recovery time for **a**) water , and **b-f**) aqueous-[Ch][Glu] solutions, using an adapted version of saturation-inversion-recovery (SIR) sequence[133]. Data was fit to an exponential inversion-recovery (red lines), $I(t) = I_0 (1 - 2\exp(-R_{1,water}t)$.

Therefore, changes in viscosity due to IL addition can be estimated by the following expression for solvent friction,

$$\rho(IL) = \frac{\eta_{IL} - \eta_0}{\eta_0} = \frac{R_{1,IL} - R_{1,0}}{R_{1,0}}, \qquad [2.5]$$

where $R_{1,0}$ and $\eta_0$ are the longitudinal relaxation of water and the viscosity in the absence of IL, respectively, and $\eta_{IL}$ is the viscosity of the sample of interest at a given IL concentration. The measured values follow a linear dependence on IL concentration until 1.0 M. Beyond this concentration, the slope begins to deviate, in agreement with all viscosity measurements. In fact, values are significantly different only for 1.5 M of IL: $\eta = 3.26 \pm 0.06$ cP measured by a viscometer, and $\eta^* = 2.94 \pm 0.07$ cP or $\eta^* = 4.7 \pm 0.3$ cP (considering $\eta_0 = 0.89$ cP) if estimate by the ratios of translational or by rotational parameters, respectively. These are a good

indication of a change in [Ch][Glu] nanostructure and/or ions self-assembly[125], in analogy to dilute-semidilute transition regime in concentrated polymer solutions.

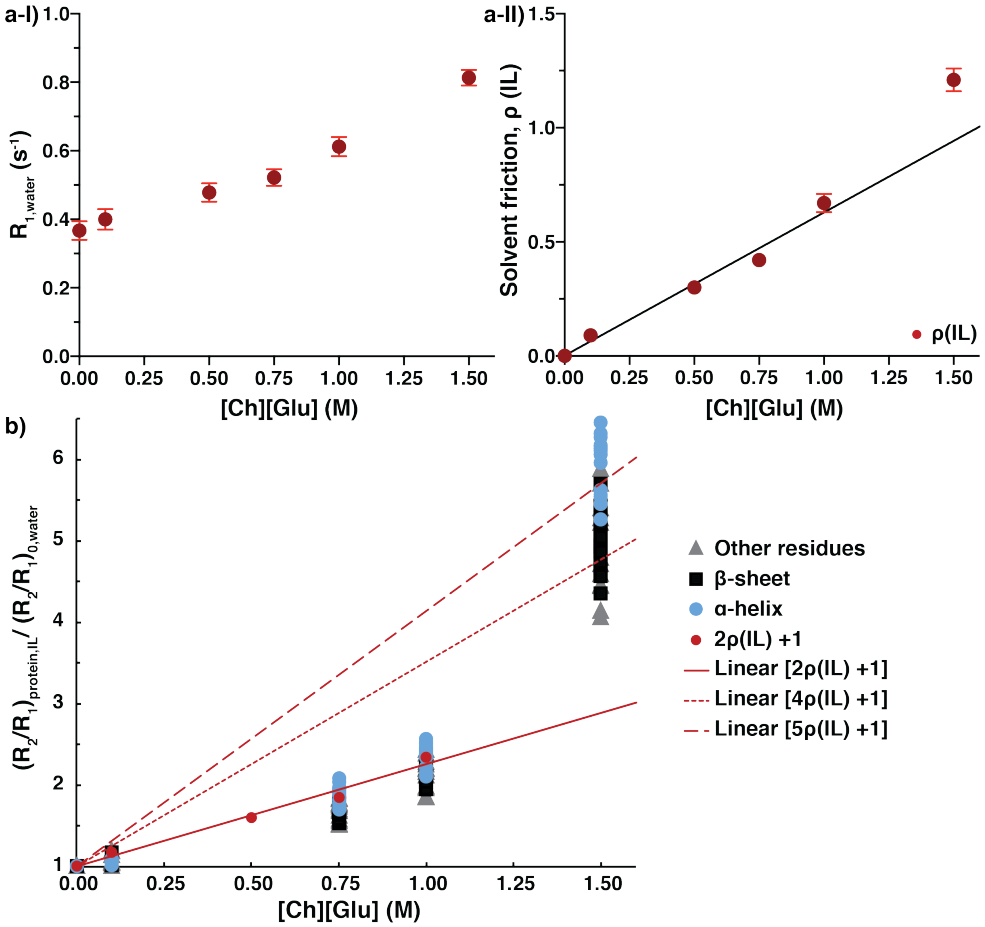



**Figure 2.37. Influence of [Ch][Glu] on dynamics of water protons and $^{15}$N backbone amide $R_2/R_1$ ratio.**

**a-I**) Longitudinal relaxation of water protons ($R_{1,\,water}$) for water and aqueous-[Ch][Glu] solutions, and **a-II**) corresponding values of solvent nanoscale viscosity / friction (black line indicates a linear dependence, from 0 to 1.0 M, whereas the slope corresponds to ϱ(IL) = 0.628). **b**) Relative changes of backbone amide $R_2/R_1$ ratio for secondary structure motifs (grey triangles, other residues; black squares, β-sheet; blue circles, α-helix) of GB1 recorded at 298.2 K and 600.13 MHz. Red line shows relative changes in dynamics of water measured by relaxation of water protons (red points and solid line: 2ϱ(C) + 1). Red square dotted line: 4ϱ(C) + 1 and red long dashed line: 5ϱ(C) + 1.

The dependence of $\tau_k$ on $\rho$(IL) measured at a given IL concentration can be interpreted by,

$$\tau_k(IL) = \tau_{k,\infty}[\varepsilon_k \rho(IL) + 1], \qquad [2.6]$$

where $\tau_{k,\infty}$ is the correlation time at infinite dilution. A strong dependence of the correlation time (estimate from $R_2/R_1$) on solvent friction was revealed for the helical segment (slowest component) at 1.5 M [Ch][Glu], while a weaker dependence was found for lower

concentrations without discrimination between secondary structure motions (Fig. 2.37b). Not surprisingly, relative changes in $^{15}$N backbone amide $R_2/R_1$ along IL titration are not explained by a simple model using a common ε. Using a value of $\varepsilon_k = 2$, the model explains the relaxation rates at 0.1, 0.75 and 1.0 M (in line with the change in their relative macroviscosity, Fig. 2.35c) but underestimates the relaxation rates that are dominated by slower motions at 1.5 M. Here, 4 and 5 times the intrinsic solvent friction are necessary to predict backbone $R_2/R_1$ measured for the β-sheet and $\alpha$-helix motif, respectively. Clearly, backbone relaxation of GB1 does not scale in a simple way with $\rho$(IL). The shift of the pattern of protein-IL interaction above 1.0 M is essentially experienced by the decrease of rotational diffusion.

Model-free

To better interpret the IL-induced changes in fast motions on ps-ns time frames and inherent μs-ms contributions from chemical exchange (or anisotropic effects) along the sequence of GB1, the relaxation data sets collected at the two fields strengths and two temperatures were analysed with the automated model-free protocol of Relax[134–136]. As mentioned in the introductory chapter (section of methods for ps-ns timescale dynamics), from the model-free analysis, fundamental parameters can be extracted: the generalized squared order parameter $S^2$, which describes the amplitude of each individual NH bond vector, providing a measure for the degree of restriction (from completely unrestricted, $S^2 = 0$ to fully rigid, $S^2 = 1$); and a characteristic and effective internal correlation time $\tau_e$, which defines motion on a time scale usually faster than the overall molecular tumbling ($\tau_m$) of the molecule. Relax also includes the extended model-free approach[137]which assumes that internal motions occur on two separable time scales (slow and fast), and the fitting of an exchange parameter $R_{ex}$, to take into account effects of slow motion on the μs-ms timescale due to chemical and/or conformational exchange[138]. A full list of the model-free parameters ($S^2$, $\tau_e$, $R_{ex}$, and the dynamic model used to fit the data) determined for 298.2 K and 310.2 K of temperature can be found in Table A9 and A10 of Appendix A, respectively.

The diffusion tensor model with best global fit, for almost all conditions, is represented by an ellipsoid (asymmetric or fully anisotropic). Spheroid diffusion tensor (axially symmetric anisotropic diffusion) was selected by relax for GB1 in water and 1.0 M at 298.2 K and 1.5 M at both temperatures, the class of spheroid is prolate with the exception of 1.0 M (298.2 K) which is oblate. Overall, slower tumbling ($\tau_m$) was observed with the increase of solvent viscosity (or decrease of temperature), as shown in Table 2.12. Deviation of the rotational correlation times from the linearity of Stokes–Einstein equations have been observed at 298 K by an initial estimate of $\tau_m$ based on the average $R_2/R_1$ (Fig. 2.35 and 2.37), however using the iteratively refined $\tau_m$, the relative values approximate to the relative bulk viscosity (the $\tau_m/\eta = 2.6 \pm 0.3$ ratios are approximately constant). On the other side, the values determined at 310.2 K reveal significant negative deviations for the high concentrations (i.e., the protein tumbles faster than predicted, as at 1.5 M and 310.2 K, $\tau_m/\tau_{m,0} = 2.2 \ll \eta/\eta_0 = 3.6$). This change in the protein tumbling to faster than expected by raising temperature from 25 to 37 ºC could reflect the difference in the solvent-cosolvent composition in the bulk solution and at the protein surface of the native state, as previously observed for protein CspB in ethylene glycol[139].

**Table 2.12. Parameters describing global dynamics of GB1 in the presence of IL.**

| [Ch][Glu] (M) | Temp. (K) | Global model | $\tau_m$ (ns) | Average order parameter, $S^2$ | | | |
|---|---|---|---|---|---|---|---|
| | | | | β-sheet | α-helix | Other residues | All residues |
| **0.0** | 298.2 | prolate | **2.37** | 0.93 ± 0.04 | **0.97 ± 0.02** | 0.86 ± 0.02 | 0.92 ± 0.07 |
| | 310.2 | ellipsoid | **2.04** | 0.84 ± 0.04 | **0.90 ± 0.04** | 0.82 ± 0.09 | 0.85 ± 0.07 |
| **0.1** | 298.2 | ellipsoid | **2.50** | 0.93 ± 0.05 | **0.96 ± 0.03** | 0.9 ± 0.1 | 0.9 ± 0.1 |
| | 310.2 | ellipsoid | **2.01** | 0.90 ± 0.05 | **0.95 ± 0.03** | 0.8 ± 0.1 | 0.89 ± 0.08 |
| **0.75** | 298.2 | ellipsoid | **4.10** | 0.94 ± 0.05 | **0.97 ± 0.01** | 0.8 ± 0.3 | 0.9 ± 0.2 |
| | 310.2 | ellipsoid | **2.94** | 0.85 ± 0.04 | **0.87 ± 0.04** | 0.82 ± 0.08 | 0.85 ± 0.06 |
| **1.0** | 298.2 | oblate | **6.32** | 0.80 ± 0.07 | **0.85 ± 0.05** | 0.6 ± 0.2 | 0.8 ± 0.1 |
| | 310.2 | ellipsoid | **3.11** | 0.92 ± 0.05 | **0.95 ± 0.00*** | 0.9 ± 0.1 | 0.9 ± 0.1 |
| **1.5** | 298.2 | prolate | **8.61** | 0.91 ± 0.07 | **0.94 ± 0.02** | 0.9 ± 0.1 | 0.90 ± 0.09 |
| | 310.2 | prolate | **4.48** | 0.8 ± 0.1 | **0.78 ± 0.04** | 0.7 ± 0.2 | 0.7 ± 0.1 |

Relaxation data collected at 600.13 and 500.34 MHz (both at 298.2 and 310.2 K) and GB1 NMR structure (PDB: 2JSV[21]) were used to perform the new model free analysis within relax. Global models for diffusion tensors were calculated following the iterative procedure. After calculation of all model-free models, the model with best global fit was chosen. Average of extracted $S^2$ is presented for each secondary structure motif, and the uncertainties are the standard deviation of the mean.

As I have previously observed in the backbone relaxation section, the highly different ratios of $R_2/R_1$ (Fig. 2.31-33) present in the alpha helix when compared to the almost geometrically orthogonal beta sheet regions suggest the existence of anisotropic rotational diffusion[112,113]. In broad agreement with these studies, when using a (prolate) anisotropic axially symmetric tensor, this apparent slow motion ($R_{ex}$ contribution to $R_2$) and the corresponding oscillation in the order parameters are eliminated, in particular for those conditions at low temperature/ high viscosity (no significative difference between average $S^2$ for the helix and β-sheet – Table 2.12). The $S^2$ and $R_{ex}$ parameters obtained for GB1 in pure water and in the presence of the IL as a function of residue number are super-positioned and plotted in Fig. 2.38. For clarity, in the following Fig. 2.39, for each condition such plot is shown for both temperatures, and slow timescales $S_s^2$ where $S^2 = S_f^2 S_s^2$ are also indicated. In addition, $S^2$ and $R_{ex}$ terms were color-coded and width-coded onto the structure of GB1 as illustrated in Fig. 2.40.

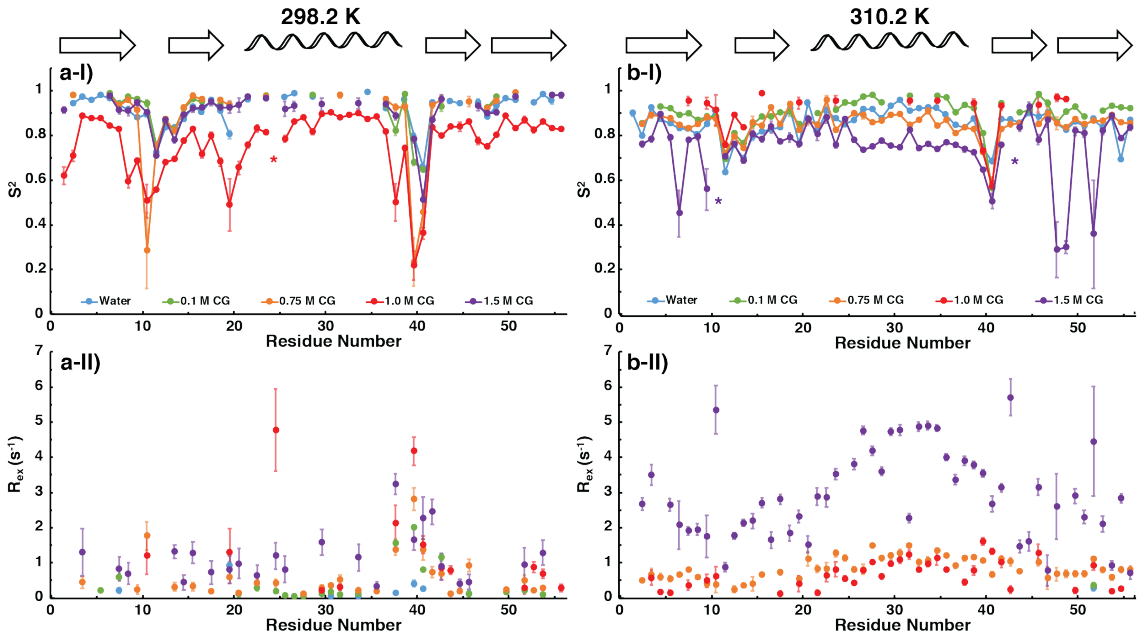



**Figure 2.38. Model-free dynamic parameters of GB1 in the presence of [Ch][Glu].**

Model-free relaxation analysis of GB1 at **a**) 298.2 K and **b**) 310.2 K using data measured at 500.34 and 600.13 MHz for water (blue) and increasing amounts of [Ch][Glu] (0.1, green; 0.75, orange; 1.0, red; 1.5 M, purple). **I**) The generalized order parameter $S^2$ provides a measure of atomic flexibility of the $^{1}H$-$^{15}N$ bond vector on the ps-ns time scale. The value of $S^2 = 0$ is indicated by symbol (*). **II**) The chemical exchange terms $R_{ex}$ at 600 MHz describe residues affected by motions occurring on the μs-ms time scale. The error bars corresponds to the uncertainty in each parameter from the model-free analysis (see material and methods).

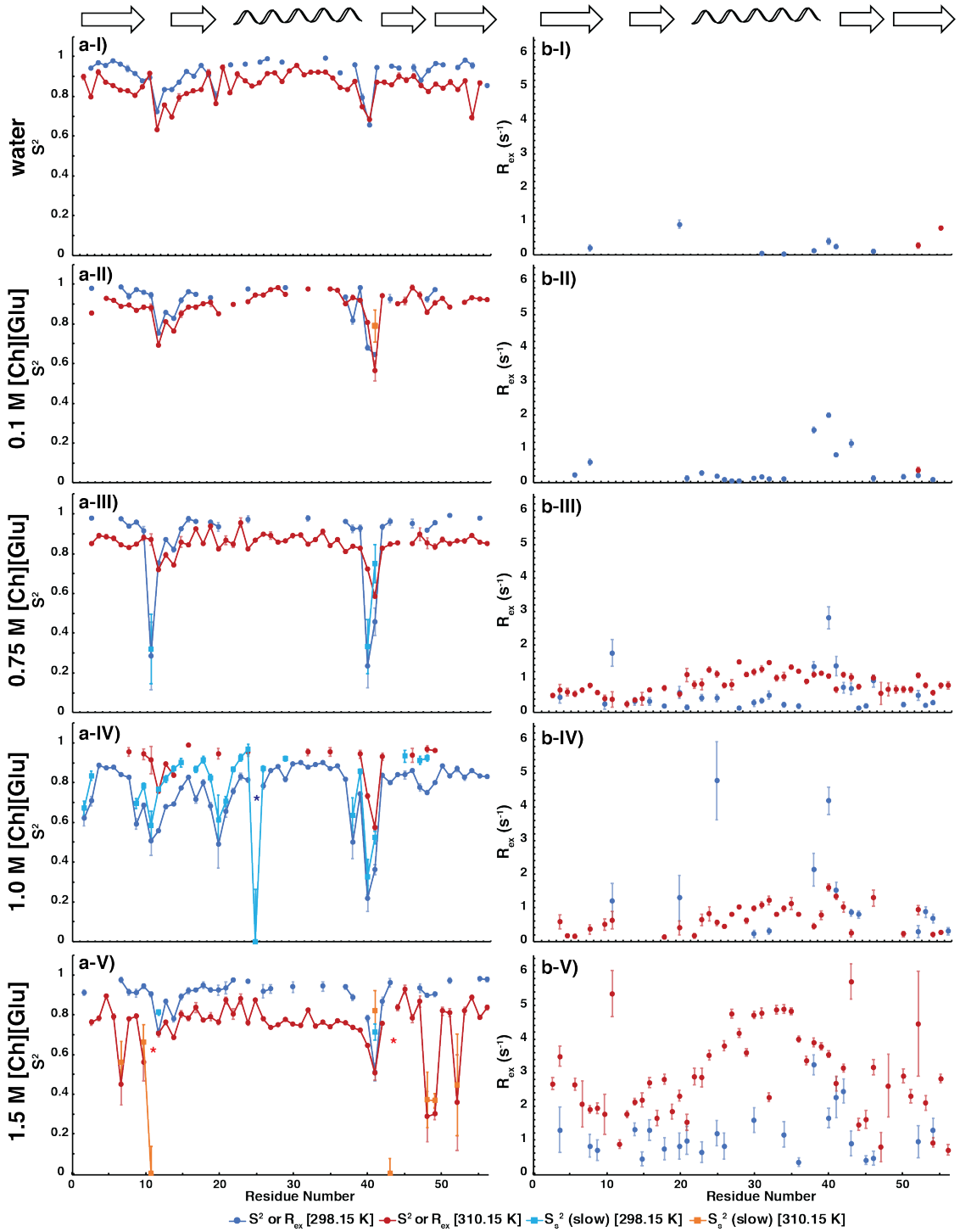



**Figure 2.39. Model-free dynamic parameters of GB1 in the presence of [Ch][Glu].**

Model-free relaxation analysis of GB1 at 298.2 K (blue) and 310.2 K (red) using data measured at 500.34 and 600.13 MHz for **I**) water and **II-V**) increasing amounts of [Ch][Glu]. The **a**) generalized order parameter $S^2$ provides a measure of atomic flexibility of the $^1H$-$^{15}N$ bond vector on the ps-ns time scale. The value of $S^2$ = 0 is indicated by symbol (*). For a few motional models, $S^2$ can be defined as $S^2 = S_f^2 S_s^2$, in which $S_f^2$ is the squared order parameter for fast time scale (< 20 ps) internal motions and $S_s^2$ is the order parameter for slow internal motions occurring on a time scale $\tau_s$ longer than ~500 ps. $S^2$ and $S_s^2$ are represented by circles or

squares, respectively. The **b**) The chemical exchange terms $R_{ex}$ at 600 MHz describe residues affected by motions occurring on the μs-ms time scale.

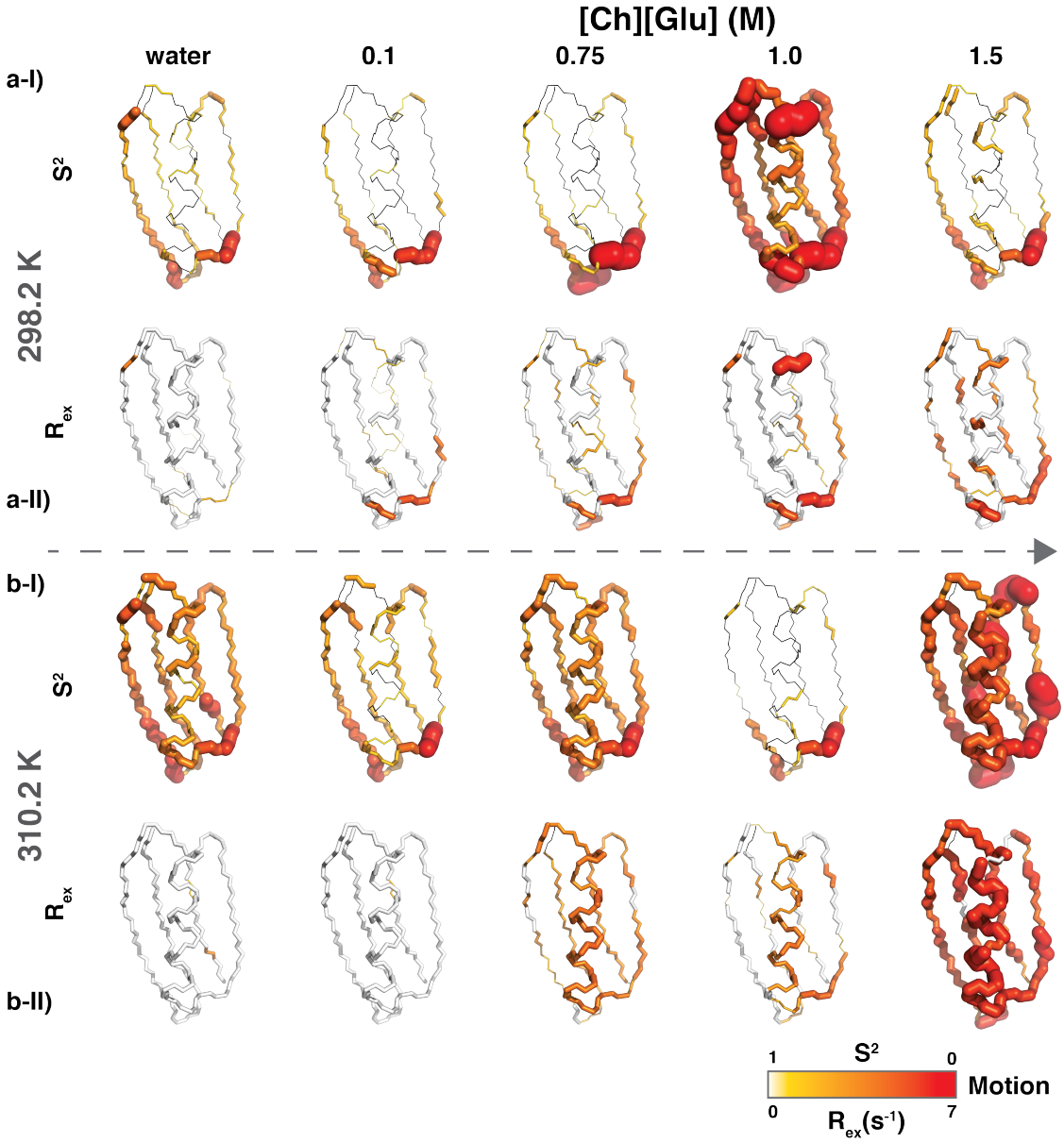



**Figure 2.40. Flexibility of GB1 on the ps-ns timescale and exchange contributions.**

**a**) The order parameter $S^2$ is colour- and width-coded onto the structure of GB1. Thicker lines and gradient from yellow to red indicate lower $S^2$ values and higher flexibility on the ps-ns time scale. **b**) $R_{ex}$ represents slower time scale motion (μs-ms) and is colour- and width-coded onto structure. The comparison of these coded structures highlights loop regions with mobility on both time scales. Structure representations were prepared in PyMOL, and colour coded structures were created using PyMOL macros automatically created by relaxGUI.

In general, high values are seen for the $S^2$ giving a picture of a largely rigid protein on the picosecond time scale, as expected from the very high proportion of residues that participate in secondary structure. The regions with the most significant mobility (low $S^2$) are those identified for relaxation data (residues around L12 and G41). The increase of temperature (and bulk viscosity) should increase the mobility of the protein in terms of $S^2$, however it is only verified for GB1 at dilute conditions (water and 0.1 M IL). Anomalous low $S^2$ values were found for the protein at high concentrations (≥ 0.75 M), in particular for the most mobile segments:

(1) For the measurements at 298.2 K, although there is an increase of $R_{ex}$ parameter, specifically in the loop connecting the helix with β-strand 3, the profile of GB1's $S^2$ at 1 M (Fig. 2.38a, Fig. 2.39a-IV, and Fig. 2.40a) is completely dominated by slow internal correlation times (> 0.5 ns), $\tau_s$ (Fig. 2.41). This suggests that internal motions are in the same nanosecond time scale as overall tumbling and somehow related with the minimum of protein stability along IL titration. Since the anisotropy of the shape of the model used is different (oblate spheroid compared to prolate or ellipsoid), artefacts of the fitting process cannot be excluded but the bond vectors that are not oriented along the long axis of the protein can be fitted with slow $\tau_s$.

(2) For the measurements at 310.2 K, the $S^2$ differences are explained by the evident increase in exchange broadening contributions from 0.75 to 1.5 M (Fig. 2.38b, Fig. 2.39a, and Fig. 2.40b). Although the free selection of an (prolate) anisotropic rotational diffusion model led to elimination of the $R_{ex}$ terms for the helix motif of GB1 at dilute solutions (as previously misinterpreted if an isotropic model is used[90,112,113]), the presence of slow chemical exchange should be true. The prolate diffusion tensor compensates the different orientation of the backbone NH bond vectors in the helix to those in beta sheet, therefore the bond vectors in the alpha helix experience a longer effective correlation time than the vectors in the beta sheet.

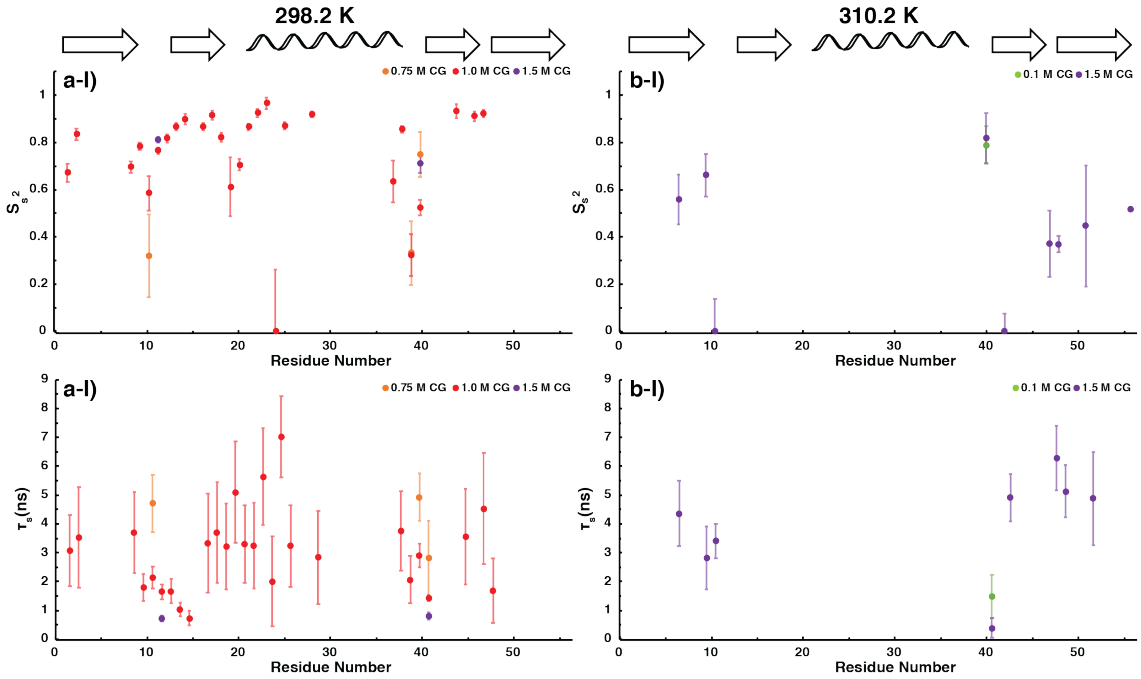



**Figure 2.41. Model-free slow dynamic parameters of GB1 in the presence of [Ch][Glu].**

**I**) $S_s^2$ and **II**) $\tau_s$ are the order parameter and corresponding correlation time obtained for slow internal motions (slower than ~500 ps) of GB1 at 298.2 K **a**) and **b**) 310.2 K using data measured at 500.34 and 600.13 MHz for [Ch][Glu] titration.

Following the early studies discussed above, according to $S^2$ values derived using relaxation-derived spectral densities (the $F(\omega)$ approach) and assuming that all NHs exhibit fluctuations on the ns timescale[114,115], the first turn of the helix, D22 to A24, is also highly flexible. This was not evident in our relaxation measurements neither in previous dynamics studies, even though it is in good agreement with our studies. Furthermore, Idiyatullin et al[114] investigated the temperature dependence (5 to 50 ºC) of both order parameters and internal correlation times of GB1. An internal motional activation energy, $E_i$, for each amide bond vector was derived from the temperature dependence of the internal correlation time. Residues with the highest $E_i$ values are all involved in hydrogen bonding and were found to correlate with heat capacity values ($C_p$). This trend of values mirrors the free energies for transient exposure of NH groups ($\Delta G_{HX}$) derived from HX measurements. The response of fast time scale motions to temperature (or other perturbation) may be related with the mechanism of exchange in a much slower macroscopic time scale.

According to the diffusion in-a-cone model of NH group motions, and under the assumption that the motions of individual NH groups are independent of each other, the contribution to the change in conformational entropy ($\Delta S_{conf}$) associated with the transition from an initial to a final state can be approximated to the changes in ps-ns timescale bond vector fluctuations in the two states (by means of the $S^2$)[140–142]. A simple relation describes $\Delta S_{conf}$ in terms of the order parameters of bond vector $j$ between the two states, denoted initial and final ($S_{j,initial}$ and $S_{j,final}$, respectively), *via*[141]:

$$\Delta S_{conf}(j) = k_B ln \left\{ \frac{3 - \left(1 + 8\, S_{j,final}\right)^{1/2}}{3 - \left(1 + 8\, S_{j,initial}\right)^{1/2}} \right\}, \qquad [2.7]$$

where $k_B$ is Boltzmann's constant, $S = \sqrt{S^2}$, and $\Delta S_{conf} = S_{conf,final} - S_{conf,initial}$.

An increase in $S^2$ of 0.1 for a single bond vector ($j$) corresponds to an increase in conformational free energy ($\Delta G_{conf} = -T\Delta S_{conf}$) of ~0.15-0.45 kcal/mol with potential contributions for the overall stability of the protein[143]. Using the order parameters of GB1 in water and in the presence of [Ch][Glu], $S_{j,nitial}$ and $S_{j,final}$, respectively (Table A9, A10 of Appendix A), $\Delta S_{conf}$ values were determined for each backbone NH bond vector at 298 or 310 K of temperature (Fig. 2.42a). Qualitatively, these data suggest that while for most conditions the $\Delta S_{conf}$ change is insignificant ($\pm$ 1 cal/mol K), for data acquired at 1.0 M and 298 K there is a significative increase of the backbone entropy ($\Delta S_{conf,298K,1M} = 3 \pm 1$ cal/mol K). While the order parameters used in this analysis have a clear influence of slow internal correlation times (as discussed above), the corresponding $\Delta G_{conf,298K,1M}$ of $-0.8 \pm 0.3$ kcal/mol agrees with the lowest stability observed in other experiments to this range of [Ch][Glu] concentrations.

Furthermore, the temperature dependence of $S^2$ was used to determine the entropy change with temperature, directly associated to the conformational heat capacity ($C_{p,conf}$), where $S_{\text{initial}}$ and $S_{\text{final}}$ are the order parameters at 298.2 K and 310.2 K of temperature (Fig. 2.42b). With exception of the 1.0 M IL condition, there is a general increase in the conformational entropy as expected with the increase of temperature. Because GB1 has a high intrinsic conformational heat capacity[112], this leads to a stabilisation of the native state. On the other hand, the reduced mean $\Delta S_{conf}$ values from the increase of temperature (298 K to 310 K) at 1.0 M IL suggests a surprisingly rigidification of the structure. This restriction of motion is presumably coupled to the formation of enthalpically favourable interactions by the IL that overcome the gain of entropy for the backbone at higher temperature.

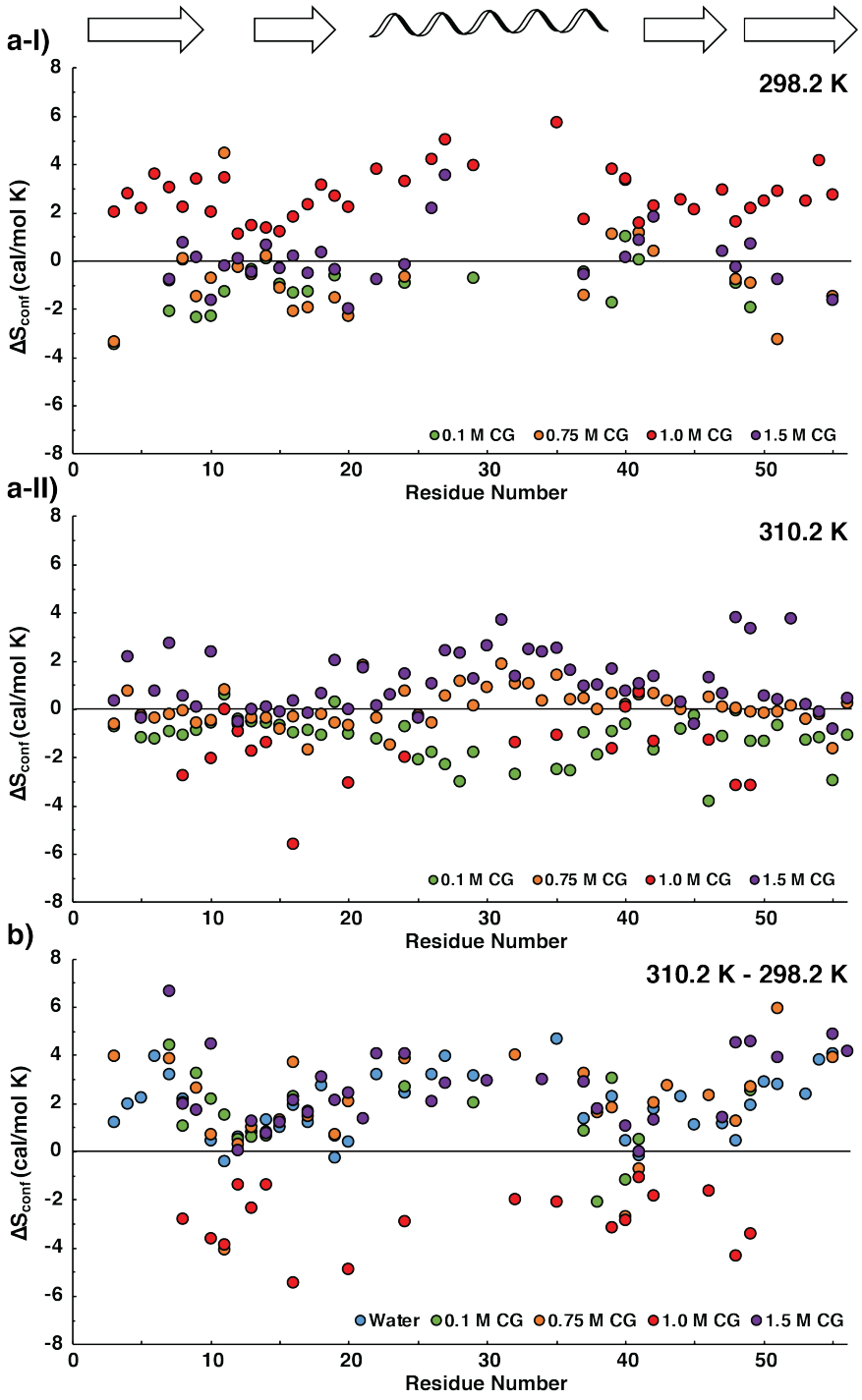



**Figure 2.42. Contributions from rapid $^{15}$N-NH backbone bond vector dynamics to the conformational entropy difference ($\Delta S_{conf}$) in the presence of IL and/or increase of temperature.**

**a**) $\Delta S_{conf}$ between GB1 in the presence of increasing amounts of [Ch][Glu] (0.1, green; 0.75, orange; 1.0, red; 1.5 M, purple) and water [$\Delta S_{conf} = S_{conf,final}$ (IL) – $S_{conf,initial}$ (water)] as a function of residue, calculated from $S^2$ values, at I) 298.2 K or II) 310.2 K. **b**) $\Delta S_{conf}$ between GB1 at 310.2 K and 298.2 K of temperature, in water (blue) and increasing amounts of [Ch][Glu] [$\Delta S_{conf} = S_{conf,final}$ (310.2 K) – $S_{conf,initial}$ (298.2 K)] as a function of residue, calculated from $S^2$ values. $\Delta S_{conf}$ values were calculated accordingly to the Eq. 2.7 and the uncertainties were propagated from $S^2$ error (if the error bar is shorter than the size of the symbol, it would not appear).

Comparing the backbone and side-chain methyl dynamics of 10 surface β-sheet mutants of GB1 (T53 as the guest position) in their folded states, Stone and his colleagues[99,144,145] estimated that $\Delta G_{conf}$ varies by more than 4 kcal/mol among the mutants, which is comparable in magnitude to the $\Delta\Delta G_u \sim 2.2$ kcal/mol[146] but it is poorly correlated with the global free energy of the domain. The variation of backbone entropy in the folded state appear to make a substantial, but not dominant, contribution to GB1 stability (unfolded state is also important as we will see in the next chapter). The fact is that assuming independent motions of different NH bond vectors lead to an overestimation of the residual entropy of proteins[145]. If the motions of two bond vectors are coupled, then a perturbation of the protein (such as mutations, different ligands or different chemical or physical conditions) that modifies the dynamics of one vector is likely to influence the coupled vector in a similar manner (i.e., covariation analysis of dynamic parameters). In this sense, it was found in GB1 an positive covariation of $S^2$ and $\tau_e$ values, which is most prevalent for the regions with spatial proximity of aminoacids (the second β-strand, the first 3-turns of the $\alpha$-helix, the extended loop and β3-β4 hairpin turn)[145]. Interestingly, the most affected two residues of GB1 by [Ch][Glu] in terms of chemical shift (A23 in the first turn of the helix and D40 in the extended loop) are those where a strong dynamic correlation is observed. These observations suggest that a long-range network of correlation motions may exist and could influence protein stability.

## Conclusion

Our data show that while denaturation induced by [Bmim][dca] is largely explained by preferential hydrophobic interactions with the hydrophobic core of GB1, the (de)stabilisation of GB1 with the increase of [Ch][Glu] concentration is a shallow parabola (minima of $T_m$ around 0.75 M) and it is not only dependent on overall solvent properties but also on the solvent-mediated interactions with charged, more flexible and solvent-accessible regions of the protein. In particular, the contact between the anion and the A23 residue (melting hotspot region) seems to be crucial to increase protein stabilisation. The [Ch][Glu]-GB1 interactions probed by means of the protection of amide protons, backbone $^{15}$N relaxation data, and translational diffusion measurements, show clear deviations from the expected exchange rate, dynamics, and protein hydrodynamic radius at high IL concentrations. These observations corroborate the hypothesis that a predominant transition exists from preferential [Ch][Glu] molecules exclusion (protein hydration) to its accumulation at the protein surface as a consequence of the increase of IL concentration. Assuming an entropic mechanism due to excluded-volume effects, while this IL accumulation (at high concentrations) around the native surface is expected to increase protein stabilisation, the IL exclusion and concomitant electrostatic screening (at low concentrations) is likely to destabilise the protein. Molecular dynamics (MD) have to be conducted to prove this hypothesis.

## Materials and Methods

Chemicals, [Ch][Glu] IL synthesis and characterization, and a detailed protocol for non-labelled, $^{15}N$, $^{15}N/^{13}C$, or $^{2}H/^{15}N/^{13}C$ isotopically labelled GB1 expression and purification can be found in this section. NMR sample preparation, data collection and analysis procedures are described in detail.

**Chemicals.** L-glutamic acid (> 98.5 % of purity) was purchased from PanReac. 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca]) and 1-butyl-3-methylimidazolium chloride ([Bmim][Cl]) ILs were sourced from IoLiTec (Denzlingen, Germany). The ILs were at least 98% pure and were dried for 24 h under vacuum at 60 ºC before solution preparation. Isotopically enriched chemical compounds, [$^{15}N$]-$NH_4Cl$, [$^{13}C$]-glucose and deuterium oxide ($D_2O$) were purchased from Cambridge Isotope Laboratories. Sodium- L-glutamate, ion exchange resin Amberlyst A-26 (OH) and SnakeSkin 3.5K MWCO dialysis tubing were purchased from Thermo Fisher Scientific. Unless otherwise described, all other chemicals were purchased from Sigma-Aldrich. Pure *milliQ* water with resistance 18.4 M$\Omega$ $cm^{-1}$ was used in all experiments. pH values are direct meter readings uncorrected for any isotope effect and were measured with Docu-pH meter (Sartorius) calibrated with standard solutions.

**Synthesis and characterization of [Ch][Glu] IL.** For the choline glutamate ([Ch][Glu]) IL synthesis, I used, with slight modifications, the method reported by De Santis, et al[147]. Based on a potentiometric titration *via* neutralization reaction, [Ch][Glu] was synthesized as follows (schematized in Fig. 2.43). Choline hydroxide ([Ch][OH]) aqueous solution was prepared from choline chloride ([Ch][Cl]) (4.75 g, 34 mmoles) in water that was passed slowly through a column packed with ~ 43 mL of anion exchange (Amberlyst A-26) resin and then washed. The absence of chloride content was tested with a silver nitrate qualitative test. Freshly [Ch][OH] aqueous solution (175 mL, 31 mmoles) was added dropwise an equimolar ratio to L-glutamic acid (4.5 g, 31 mmoles) in water. The reaction was monitored by reading the solution pH until it reached the equivalence point (pH=7.00) and the stoichiometry 1:1 [Ch]/[Glu] ratio was confirmed by characterization of the solution by $^{1}H$ NMR using the integrals ratio. The mixture was stirred at about 25 °C for 24 h in the dark and subsequently the solvent was evaporated under reduced pressure. The product was dried *in vacuo* for 24 h at 60 °C to yield [Ch][Glu] (8.5 g, >99%) as a slightly yellow oily or glassy compound at room temperature.

The structure of the resulting [Ch][Glu] IL was confirmed by $^{1}H$ and $^{13}C$ NMR spectroscopy (see below the chemical shifts). The spectra were recorded at 25 ºC on a 400 MHz Bruker AVANCE II+ instrument operating at 400.15 MHz for protons and 100.6 MHz for $^{13}C$, equipped with a 5 mm high-resolution BBO probe with pulsed gradient units. Solution was prepared by dissolving [Ch][Glu] in 99.9% $D_2O$ to a concentration of 0.5 M. The glass transition ($T_g$) and decomposition temperature ($T_d$) of pure [Ch][Glu] IL were determined to -21.3 ºC and 221.3 ºC, respectively. The measurements were conducted using a differential thermal analyser DSC 131 (Setaram) and a thermogravimetric analyzer Labsys EVO (Setaram).

$^{1}$H NMR (400.15 MHz, $D_2O$, 25 ºC) $\delta_H$ (ppm): 1.84 – 2.02 (m, 2H, $CH_2$, Glu), 2.20 (apparent q, 2H, $CH_2$, Glu), 3.05 (s, 9H, $CH_3$, $CH_3$, $CH_3$, Ch), 3.37 (apparent t, 2H, $CH_2$, Ch), 3.60 (q, J = 4.87, 7.12 Hz, 1H, CH, Glu), 3.88 – 3.93 (m, 2H, $CH_2$, Glu).

$^{13}$C NMR (100.6 MHz, $D_2O$, 25 ºC) $\delta_C$ (ppm): 27.34 ($CH_2$, Glu), 33.6 ($CH_2$, Glu), 53.90 ($CH_3$, $CH_3$, CH3, Ch), 54.8 (CH, Glu), 55.61 ($CH_2$, Ch), 67.45 ($CH_2$, Ch), 175.1 (CO, Glu), 181.36 (CO, Glu).

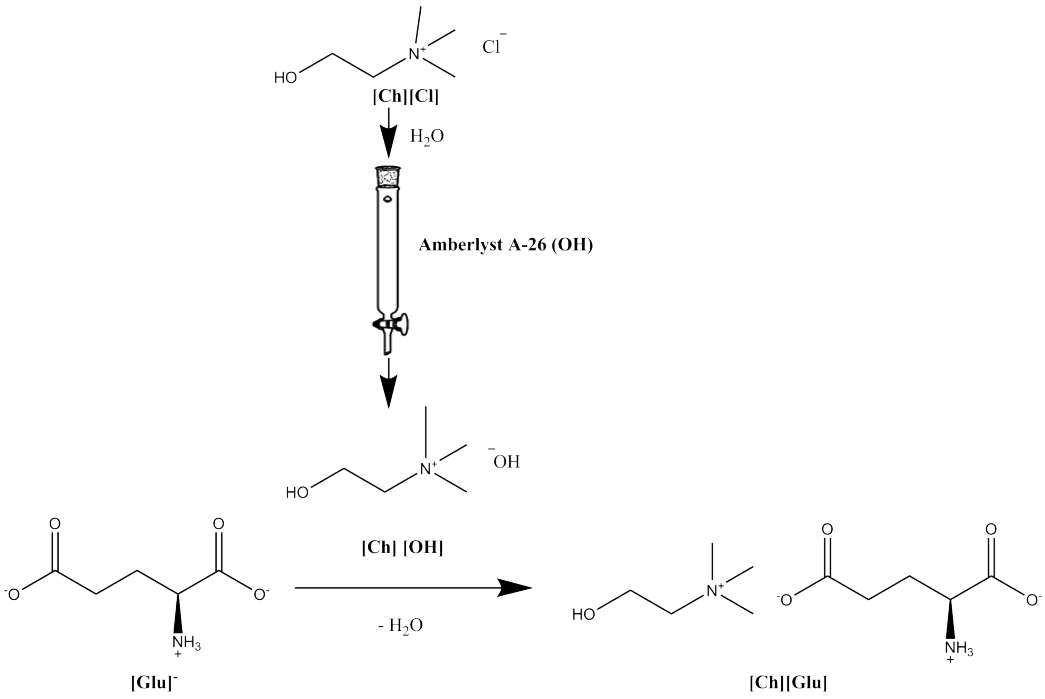



**Figure 2.43. Schematic synthetic procedure for the preparation of [Ch][Glu] IL.**

**Viscosity measurements and extrapolations.** Viscosity of pure water at 298.2 K and 310.2 K is 0.890 cP and 0.691, respectively[118]. Viscosities of 0.1 M, 0.5 M, 1.0 and 1.5 M aqueous-[Ch][Glu] IL were measured from 294.2 to 312.2 K in 2 K increment at atmospheric pressure using an automated SVM 3000 rotational Stabinger viscometer–densimeter (Anton Paar). The SVM 3000 uses Peltier elements for fast and efficient thermal stability. At each concentration, to estimate an uncertainty from standard deviation, a minimum of three viscosity values were measured for three temperatures (298.2, 304.2 and 310.2 K). Table 2.13 shows the measured viscosities in triplicate for each [Ch][Glu] concentration for three temperatures. For each condition, the viscosity as a function of $T$ was fit to an exponential function, which was used to extrapolate the $\eta$ at 298.2 K and 310.2 K. Likewise, the dependency of the solvent viscosity as a function of [Ch][Glu] concentration ([Ch][Glu]) at 298.2 K and 310.2 K can be described by empirical exponential equations,

$$\eta\,(298.2\ K, [Ch][Glu]) = 0.946491 \exp(0.809821\ [Ch][Glu]), \quad [2.8]$$

$$\eta\,(310.2\ K, [Ch][Glu]) = 0.751187 \exp(0.785736\ [Ch][Glu]), \quad [2.9]$$

these equations have no physical meaning but describe the dependence of viscosity on [Ch][Glu] concentration very well at 293.2 K and 310.2 K. Viscosity of 0.75 M [Ch][Glu] at 298.2 K and 310.2 K was calculated to 1.737 cP and 1.354 cP, respectively.

**Table 2.13. Measured viscosities for aqueous-[Ch][Glu] solutions.**

| | Aqueous-[Ch][Glu] viscosity (cP) | | | |
|---|---|---|---|---|
| **Temperature (K)** | **0.1 M** | **0.5 M** | **1.0 M** | **1.5 M** |
| **298.2** | 1.113 ± 0.006 | 1.417 ± 0.005 | 2.04 ± 0.01 | 3.26 ± 0.06 |
| **304.2** | 0.993 ± 0.007 | 1.26 ± 0.01 | 1.797 ± 0.009 | 2.82 ± 0.03 |
| **310.2** | 0.89 ± 0.01 | 1.13 ± 0.02 | 1.598 ± 0.006 | 2.47 ± 0.02 |

The uncertainties are the standard deviation from the triplicate dataset.

**GB1 expression and purification**

Vector. Professor Gary Pielak from University of North Carolina at Chapel Hill kindly provided the pET11a plasmid containing the gene encoding T2Q B1 immunoglobulin G binding domain of *streptococcal* protein G (GB1). The T2Q mutation prevents N-terminal deamidation (i.e., cleavage of the N-terminal Met residue)[146]. I refer to this variant as the wild-type protein or as only "GB1".

Protein expression and purification. Non-labelled, $^{15}N$, $^{15}N/\ ^{13}C$, or $^{2}H/^{15}N/^{13}C$ isotopically enriched GB1, containing 56 residues, was overexpressed and purified based on previously described protocols[26,28]. *Escherichia coli* BL21 (DE3) competent cells (NZYTech) were transformed with the pET-11a plasmid vector containing the gene encoding T2Q GB1 protein by heat shock. Cells were grown with shaking at 37 ºC and 180 rpm in LB medium or M9 minimal medium supplemented with 100 μM $FeSO_4$, 100 μM $CaCl_2$, 2 mM MgSO4, 10 mg/L Thiamine-HCl, 0.5% MEM Vitamins, 2.5 g/L of $^{15}NH_4Cl$ and 4 g/L of glucose or $^{13}C$-glucose, and 100 mg/L ampicillin. For uniform $^{2}H$-labeling scheme, cells were grown in 99.9% $D_2O$-based M9 minimal media: cell culture was adapted to grow in $D_2O$ by increasing the amount of $D_2O$ in the growth medium from zero to 99.9% in each growth cycle. Protein expression was induced by adding 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG, NZYTech) at $OD_{600}$ (optical density at 600 nm) = 0.6. After 3 h induction in the same conditions, cells were harvested [6,400 x *g*, 12 min, 4ºC in a JA-10 rotor (Avanti J-26S XPI, Beckman Coulter)] and frozen at -20 ºC overnight. Pellet was resuspended in 30 mL of lysis buffer per liter of culture [10 mM Tris-HCl (pH 7.5), 1 mM EDTA] preheated to 80 ºC. The sample was stirred and heated until the temperature reached 80 °C for 5 min. The lysed cells were cooled on ice for 10 min and then centrifuged [29,000 x *g*, 30 min, 4ºC in a JA-25.50 rotor (Avanti J-26S XPI, Beckman Coulter)]. Supernatant was dialyzed (SnakeSkin 3.5K MWCO) overnight at 4ºC against Buffer A [20 mM Tris-HCl (pH 7.5)]. The dialyzed sample was loaded to a HiTrap Q HP anion-exchange column (GE Healthcare) on an ÄKTA start system (GE Healthcare), and buffer B [20 mM Tris-HCl (pH 7.5), 1 M NaCl] was used to produce a

gradient 0 - 400 mM NaCl where GB1 elutes at around 200 mM NaCl since the binding with the column is weak.

Fractions containing the protein of interest were pooled and concentrated with Vivaspin turbo 15 3 K MWCO centrifugal concentrator (Sartorius). In a final step, a size exclusion chromatography step [Superdex 75 10/300 GL column (GE Healthcare) in Shimadzu prominence machine, at 4ºC] was applied to further purify and adjust GB1 to a different buffer [20 mM potassium phosphate buffer (pH 6.0) and 50 mM NaCl]. Purity was analysed in each step by SDS-PAGE (BioRad). Pure samples were extensively desalted by dialysis in *milliQ* water, flash frozen, lyophilized (Edwards Modulyo Freeze Dyer) and stored at -20 ºC until usage. Final yield of purified protein was 30 mg of GB1 per litre of M9 minimal medium. Concentrations were assessed spectrophotometrically by absorption measurements at 280 nm ($\varepsilon$= 9,970 $M^{-1}$ $cm^{-1}$) using NanoDrop ND-1000 UV-Vis (Thermo Fisher Scientific).

**Protein melting temperature ($T_m$) determination**

Aqueous-solution pH. To investigate how the different cosolutes affect the overall solution pH, I measured the pH for each IL or salt at the different concentrations. Stock solutions for each cosolute ([Ch][Glu], [Bmim][dca], [Bmim]Cl, [Ch]Cl, Na[Glu], Na[dca], and NaCl) were prepared in *milliQ* water and then diluted for the desired concentration. Additionally [Ch][Glu] and [Bmim][dca] stock solutions and their dilutions were also prepared in buffer 0.05 M sodium phosphate pH 7.2. pH of stock solutions was adjusted for 7.2 ± 0.2 with negligible microliter addition of HCl or NaOH solutions. pH values of each sample were measured with Docu-pH meter (Sartorius) calibrated with standard solutions.

Fluorescence spectroscopy. Steady state fluorescence measurements were carried out on a Cary Eclipse fluorescence spectrophotometer (Varian) using 10 mm pathlength fluorescence quartz cuvettes. The concentration of non-labelled GB1 was kept at 52 μM, previously determined by NanoDrop ND-1000 (Thermo Fisher Scientific). For the titration experiments a series of solutions with varying IL ([Ch][Glu], [Bmim][dca], [Bmim]Cl) or salt ( [Ch]Cl, Na[Glu], Na[dca], and NaCl) concentration from 0 to 2.0 M or 2.5 M (e.g., 0, 0.05, 0.10, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 2.00, 2.50 M) were prepared from each solution stock in *milliQ* water. Additionally [Ch][Glu] and [Bmim][dca] solutions were also prepared in 0.05 M sodium phosphate pH 7.25. The intrinsic tryptophan fluorescence was monitored by exciting the protein at 280 nm and the emission was collected at 350 nm. The spectral slit widths employed were 5 nm for excitation and emission. For titrations with [Ch][Glu], Na[Glu], [Ch]Cl, NaCl, thermal scans were obtained from 50 to 100 ºC with 1 ºC/min rate and 0.25 of data interval collection. For titrations with [Bmim][dca], Na[dca], and [Bmim]Cl, thermal scans were obtained from 10 to 100 ºC with 1 ºC/min rate and 0.50 of data interval collection. Temperature was maintained through a Peltier based controller. For each experiment, a suitable temperature range was chosen to normalize the data. $T_m$ values were determined by fitting the intensity decay to a sigmoidal function using Prism 8 (GraphPad Software). The standard error was obtained from the fitting.

Differential Scanning Calorimetry. DSC data was recorded using a NanoDSC (TA instruments). The concentration of non-labelled GB1 was kept at 2 mg/mL, previously

determined by NanoDrop 2000 (Thermo Fisher Scientific). Protein samples were prepared in water, 0.1 M sodium phosphate, and 0.1 M, 1 M, and 2 M [Ch][Glu]. A scan rate of 1.0 ºC/min, from 25 to 100 ºC with an equilibration period of 10 min was used with a constant cell pressure of 3 atm, 12 – 15 baselines (consecutive heating scans) were acquired before each experiment. Samples for DSC analysis were degassed prior to any DSC measurement to avoid bubble formation during the temperature scan.

Data were analysed using the NanoAnalyze software package within NanoDSC. Baselines were corrected by subtracting the no-protein scan from the protein scan. A nonlinear least squares analysis was performed with a two-state transition model. $T_m$ values were measured for all conditions, but an accurate GB1 stability curve could be constructed just for low concentrations of [Ch][Glu] due to inadequate post transitional baselines for robust fitting. The GB1 free energy of denaturation ($\Delta G_u^{0'}$) was extrapolated to 37 ºC using the experimental calorimetric enthalpy [$\Delta H_u^{0'}(T_m) \approx \Delta H_{vH}^{0'}(T_m)$] and the constant heat capacity change upon unfolding, $\Delta C_{p,u}^{0'} = C_{p,U}^{0'} - C_{p,F}^{0'} = 2.9 \pm 0.3$ kJ/mol K[20], according to the integrated Gibbs-Helmholtz equation[98],

$$\Delta G_u^{0'}(T) = \Delta H_u^{0'}(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_{p,u}^{0'}\left[(T - T_m) - T\,ln\left(\frac{T}{T_m}\right)\right], \quad [2.10]$$

where $T_m$ is the melting temperature, $\Delta H_u^{0'}(T_m)$ and $\Delta S_u^{0'}(T_m)$ are the enthalpy and entropy at $T_m$ [$\Delta S_u^{0'}(T_m) = \Delta H_u^{0'}(T_m)/T_m$]. Calorimetric data have established that over a reasonably small temperature range the change in $C_p$ is rather small[148], and that considering $\Delta C_{p,u}^{0'}(T_m) = \Delta C_{p,u}^{0'}$ (T) leads to meaningful results for $\Delta G_u^{0'}(T)$.

Differential Scanning Fluorimetry. DSF experiments were carried out with an RT-PCR instrument StepOnePlus Real-Time PCR System (Applied Biosystems). The protocol used was optimized from previous studies[54,149]. PCR 96-well microplate (Multiplate™ white, Bio-Rad) was filled with 20 $\mu$L of solution per well (i.e., protein, ligand/buffer/water and dye). The final concentration of lysozyme from chicken egg white (purity > 90%) per well was 4.4 μg (15.4 μM) as determined by NanoDrop ND-1000. Stock solution of IL ([Ch][Glu]) and salts ([Ch]Cl, [Na][Glu], NaCl) were prepared to titrate the protein from 0 to 2.5 M of concentration (e.g., 0, 0.05, 0.10, 0.20, 0.30, 0.50, 0.60, 0.75, 1.0, 1.50, 2.0, 2.5) in 0.05 M sodium phosphate pH 7.25. Additionally, [Ch][Glu] titration was also performed in *milliQ* water. Fluorescent dye SYPRO orange protein gel stain was used at 5x concentrated final solution and it was the last component to add. Each condition was made in triplicate and controls were performed without protein. The microplates were sealed with an adhesive optical clear seal (Microseal® 'B' seal, Bio-Rad) and centrifuged at 1000 rpm for 1 min before being heated. Data were collected using the channel for the reporter ROX™ using the fast ramp speed from 25ºC to 99ºC. The data was analysed using the Protein Thermal Shift™ Software v1.1 (Applied Biosystems). The transition midpoint of the sigmoidal curve was calculated automatically, and the error was measured from the triplicate of each condition.

**NMR spectroscopy**

Sample preparation, data acquisition and processing. All samples for NMR spectroscopy were prepared in *milliQ* water with 10% (v/v) $D_2O$, 0.1% (v/v) $NaN_3$ and 50 $\mu$M

sodium-2,2-dimethyl-2-silapentane-5-sulfonate-$d_6$ (DSS, Eurisotop), pH was adjusted for 7.1 ± 0.1 with negligible microliter addition of HCl or NaOH solutions. Except where stated, spectra were recorded at 298.2 K on a 600 MHz Bruker AVANCE II+ spectrometer equipped with a 5-mm inverse detection triple-resonance z-gradient TCI cryoprobe, operating at a $^{1}H$ Larmor frequency of 600.13 MHz, 60.81 MHz for $^{15}N$, and 150.90 MHz for $^{13}C$. Spectrometer temperature was calibrated using a pure methanol-$d_4$ standard Bruker sample[150]. Further specific details of data acquisition (pulse programs, spectral windows, etc) and processing for each experiment are given in the following subsections. Proton chemical shifts were referenced against internal DSS while nitrogen and carbon chemical shifts were referenced indirectly to DSS using the absolute frequency ratio[151]. Processing and visualization of NMR spectra were performed with NMRPipe (version 10.1)[152] and CcpNmr Analysis 2.5[153]. Backbone amide assignments are based on published work[19,154]. The BMRB databank under the accession code 25909[100] shows $^{1}H$, $^{13}C$, $^{15}N$ chemical shift assignments for GB1 in similar conditions for those studied here.

Backbone chemical shift titrations. The standard Bruker $^{1}H$-$^{15}N$ HSQC (hsqcetfpf3gpsi2) pulse sequence employs a sensitivity-enhanced pulse field gradient[155]. The spectra were acquired with 2048 ($^{1}H$) and 128 ($^{15}N$) complex points for a spectral width (SW) of 13 ppm ($^{1}H$) and 37 ppm ($^{15}N$), with 4 scans (or 8 scans for low GB1 concentration).

Two-dimensional $^{1}H$-$^{15}N$ HSQC experiments were successively collected for GB1 with increasing concentration of two different ionic liquids ([Ch][Glu], [Bmim][dca]) and salts related with [Ch][Glu] moieties ([NaGlu], and [ChCl]). A solution of uniformly $^{15}N$ or $^{15}N/^{13}C$-labelled GB1 (*milliQ* water with 10% $D_2O$, 0.1% $NaN_3$ and 50 $\mu$M DSS) was titrated individually with different ILs/ salts (e.g., 0, 0.05, 0.1, 0.25, 0.50, 1.00, 1.25 and 1.50 M of concentration), in order to maintain the protein concentration constant throughout the titration procedure.

One $^{1}H$-$^{15}N$ HSQC spectrum was acquired for each cosolute concentration sample. A 0.6 mM $^{15}N/^{13}C$ GB1 was used for [Ch][Glu] titration regarding the further side-chain correlation experiments (see below). Similarly, $^{1}H$-$^{15}N$ HSQC spectrum for GB1 dissolved in buffer 0.1 M sodium phosphate (pH 7.15) was also recorded. A 0.6 mM $^{15}N$ GB1 was used for Na[Glu] and [Ch]Cl salts to compare with [Ch][Glu] IL. A 0.25 mM $^{15}N$ GB1 was used for [Bmim][dca] titration. Additionally, $^{1}H$-$^{15}N$ HSQC spectrum of 0.25 mM $^{15}N$ GB1 was obtained individually for [Ch][Glu], Na[Glu], [Ch]Cl and NaCl at 1.0 M of concentration.

Combined $^{1}H$-$^{15}N$ chemical shift differences of the amide in 2D [$^{1}H$-$^{15}N$]-HSQC spectra, $\Delta\delta_{comb}$ or chemical shift perturbations (CSP), were calculated as

$$\Delta\delta_{comb} = \sqrt{(\Delta\delta(H_N)^2 + (\alpha\Delta\delta N_H)^2}, \quad [2.11]$$

where $\Delta\delta H_N$ and $\Delta\delta N_H$ are the $^{1}H$ and $^{15}N$ chemical shift differences of GB1 in the presence of added IL/ salt minus the same resonance in the absence of added species, normalized with the scaling factor $\alpha$= 0.14 for most residues but $\alpha$= 0.2 for glycine[59]. To decide whether a given residue belongs to the class of interacting or non-interacting residues, I have calculated a cut-off value, based on the corrected standard deviation to zero, according to the method developed by Schumann, et al[60]. Since the total concentrations of ligand and protein are known during titration, $[L]_t$ and $[P]_t$, for residues above the cut-off, the $\Delta\delta_{comb}$ values were

used to obtain the dissociation constant ($K_d$) from the titration experiments ([Ch][Glu] and [Bmim][dca]) by non-linear regression analysis according to[59],

$$\Delta\delta_{comb} = \Delta\delta_{max}\frac{(K_d + [L]_t + [P]_t) - \sqrt{(K_d + [L]_t + [P]_t)^2 - (4[P]_t\,[L]_t)}}{2[P]_0}, \quad [2.12]$$

where $\Delta\delta_{comb}$ is the combined chemical shift deviation from the free state defined by Eq. 2.11 and $\Delta\delta_{max}$ is the maximum chemical shift change on saturation. Data are fitted to a single site binding model, using a least square fitting search of Microsoft Excel Solver to find the values of $K_d$ and the chemical shift of the fully saturated protein.

Amide proton temperature coefficients. 2D $^{1}$H-$^{15}$N HSQC were measured at 4 temperatures: 283.2, 298.2, 310.2 and 323.2 K, for water and 1.5 M [Ch][Glu] samples at magnetic field strength of 600.13 MHz ($^{1}$H frequency). Each 2D spectrum was recorded as a complex data matrix of 2048 ($^{1}$H) and 128 ($^{15}$N) points for a SW of 14 ppm ($^{1}$H) and 38 ppm ($^{15}$N), with 8 scans. The amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) were determined from the change in the upfield $^{1}H_N$ chemical shifts with increasing temperature[72]. The $^{1}H_N$ chemical shifts were extracted from peak assignments for each $^{1}$H-$^{15}$N HSQC with CcpNmr2.5. The values and their uncertainties were determined using a linear least-squares fit of $^{1}H_N$ chemical shifts from temperature increments. The fitting was performed with Prism 8 (GraphPad Software).

NMR Side-Chain correlation experiments. Complementary 2D $^{1}$H-$^{15}$N H2(C)N[57] and $^{1}$H-$^{13}$C H(C)CO[58] correlation spectra were collected at each [Ch][Glu] titration point at 298.2 K.

To follow $^{15}$N chemical shift of GB1's lysine side chain $NH_3^+$ groups, I carried out a triple resonance experiment that correlates lysine $^{1}$Hε, $^{13}$Cε and $^{15}$Nζ resonances. The pulse sequence used is essentially identical to the H2(C)N experiment by André et al[57] except that 15N ζ - selective SNOB 180° pulses were incorporated[156,157]. The $^{1}H^{\varepsilon}$-$^{15}N^{\zeta}$ spectra were measured with the 2D version of the experiment, for which cross peaks are with lysine $^{1}$Hε and $^{15}$Nζ resonances. The $^{1}$H carrier position was set at the water resonance, $^{15}$N carrier position at 32.25 ppm, and $^{13}$C pulses were centred at 40 ppm. WALTZ-16 $^{1}$H decoupling was applied with 70 $\mu$s of pulse length during nitrogen chemical shift evolution to suppress line broadening due to scalar relaxation resulting from proton exchange with the solvent. Selective $^{15}$Nζ r-SNOB and $^{13}$Cε r-SNOB 180° pulses were applied with 1.00 and 0.98 ms, respectively. The spectra were acquired with 2048 ($^{1}$H) and 100 ($^{15}$N) complex points for a spectral width (SW) of 13 ppm ($^{1}$H) and 4 ppm ($^{15}$N), with 32 scans. Since this experiment detects non-labile $^{13}$C-attached $^{1}$H nuclei, cross peaks from all lysine residues show up regardless of hydrogen exchange rates for the amine groups. Peak assignments were based on chemical shift assignments previously published[62].

To follow $^{13}$C chemical shift of GB1's side chain carboxylate groups (Gln, Glu, Asp and Asn residues), I carried out a triple resonance experiment that correlates carboxyl/carbonyl ($^{13}$CO) with $^{1}$Hβ or $^{1}$Hγ protons of Asx (aspartic acid/asparagine) or Glx (glutamic acid/glutamine), respectively. The $^{1}H^{\beta/\gamma}$-$^{13}$CO correlation spectra were measured with an adapted H(C)CO-type version[58,158,159] of the 2D ct-H(CA)CO experiment[160,161]. The $^{1}$H and $^{15}$N carriers are centred on the water frequency and 118 ppm, respectively. The $^{13}$C transmitter is

positioned at 179.5 ppm. Selective shaped pulses were applied with RE-BURP profile at aliphatic $^{13}C^{\beta/\gamma}$ (180º pulse, 2.28 ms, centred at 36.0 ppm covering a bandwidth of 10 ppm) and Q5 gaussian cascade at $^{13}CO$ (90º pulse, 1.8 ms, centred at 176.0 ppm). The spectra were acquired with 2048 ($^{1}H$) and 100 ($^{13}C$) complex points for a spectral width (SW) of 13 ppm ($^{1}H$) and 12 ppm ($^{13}C$), with 32 scans. NMR spectra were assigned by comparison to literature values[29,62,162].

NMR Intensity analysis. For [Ch][Glu] titration, 2D $^{1}H$-$^{15}N$ HSQC and complementary 2D $^{1}H$-$^{15}N$ H2(C)N and $^{1}H$-$^{13}C$ H(C)CO spectra were analysed concerning peak intensity. The assignment was prepared with CcpNmr2.5 and the integration was done with PINT[64,88] using a gaussian line shape fitting analysis, optimizing peak positions, line widths and intensity. Peak heights and peak volumes were extracted. The data were normalized according to $H_i/H_0$ (or $V_i/V_0$) where "0" denotes the reference – water.

GB1 backbone assignments. Protein backbone resonance assignment experiments (Bruker standard pulse sequences) were recorded for 0.6 mM $^{15}N^{13}C$ GB1 in water and at 1.5 M [Ch][Glu] (90% $H_2O$/10% $D_2O$, 0.1% $NaN_3$ and 50 $\mu$M DSS) on a 500 MHz Bruker Bruker NEO spectrometer. Another sample was prepared for 1.2 mM $^{15}N^{13}C$ GB1 at 3.5 M [Bmim][dca] (90% $H_2O$/10% $D_2O$, 0.1% $NaN_3$ and 100 $\mu$M DSS) to be measured at 600.13 MHz. Spectra were recorded in 3 mm tubes at 298.2 K and comprises 2D $^{1}H$-$^{15}N$ HSQC and 3D triple resonance experiments HNCACB, HNcoCACB/ CBCAcoNH, HNCO, HNcaCO spectra. All spectra were collected using Echo/Antiecho-TPPI gradient selection, which was efficient at suppressing signals from ionic liquid.

For folded GB1, acquisition parameters for 2D $^{1}H$-$^{15}N$ HSQC experiments were 2048 ($^{1}H$) and 256 ($^{15}N$) complex points for a SW of 13 ppm ($^{1}H$) and 37 ppm ($^{15}N$), with 16 scans. For 3D experiments, 2048 ($^{1}H$), 40 ($^{15}N$) and 128 ($^{13}C$) complex points for a SW of 13 ppm ($^{1}H$), 37 ppm ($^{15}N$) and 75 ppm for HNCACB/ CBCAcoNH or 18 ppm for HNCO ($^{13}C$), were used with 16 scans except HNCACB with 40 scans. For unfolded GB1, acquisition parameters for 2D $^{1}H$-$^{15}N$ HSQC experiments were 2048 ($^{1}H$) and 512 ($^{15}N$) complex points for a SW of 13 ppm ($^{1}H$) and 30 ppm ($^{15}N$), with 16 scans. For 3D experiments, 2048 ($^{1}H$), 40 ($^{15}N$) and 256 ($^{13}C$) complex points for a SW of 13 ppm ($^{1}H$), 30 ppm ($^{15}N$) and 75 ppm for HNCACB/ HNcoCACB or 20 ppm for HNCO/ HNcaCO ($^{13}C$), were used with 32 scans except HNCACB with 64 scans.

Non-uniform sampling (NUS) was used to optimize resolution of the indirect dimensions in the available experiment time, where just 20-25% of sparse data was recorded. NUS acquired data were processed with SMILE algorithm within NMRPipe for spectra reconstruction[163]. Sequential connectivities were performed with CcpNmr AnalysisAssign 3.0[164] followed by manual verification. Backbone assignments for folded GB1 in water/[Ch][Glu] are consistent with previous assignments[19,77,100]. The backbone assignment for GB1 $U_{[Bmim][dca]}$ state led to unambiguous assignment in the 2D $^{1}H$-$^{15}N$ HSQC even for very overlapped peaks.

Secondary structure propensity (SSP). SSPs were calculated using $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts of GB1 as input according to the SSP protocol by Marsh et al[82]. Positive SSP values ranging from 0 to 1 and negative values from 0 to -1 represent the propensities of $\alpha$

and β secondary structures, respectively. Reference values for random coil, secondary structure chemical shifts and standard deviations came from RefDB[165].

GB1 stability through methyl groups. A 0.3 mM $^{15}N^{13}C$ GB1 was prepared in 99.9% $D_2O$ and 50 $\mu$M TSP, and with 0.1 M sodium phosphate or 0.1 M, 0.25 M, 0.5 M, and 1.5 M of [Ch][Glu] ($pH_{read} \approx 7.5$, uncorrected for any isotope effect). A series of 2D $^{1}H,^{13}C$ HSQCs (hmqcphpr, standard Bruker pulse sequence) were measured at 298.2, 323.2, 328.2, 330.7 and from 333.2 to 353.2 in 2 K step with 5 min equilibrium time between each measurement to ensure temperature stability. NMR spectra were acquired on a Bruker Avance III 400 spectrometer equipped with a multinuclear direct detection probe head (BBO) a temperature control unit, and a pulse gradient unit capable of producing magnetic field pulsed gradients in the z direction of 56.0 G/cm operating at 400.15 MHz for $^{1}H$, and 100.62 MHz for $^{13}C$. Each 2D spectrum was recorded as a complex data matrix of 2048 ($^{1}H$) and 128 ($^{13}C$) points for a SW of 12 ppm ($^{1}H$) and 40 ppm ($^{13}C$), centred at 3.1 ppm and 25.0 ppm, respectively, with 8 scans per $t_1$ increment.

All time domain data were processed with Bruker TopSpin4.0. The temperature dependence of peak height was analysed with CcpNmr2.5. Side chain methyl group assignment was based on previous publications[99,100] and on backbone chemical shifts determined here (see above). Non-exposed side chain methyl groups were selected as probes for protein stability: A26 and A34 $^{13}C\beta$, and V54 $^{13}C\gamma$. For each experiment, a suitable temperature range was chosen to normalize the data. $T_m$ values were determined by fitting the intensity decay to a sigmoidal function using Prism 8 (GraphPad Software). The standard error was obtained from the fitting.

Self-Diffusion coefficients. The samples used to follow chemical shift perturbations in [Ch][Glu] were used to determine GB1, IL and water self-diffusion: 0.6 mM $^{15}N^{13}C$ GB1 (90% $H_2O$/10% $D_2O$, 0.1% $NaN_3$ and 50 $\mu$M DSS) in the presence of 0.1, 0.5, 0.75, 1.0, 1.5 M of [Ch][Glu] concentration, pH 7.2 ± 0.2. The experiments were measured at 600 MHz and 298.2 K in 3-mm tubes.

For water and [Ch][Glu] protons self-diffusion coefficients determination, DOSY (diffusion ordered spectroscopy) experiments were performed using the Bruker pulse sequence (ledbpgp2s) with a bipolar pulse longitudinal eddy current delay (BPPLED) for diffusion[166]. The duration of the magnetic field pulse gradients ($\delta$) was maintained constant throughout, and the diffusion times ($\Delta$) were optimized for each sample in order to obtain complete dephasing of the signals with the maximum gradient strength. Typically, in each PFG NMR (pulsed field gradient) experiment, a series of 32 BPPLED spectra on 16 K data points were collected, the values of $\Delta$ were 180 to 400 ms duration for water or 240 to 650 ms for [Ch][Glu] self-diffusion coefficients. In all experiments, the $\delta$ and the eddy current delay ($T_e$) were set to 1.5 ms and 5 ms. A sine shaped (shape factor $s = 0.637$) pulsed gradient ($g$) was incremented from 5 to 95 % of the maximum gradient strength (53.2 G/cm) in a linear ramp. The degree of attenuation is a function of the magnetic gradient pulse amplitude ($g$) and occurs at a rate proportional to the diffusion coefficient ($D$) of the molecule according to Stejskal-Tanner equation[167],

$$I(G) = I_0 \exp\left[-D(\gamma g \delta)^2 \left(\Delta - \frac{\delta}{3}\right)\right], \qquad [2.13]$$

where *I(G)* is the observed intensity, $I_0$ the reference intensity (unattenuated signal intensity), *D* the diffusion coefficient, $\gamma$ the gyromagnetic ratio of the observed nucleus ($\gamma/(2\pi)$ = 4257.64 Hz/G for protons), *g* the gradient strength, $\delta$ the length of the gradient, and $\Delta$ the diffusion time. The spectra were first processed in the F2 dimension by standard Fourier transform and baseline correction. Then the data were analysed using the variable gradient fitting routines in Bruker TopSpin4.0 software. The diffusion coefficients are calculated by exponential fitting of the experimental data using Prism 8 (GraphPad Software).

For GB1 protein translational diffusion determination, I used an adapted version of $^{15}$N-edited X-STE (heteronuclear stimulated echoes) diffusion experiment[119]. In this experiment, spatial encoding/decoding occurs on $^{1}$H spins with storage as longitudinal magnetisation on $^{15}$N during the diffusion delay. Thus, storage as $N_z$ in the stimulated echo not just provides isotope editing, but also takes advantage of the longer $^{15}$N $T_1$ to reduce longitudinal relaxation losses during the diffusion delay $\Delta$. Most importantly, the X-STE is not sensitive to amide hydrogen exchange during long diffusion delays.

$^{15}$N-edited X-STE measurements were acquired with 32,768 transients using a $^{1}$H stimulated echo Δ from 165 to 400 ms and sine gradient pulses with a total length, $\delta$, of 3 ms. The INEPT delays were $\tau$ = 2.72 ms ≈ $|4J_{HN}|^{-1}$ ($J_{HN}$ ≈ - 92 Hz). Water flip-back scheme (selective sinc-shaped 90º pulses, 1.2 ms) was used to minimize water saturation during the pulse sequence, and WATERGATE method was employed to supress the solvent signal. Echoes were acquired in 64 scans (except for 1.5 M [Ch][Glu] with 128 scans) at 64 gradient strengths ranging linearly from 5 to 95% of the maximum gradient. NMR data were processed with a linear baseline using the variable gradient fitting routines in Bruker TopSpin4.0 software. 4 sections from amide envelope were extracted and integrated. Using Prism 8 (GraphPad Software), the data were fitted according to[119,168],

$$I(G) = I_0 \exp[-D(\gamma g \delta)^2(\Delta + 6\tau)]. \quad [2.14]$$

Longitudinal relaxation of water. $R_{1,water}$ was measured using an adapted version of saturation-inversion-recovery (SIR) sequence[133]. The SIR sequence is the 3-pulse sequence: 90º – τ – 180º – τ – 90º – acquisition. To avoid radiation damping, a weak gradient (5 %) was applied during recovery delay, τ[169]. The SIR measurements were acquired with 16 τ-delays in the range 0.002 to 15 s (optimized for each sample), a saturation delay of 5 ms (immediately before to the first 90º pulse, without saturation pulse train) and a relaxation delay of 6 s to ensure complete relaxation of the spin magnetization between each scan. Data were analysed using the relaxation fitting routines in Bruker TopSpin4.0 software, and integrals were exported and fitted in Prism 8 (GraphPad Software) according to exponential inversion-recovery,

$$I(\tau) = I_0\left[1 - 2\exp\left(-R_{1,water}\tau\right)\right] \quad [2.15]$$

CLEANEX-PM experiments. Values of $k_{ex}$ were obtained from (CLEANEX-PM)-FHSQC (fhsqccxf3gpph) experiments[101] for the samples above mentioned at 298.2 K and 310.2 K in a 500 MHz Bruker NEO spectrometer operating at 500.34 MHz for $^{1}$H and 50.70 MHz for $^{15}$N. By applying Phase-Modulated CLEAN chemical EXchange (CLEANEX-PM) spin-locking module during mixing period in $^{1}$H-$^{15}$N FHSQC (fhsqcf3gpph) sequence[170], common artifacts can be eliminated (i.e., intramolecular NOEs, and/or TOCSY transfer from

C$\alpha$ protons coincident with the water frequency, or exchange-relayed NOEs from fast exchanging hydroxyl or amine protons), revealing an unambiguous water–NH exchange spectrum.

A 7.5 ms Gaussian shaped selective pulse was applied to selectively perturb water magnetization at the start of the scheme. Each 2D data set was recorded with 16 transients/FID (except for 1 M sample at 298 K, 1.5 M at 310 K and 298 K with 32, 32 and 64 scans, respectively), a relaxation delay of 2 s, and (2048, 128) complex points in ($t_2$, $t_1$). The series of measurements using 16 different $\tau_{mix}$ values in the range 1.8–100 ms (i.e., 1.8, 5.1, 9.9, 12.4, 15.0, 20.1, 25.2, 29.9, 35.0, 40.1, 50.0, 59.9, 70.1, 90.1, 100.0 ms, the 5.1 ms experiment was recorded in duplicate) were chosen in order to measure the amide $^1H^N$ magnetization build-up curve. The data was acquired as single scan interleaving manner to minimize the effects of sample heating. A separate reference 2D plane (FHSQC spectrum) was recorded using the same experimental settings as in the CLEANEX sequence, to estimate the equilibrium magnetization at each site. An additional 1D version of CLEANEX-PM)-FHSQC experiment was implemented to find the saturation degree of water magnetization. Two experiments with recycling delay times of 2 s and 30 s, and 16 dummy scans were used to go through the entire sequence with 5.84 ms mixing. Then the area of the water peak[169] immediately prior to mixing was measured. The ratio of the water area at 2 s (partially saturated $H_2O$ signal) and 30 s (maximum water signal) of recycling delay corresponds to the fraction of unsaturated water.

The dependence of peak volumes on mixing time ($\tau_{mix}$) is given by the following equation,

$$\frac{V(\tau_{mix})}{V_0} = \frac{f k_{ex}}{k_{ex} + R_{H_N} - R_w}\left[e^{-R_w \tau_{mix}} - e^{-(k_{ex}+R_{H_N})\,\tau_{mix}}\right], \qquad [2.16]$$

where $V$ is the peak volume of an amide correlation as a function of mixing time $\tau_{mix}$, $V_0$ is the intensity of the corresponding peak in a fully relaxed spectrum. $R_w$ and $R_{HN}$ parameters are the water and amide relaxation rate during the CLEANEX-PM element, respectively. The water preservation factor corresponds to $f$ = (1/fraction of unsaturated water). At short mixing time, $k_{ex}$ in Eq. 2.16 reflects the initial slope ($V/V_0$ vs $\tau_{mix}$) and can be approximated by,

$$\frac{V(\tau_{mix})}{V_0} = f k_{ex}\, \tau_{mix}, \qquad [2.17]$$

where the exchange rate obtained can be corrected by simply multiplying $k_{ex}$ by $f$, assuming that the saturation of water leads to the decrease in intensity for pure exchange peaks in the same proportion. The $f$ for the different samples, water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu], obtained for 298.2 K and 310.2 K measurements were $f$ = 0.74, 0.51, 0.62, 0.66, 0.79, respectively, and $f$ = 0.61, 0.40, 0.47, 0.52, 0.67, respectively.

The spectrum originally recorded in pseudo-3D fashion (each plane corresponds to a different $\tau_{mix}$) was processed in NMRPipe and assigned with CcpNmr2.5. The quantitation of the peak volumes were performed within PINT[64,88], assuming Gaussian peak shapes. The data with a linear behaviour up to 20-50 ms of mixing time were normalized and fitted with a simple linear regression in Prism 8.

Backbone relaxation. Samples were prepared with 0.6 mM $^{15}N$ or $^{13}C^{15}N$ enriched GB1 for water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentrations (analogous to experiments above described). All relaxation NMR experiments were performed at two $^{1}H$ Larmor precessions frequencies (500.34 and 600.13 MHz) on Bruker NEO 500 and AVANCE II+ 600 NMR spectrometers equipped with triple-resonance probes. Spin-lattice ($R_1$) and spin-spin relaxation ($R_2$) rate constants and {$^{1}H$}-$^{15}N$ steady-state heteronuclear Overhauser enhancements (NOE) were measured using two-dimensional pulse field gradient sensitivity-enhanced experiments[155] as described[171]. Standard Bruker pulse sequences were used to collect $R_1$, $R_2$ and NOE data (hsqct1etf3gpsi3d.2, hsqct2etf3gpsi3d, hsqcnoef3gpsi). The three types of experiments were performed at 298.2 and 310.2 K for the two magnetic fields. Additionally, water and 1.5 M [Ch][Glu] samples were also measured for other two temperatures, 283.2 and 323.2 K, at 600.13 MHz. The sample temperature was calibrated using a methanol standard and regulated by the Bruker VT unit.

All spectra were acquired with spectral widths of 7812 Hz and 2250 Hz in the $^{1}H$ and $^{15}N$ dimensions, respectively. $R_1$, $R_2$ spectra were acquired as 2048 x 128 complex matrices with 4 to 8 scans per $t_1$, depending on the sample. List of 12-16 relaxation delays were adjusted for each experiment, a typical range for $R_1$ and $R_2$ is between 0.01 to 1.20 s and from 0.02 to 0.42 s, respectively. A recycle delay of 1.5 s was employed when recording these experiments. The $R_2$ data was acquired as single scan interleaving manner to minimize the effects of sample heating. To permit the estimation of noise levels, 1-2 duplicate spectra were recorded in each experiment. NOE spectra were acquired as 2048 x 512 complex matrices with 16 to 32 scans. Steady-state $^{1}H$-$^{15}N$ NOE values were determined from spectra recorded in the presence and absence of proton saturation. Saturation was achieved by the application of $^{1}H$ 120° pulses every 5 ms during the 5 s recycle period.

All data were processed using NMRPipe10.1[152]. Resonances were picked with CcpNmr2.5[153] and the integration of peak intensities and extraction of relaxation rates were performed with PINT[88]. Peak volumes were fitted to a simple exponential function [$I(t)=I_0 \exp(R_{1,2}\,\tau)$] using curve fitting routines (*-fitEXP* flag). Heteronuclear NOE was calculated with another routine (*-calcNOE* flag) which corresponds to the ratio of peak volume with and without proton saturation ($NOE = V_{sat}/V_{unsat}$). Errors in the parameters are estimated with the jackknife resampling method within the software package. The relaxation dataset measured at 298.2 K and 310.2 K of temperature (and 600.13 MHz of magnetic field strength) can be found in the Table A7 and A8 of Appendix A, respectively.

NMR-Derived order parameters, diffusion tensor and conformational entropy. The relation between order parameters derived from NMR spin relaxation experiments and the contribution to conformational entropy was investigated.

Order parameters ($S^2$), rotation correlation times ($\tau_m$), conformation exchange ($R_{ex}$) and effective internal correlation time ($\tau_e$) were obtained from an automated model-free analysis conducted by relax[134–136]. Spin relaxation and NOE data without residue exclusion, collected at the two fields strengths, were applied in relax for the different conditions at 298.2 and 310.2 K of temperature. Values for the $^{15}N$ chemical shift anisotropy and N-H bond length were set at -172 ppm and 1.02 Å, respectively. The spin system information was extracted from PDB 2JSV.pdb. The fully automated protocol mode was used with 11 grid search

increments, 500 monte carlo simulation and 30 maximum interactions. The automatic mode in which relaxGUI[136] fully controls the calculation were chosen, it starts with the optimization of the 9 local $\tau_m$ models for each residue/spin system as by the new model-free protocol previously described[135]. After this calculation, relaxGUI automatically starts the calculation of the different global models of the spherical, prolate spheroidal, oblate spheroidal and ellipsoidal diffusion tensors. The different diffusion models are then iteratively optimized until convergence. Finally, relaxGUI performs model selection for the best global model. All the calculated internal mobility parameters ($S^2$, $\tau_e$, $R_{ex}$, and the dynamic model used to fit the data) can be found in the Table A9 and A10 of Appendix A for 298.2 K and 310.2 K dataset, respectively.

The entropy ($S_p$) associated with angular fluctuations of each NH group at each condition can be calculated from the order parameter according to the equation[141],

$$S_{conf} = k_B ln\left\{\pi\left[3 - (1 + 8\,S)^{1/2}\right]\right\}, 0 \leq S^2 \leq 1, \quad [2.18]$$

in which S is the generalized order parameter and $k_B$ is Boltzmann's constant.

Amide proton exchange. Before hydrogen exchange (HX) experiments, the spectrometer was first tuned and shimmed with a sample with the same contents from those in the sample to be measured but without the protein. The first sample for the exchange experiment was prepared by dissolving 4 mg of lyophilized $^{15}$N-labelled GB1 in 600 $\mu$L of $D_2O$ "100%" (> 99,96% isotopic purity, Cambridge Isotope Laboratories) to a final concentration of 1.0 mM. Samples for [Ch][Glu] titration were prepared by dissolving the appropriate quantity of [Ch][Glu] "dry" (previously exchanged in $D_2O$ and lyophilized in at least 3 cycles) with $D_2O$ "100%" to final concentrations of 0.1, 0.75, 1.0, and 1.5 M of [Ch][Glu] ($pH_{read}$ ~ 6.8) in 600 $\mu$L of final volume. Then, 4 mg of lyophilized $^{15}$N-GB1 was dissolved in this [Ch][Glu]-aqueous solution. An additional sample, 0.1 M [Ch][Glu] at pH ~ 6 was prepared to confirm EX2 mechanism (see below). All samples were pre-equilibrated at the desired temperature. Unless otherwise stated, pH readings are uncorrected for deuterium isotope effects[91].

The exchange rates of GB1 were measured at two temperatures, 298.2 and 310.2 K. The samples for these experiments comprise pH readings about 6.7 ± 0.2 and 6.9 ± 0.1, respectively, which corresponds to a pH correct ($pD_{corr}$) of 7.1 ± 0.2 and 7.3 ± 0.1, accordingly to $pD_{corr} = pH_{read} + 0.4$[91]. For each experiment, the decay of the amide proton signal intensities due to hydrogen exchange with $D_2O$ was followed by a series of 105-110 standard $^1$H-$^{15}$N HSQC spectra over an 8-hour period. Spectra were acquired at 600.13 MHz with 2048 x 128 complex points, recycle delay of 1 s in a spectral window of 8417 Hz (centred on water) x 2311 (centred at 118 ppm) in $^1$H and $^{15}$N sweep's width, respectively. Consecutive spectra were collected with the increasing number of scans, due to the loss of signal intensity and consequent decrease of the signal/ noise ratio. In this way, acceptable signal/noise ratios were maintained over the experiment. Typically, the first spectrum is acquired within ~ 4 mins (~ 2 min of protein dissolution) and the following number of scans for each $t_1$ increment were used: 90 spectra with 1 scan, 10 with 8 scans and 5-10 with 32 scans.

Processing was performed with NMRPipe10.1[152]. Resonances were picked for first spectrum in each experiment with CcpNmr2.5[153]. Peak integration and downstream fitting was carried out with PINT[64,88]. For each amide proton, cross-peak height intensity was

extracted and normalized over the number of scans of each $^{1}H$-$^{15}N$ HSQC spectrum. Rates of exchange, $k_{ex}$, were obtained from plots of peak intensity against exchange time fitted to a single exponential function ($I(t)=I_0$ exp ($-k_{ex}$ $t$) + Plateau) with Prism 8. The exchange time was defined as the period from protein dissolution in $D_2O$ solution to the midpoint of the estimate taken to acquire each spectrum.

Free energies of exchange. The observed rate of amide proton (N–H) exchange, $k_{ex}$, is related to equilibrium stability by considering a protein in which each N–H exists in an open (exposed, exchange-competent) state, or a closed (protected, exchange-incompetent)[85,89]:

$$closed\ (N-H) \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} open\ (N-H) \xrightarrow{k_{int}} open\ (N-D). \qquad [2.19]$$

Each position opens and closes with rate constants, $k_{op}$ and $k_{cl}$ (equilibrium constant for structural opening by $K_{op} = k_{op}/k_{cl}$), and exchange from the open state occurs with intrinsic rate constant, $k_{int}$. Values for $k_{int}$ are based on exchange data from unstructured polypeptides based on primary amino acid sequence and ambient conditions[87,172]. If the protein is stable (i.e., $k_{cl} \gg k_{op}$), the exchange occurs within two limits[85]. At the so-called EX1 limit closing is rate determining ($k_{cl} \ll k_{int}$ , $k_{ex} = k_{op}$) while at the usually observed EX2 kinetics, the exchange from the open state is rate limiting, $k_{cl} \gg k_{int}$. According to Orban et al[173], deuterium exchange process on GB1 occurs by EX2 mechanism as most of globular proteins under stabilising conditions. As the reclosing is faster than the exchange, the structural opening reaction appears as a pre-equilibrium step prior to the rate-limiting chemical exchange, and the observed rate constant ($k_{ex}$) is $k_{ex} = K_{op}\, k_{int}$. The amide proton protection factors ($Pf$) can be estimated by $Pf = k_{int}/k_{ex}$. With the values from the exchange rate from the open state available ($k_{int}$), stability is measured as the free energy of exchange or opening $\Delta G_{HX}$[84],

$$\Delta G_{HX} = -RTln(K_{op}) = -RTln\left(\frac{k_{ex}}{k_{int}}\right), \qquad [2.20]$$

where RT is the molar gas constant multiplied by the absolute temperature (K). Values for $k_{int}$ were determined using the online server program (https://protocol.fccc.edu/research/labs/roder/sphere/sphere.html) for hydrogen exchange rate estimation, SPHERE[87], with default activation energies: $E_{acid}$ = 15 kcal/mol, $E_{base}$ = 2.6 kcal/mol. The higher the value of $\Delta G_{HX}$ the more protected is the residue for exchange. Eq. 2.20 was also used with $k_{ex}$ values from CLEANEX experiments, obtaining $\Delta G_{CLEANEX}$ values. Changes in stability caused by [Ch][Glu] IL were calculated by $\Delta\Delta G_{HX/\ CLEANEX} = \Delta G_{HX/\ CLEANEX,\ cosolute} - \Delta G_{HX/\ CLEANEX,\ water}$.

Protein-water interactions. Aqueous uniformly $^{2}H$,$^{15}N$,$^{13}C$-labelled GB1 samples were composed of 0.65 mM protein, 10% $D_2O$, 0.1% $NaN_3$, 50 $\mu$M TSP, pH 7.2 in the presence or absence of 1 M [Ch][Glu]. NMR experiments were performed at 298.2 K using Bruker Avance II+ NMR spectrometer operating at 600.13 MHz equipped with a cryogenically cooled probe. Three-dimensional sensitivity-enhanced $^{15}N$-resolved NOESY-HSQC[174,175] (noesyhsqcfpf3gpsi3d) and $^{15}N$-resolved ROESY-HSQC[176] (modified from standard Bruker TOCSY-HSQC, dipsihsqcf3gpsi3d) were obtained on the same sample. The ROESY experiment employed an 8.6 kHz CW spin-lock field. The NOE/ROE mixing time was 40 ms with recycle delay of 1 s. The spectra were collected with 8 transients, 40 complex increments with

gradient selection for quadrature in the indirect $^{15}N$ dimension, and 128 complex increments with States-TPPI quadrature detection in the indirect $^{1}H$ dimension. The data were processed in TopSpin4.0 and $^{1}H$-$^{15}N$ plane at the water $^{1}H$ resonance was extracted and analysed with CcpNmr2.5.

ROESY is analogous to NOESY experiment, except that instead of generating cross-peaks by cross relaxation between the *z*-magnetization of different spins, the cross-peaks in ROESY arise from cross-relaxation between spin-locked transverse magnetization. Dipolar exchange is identified by the opposite phase of the cross-peaks in the orthogonal frames (when diagonal is positive: positive NOE, negative ROE), whereas direct hydrogen exchange with solvent gives rise to peaks of identical phase (positive in both spectra). NOE obtained at the 40-ms mixing time used here placed it in the linear regime and corresponds to a maximum NOE-detected distance of approximately 4.3 Å. These cross-peaks can therefore be unequivocally assigned to direct NOE interactions between the hydration water and amide hydrogens of the protein[109]. Amide hydrogens that showed a cross peak to water with sufficient resolution were qualitatively compared in NOESY and ROESY $^{1}H$-$^{15}N$ plane at the water $^{1}H$ resonance.

# References


1. Fujita, K., MacFarlane, D. R. & Forsyth, M. Protein solubilising and stabilising ionic liquids. *Chem. Commun.* **70**, 4804 (2005).
2. Byrne, N., Wang, L.-M., Belieres, J. & Angell, C. A. Reversible folding–unfolding, aggregation protection, and multi-year stabilization, in high concentration protein solutions, using ionic liquids. *Chem. Commun.* 2714–2716 (2007).
3. Constantinescu, D., Weingärtner, H. & Herrmann, C. Protein Denaturation by Ionic Liquids and the Hofmeister Series: A Case Study of Aqueous Solutions of Ribonuclease A. *Angew. Chemie Int. Ed.* **46**, 8887–8889 (2007).
4. Constatinescu, D., Herrmann, C. & Weingärtner, H. Patterns of protein unfolding and protein aggregation in ionic liquids. *Phys. Chem. Chem. Phys.* **12**, 1756 (2010).
5. Weingärtner, H., Cabrele, C. & Herrmann, C. How ionic liquids can help to stabilize native proteins. *Phys. Chem. Chem. Phys.* **14**, 415–426 (2012).
6. Figueiredo, A. M., Sardinha, J., Moore, G. R. & Cabrita, E. J. Protein destabilisation in ionic liquids: the role of preferential interactions in denaturation. *Phys. Chem. Chem. Phys.* **15**, 19632 (2013).
7. Yang, Z. Hofmeister effects: an explanation for the impact of ionic liquids on biocatalysis. *J. Biotechnol.* **144**, 12–22 (2009).
8. Zhang, Y. & Cremer, P. S. The inverse and direct Hofmeister series for lysozyme. *Proc. Natl. Acad. Sci.* **106**, 15249–15253 (2009).
9. Naushad, M., ALOthman, Z. A., Khan, A. B. & Ali, M. Effect of ionic liquid on activity, stability, and structure of enzymes: A review. *Int. J. Biol. Macromol.* **51**, 555–560 (2012).
10. Senske, M. *et al.* The temperature dependence of the Hofmeister series: thermodynamic fingerprints of cosolute–protein interactions. *Phys. Chem. Chem. Phys.* **18**, 29698–29708 (2016).
11. Zhao, H. Protein stabilization and enzyme activation in ionic liquids: specific ion effects. *J. Chem. Technol. Biotechnol.* **91**, 25–50 (2016).
12. Shukla, S. K. & Mikkola, J.-P. Use of Ionic Liquids in Protein and DNA Chemistry. *Front. Chem.* **8**, 1–23 (2020).
13. Jungwirth, P. & Cremer, P. S. Beyond Hofmeister. *Nat. Chem.* **6**, 261–263 (2014).
14. Okur, H. I. *et al.* Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **121**, 1997–2014 (2017).
15. Bui-Le, L. *et al.* Revealing the complexity of ionic liquid–protein interactions through a multi-technique investigation. *Commun. Chem.* **3**, 55 (2020).
16. Rodrigues, J. V, Prosinecki, V., Marrucho, I., Rebelo, L. P. N. & Gomes, C. M. Protein stability in an ionic liquid milieu: on the use of differential scanning fluorimetry. *Phys. Chem. Chem. Phys.* **13**, 13614 (2011).
17. Pegram, L. M. *et al.* Why Hofmeister effects of many salts favor protein folding but not DNA helix formation. *Proc. Natl. Acad. Sci.* **107**, 7716–7721 (2010).
18. Sukenik, S., Sapir, L., Gilman-Politi, R. & Harries, D. Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday Discuss.* **160**, 225–237 (2013).
19. Gronenborn, A. *et al.* A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science (80-. ).* **253**, 657–661 (1991).
20. Alexander, P., Fahnestock, S., Lee, T., Orban, J. & Bryan, P. Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. *Biochemistry* **31**, 3597–3603 (1992).
21. Franks, W. T. *et al.* Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. *Proc. Natl. Acad. Sci.* **105**, 4621–4626 (2008).
22. Kuszewski, J., Clore, G. M. & Gronenborn, A. M. Fast folding of a prototypic polypeptide: the immunoglobulin binding domain of streptococcal protein G. *Protein Sci.* **3**, 1945–1952 (1994).
23. Orban, J., Alexander, P., Bryan, P. & Khare, D. Assessment of Stability Differences in the Protein G B1 and B2 Domains From Hydrogen-Deuterium Exchange: Comparison with Calorimetric Data. *Biochemistry* **34**, 15291–15300 (1995).
24. Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W. & Sosnick, T. R. Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J. Mol. Biol.* **324**, 359–371 (2002).
25. Ding, K., Louis, J. M. & Gronenborn, A. M. Insights into Conformation and Dynamics of Protein GB1 During Folding and Unfolding by NMR. *J. Mol. Biol.* **335**, 1299–1307 (2004).
26. Monteith, W. B. & Pielak, G. J. Residue level quantification of protein stability in living cells. *Proc. Natl. Acad. Sci.* **111**, 11335–11340 (2014).
27. Spitzer, J. & Poolman, B. The Role of Biomacromolecular Crowding, Ionic Strength, and Physicochemical Gradients in the Complexities of Life's Emergence. *Microbiol. Mol. Biol. Rev.* **73**, 371–388 (2009).
28. Lindman, S. *et al.* Salting the Charged Surface: pH and Salt Dependence of Protein G B1 Stability. *Biophys. J.* **90**, 2911–2921 (2006).
29. Lindman, S., Linse, S., Mulder, F. A. A. & André, I. pKa Values for Side-Chain Carboxyl Groups of a PGB1 Variant Explain Salt and pH-Dependent Stability. *Biophys. J.* **92**, 257–266 (2007).

30. Park, S. H., Shastry, M. C. R. & Roder, H. Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. *Nat. Struct. Biol.* **6**, 943–947 (1999).

31. Shimada, J. & Shakhnovich, E. I. The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation. *Proc. Natl. Acad. Sci.* **99**, 11175–11180 (2002).

32. Morrone, A. *et al.* GB1 Is Not a Two-State Folder: Identification and Characterization of an On-Pathway Intermediate. *Biophys. J.* **101**, 2053–2060 (2011).

33. McCallister, E. L., Alm, E. & Baker, D. Critical role of β-hairpin formation in protein G folding. *Nat. Struct. Biol.* **7**, 669–673 (2000).

34. Blanco, F. J., Ortiz, A. R. & Serrano, L. Role of a nonnative interaction in the folding of the protein G B1 domain as inferred from the conformational analysis of the α-helix fragment. *Fold. Des.* **2**, 123–133 (1997).

35. Cohen, R. D. & Pielak, G. J. Quinary interactions with an unfolded state ensemble. *Protein Sci.* **26**, 1698–1703 (2017).

36. Timasheff, S. N. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci.* **99**, 9721–9726 (2002).

37. Tomlinson, J. H., Green, V. L., Baker, P. J. & Williamson, M. P. Structural origins of pH-dependent chemical shifts in the B1 domain of protein G. *Proteins Struct. Funct. Bioinforma.* **78**, 3000–3016 (2010).

38. Zbacnik, T. J. *et al.* Role of Buffers in Protein Formulations. *J. Pharm. Sci.* **106**, 713–733 (2017).

39. Buchfink, R., Tischer, A., Patil, G., Rudolph, R. & Lange, C. Ionic liquids as refolding additives: Variation of the anion. *J. Biotechnol.* **150**, 64–72 (2010).

40. Kowacz, M. *et al.* Hofmeister effects of ionic liquids in protein crystallization: Direct and water-mediated interactions. *CrystEngComm* **14**, 4912 (2012).

41. Silva, M. S., Figueiredo, A. M. & Cabrita, E. J. Design of bio-inspired ionic liquids for protein stabilisation. (Universidade Nova de Lisboa, 2015).

42. Vrbka, L., Jungwirth, P., Bauduin, P., Touraud, D. & Kunz, W. Specific Ion Effects at Protein Surfaces: A Molecular Dynamics Study of Bovine Pancreatic Trypsin Inhibitor and Horseradish Peroxidase in Selected Salt Solutions. *J. Phys. Chem. B* **110**, 7036–7043 (2006).

43. Mason, P. E. *et al.* Specificity of Ion−Protein Interactions: Complementary and Competitive Effects of Tetrapropylammonium, Guanidinium, Sulfate, and Chloride Ions. *J. Phys. Chem. B* **113**, 3227–3234 (2009).

44. Cheng, X. *et al.* Basis of Protein Stabilization by K Glutamate: Unfavorable Interactions with Carbon, Oxygen Groups. *Biophys. J.* **111**, 1854–1865 (2016).

45. Collins, K. D. Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Methods* **34**, 300–311 (2004).

46. Dempsey, C. E., Mason, P. E., Brady, J. W. & Neilson, G. W. The reversal by sulfate of the denaturant activity of guanidinium. *J. Am. Chem. Soc.* **129**, 15895–15902 (2007).

47. Lockhart, D. J. & Kim, P. S. Electrostatic screening of charge and dipole interactions with the helix backbone. *Science (80-. ).* **260**, 198–202 (1993).

48. Bauer, M., Xue, W.-F. & Linse, S. Protein GB1 Folding and Assembly from Structural Elements. *Int. J. Mol. Sci.* **10**, 1552–1566 (2009).

49. MacFarlane, D. R. *et al.* Ionic liquid "buffers"—pH control in ionic liquid systems. *Chem. Commun.* **46**, 7703 (2010).

50. MacFarlane, D. R., Pringle, J. M., Johansson, K. M., Forsyth, S. A. & Forsyth, M. Lewis base ionic liquids. *Chem. Commun.* 1905 (2006). doi:10.1039/b516961p

51. Lange, C., Patil, G. & Rudolph, R. Ionic liquids as refolding additives: N'-alkyl and N'-(ω-hydroxyalkyl) N-methylimidazolium chlorides. *Protein Sci.* **14**, 2693–2701 (2005).

52. Weaver, K. D. *et al.* Structure and function of proteins in hydrated choline dihydrogen phosphate ionic liquid. *Phys. Chem. Chem. Phys.* **14**, 790 (2012).

53. Singh, O., Lee, P.-Y., Matysiak, S. & Bermudez, H. Dual mechanism of ionic liquid-induced protein unfolding. *Phys. Chem. Chem. Phys.* **22**, 19779–19786 (2020).

54. Pantoliano, M. W. *et al.* High-Density Miniaturized Thermal Shift Assays as a General Strategy for Drug Discovery. *J. Biomol. Screen.* **6**, 429–440 (2001).

55. Salis, A. & Ninham, B. W. Models and mechanisms of Hofmeister effects in electrolyte solutions, and colloid and protein systems revisited. *Chem. Soc. Rev.* **43**, 7358–7377 (2014).

56. Silva, M., Figueiredo, A. M. & Cabrita, E. J. Epitope mapping of imidazolium cations in ionic liquid–protein interactions unveils the balance between hydrophobicity and electrostatics towards protein destabilisation. *Phys. Chem. Chem. Phys.* **16**, 23394–23403 (2014).

57. André, I., Linse, S. & Mulder, F. A. A. Residue-Specific p K a Determination of Lysine and Arginine Side Chains by Indirect 15 N and 13 C NMR Spectroscopy: Application to apo Calmodulin. *J. Am. Chem. Soc.* **129**, 15805–15813 (2007).

58. Pellecchia, M., Iwai, H., Szyperski, T. & Wüthrich, K. The 2D NMR Experiments H(C)CO2 and HCCO2 for Assignment and pH Titration of Carboxylate Groups in Uniformly 15N/13C-Labeled Proteins. *J. Magn. Reson.* **124**, 274–278 (1997).

59. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).

60. Schumann, F. H. *et al.* Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. *J. Biomol. NMR* **39**, 275–289 (2007).

61. Gallagher, T., Alexander, P., Bryan, P. & Gilliland, G. L. Two Crystal Structures of the B1 Immunoglobulin-Binding Domain of Streptococcal Protein G and Comparison with NMR. *Biochemistry* **33**, 4721–4729 (1994).

62. Tomlinson, J. H., Ullah, S., Hansen, P. E. & Williamson, M. P. Characterization of Salt Bridges to Lysines in the Protein G B1 Domain. *J. Am. Chem. Soc.* **131**, 4674–4684 (2009).

63. Palmer, A. G., Kroenke, C. D. & Patrick Loria, J. Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules. in *Methods in Enzymology* **339**, 204–238 (Elsevier Masson SAS, 2001).

64. Niklasson, M. *et al.* Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* **69**, 93–99 (2017).

65. Faustino, A. F. *et al.* Fast NMR method to probe solvent accessibility and disordered regions in proteins. *Sci. Rep.* **9**, 1647 (2019).

66. Wishart, D. S., Sykes, B. D. & Richards, F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311–333 (1991).

67. Zhou, N. E., Zhu, B. Y., Sykes, B. D. & Hodges, R. S. Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic .alpha.-helical peptides. *J. Am. Chem. Soc.* **114**, 4320–4326 (1992).

68. Frericks Schmidt, H. L. *et al.* Crystal Polymorphism of Protein GB1 Examined by Solid-State NMR Spectroscopy and X-ray Diffraction. *J. Phys. Chem. B* **111**, 14362–14369 (2007).

69. Monteith, W. B., Cohen, R. D., Smith, A. E., Guzman-Cisneros, E. & Pielak, G. J. Quinary structure modulates protein stability in cells. *Proc. Natl. Acad. Sci.* **112**, 1739–1742 (2015).

70. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).

71. Baxter, N. J. & Williamson, M. P. Temperature dependence of 1H chemical shifts in proteins. *J. Biomol. NMR* **9**, 359–369 (1997).

72. Cierpicki, T. & Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **21**, 249–261 (2001).

73. Tomlinson, J. H. & Williamson, M. P. Amide temperature coefficients in the protein G B1 domain. *J. Biomol. NMR* **52**, 57–64 (2012).

74. Khare, D. *et al.* p K a Measurements from Nuclear Magnetic Resonance for the B1 and B2 Immunoglobulin G-Binding Domains of Protein G: Comparison with Calculated Values for Nuclear Magnetic Resonance and X-ray Structures †. *Biochemistry* **36**, 3580–3589 (1997).

75. Tomlinson, J. H. *et al.* Dimerization of protein G B1 domain at low pH: A conformational switch caused by loss of a single hydrogen bond. *Proteins Struct. Funct. Bioinforma.* **78**, 1642–1661 (2010).

76. Gronenborn, A. M., Frank, M. K. & Clore, G. M. Core mutants of the immunoglobulin binding domain of streptococcal protein G: Stability and structural integrity. *FEBS Lett.* **398**, 312–316 (1996).

77. Warner, L. *et al.* The effects of high concentrations of ionic liquid on GB1 protein structure and dynamics probed by high-resolution magic-angle-spinning NMR spectroscopy. *Biochem. Biophys. Reports* **8**, 75–80 (2016).

78. Frank, M. K., Clore, G. M. & Gronenborn, A. M. Structural and dynamic characterization of the urea denatured state of the immunoglobulin binding domain of streptococcal protein G by multidimensional heteronuclear NMR spectroscopy. *Protein Sci.* **4**, 2605–2615 (1995).

79. Sapir, L. & Harries, D. Macromolecular Stabilization by Excluded Cosolutes: Mean Field Theory of Crowded Solutions. *J. Chem. Theory Comput.* **11**, 3478–3490 (2015).

80. Dyson, H. J. & Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev.* **104**, 3607–3622 (2004).

81. Wishart, D. S. & Case, D. A. Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* **338**, 3–34 (2001).

82. Marsh, J. A., Singh, V. K., Jia, Z. & Forman-Kay, J. D. Sensitivity of secondary structure propensities to sequence differences between $\alpha$- and $\gamma$-synuclein: Implications for fibrillation. *Protein Sci.* **15**, 2795–2804 (2006).

83. Ball, P. & Hallsworth, J. E. Water structure and chaotropicity: their uses, abuses and biological implications. *Phys. Chem. Chem. Phys.* **17**, 8297–8305 (2015).

84. Englander, S. W. & Mayne, L. Protein Folding Studied Using Hydrogen-Exchange Labeling and Two-Dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 243–265 (1992).

85. Hvidt, A. & Nielsen, S. O. Hydrogen Exchange in Proteins. in *Advances in Protein Chemistry* **21**, 287–386 (1966).

86. Englander, S. W. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521–655 (1983).

87. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **17**, 75–86 (1993).

88. Ahlner, A., Carlsson, M., Jonsson, B.-H. H. & Lundström, P. PINT: A software for integration of peak volumes and extraction of relaxation rates. *J. Biomol. NMR* **56**, 191–202 (2013).

89. Krishna, M. M. G., Hoang, L., Lin, Y. & Englander, S. W. Hydrogen exchange methods to study protein folding. *Methods* **34**, 51–64 (2004).

90. Barchi, J. J., Grasberger, B., Gronenborn, A. M. & Clore, G. M. Investigation of the backbone dynamics of the igg-binding domain of streptococcal protein g by heteronuclear two-dimensional 1 H- 15 N nuclear magnetic resonance spectroscopy. *Protein Sci.* **3**, 15–21 (1994).

91. Glasoe, P. K. & Long, F. A. Use of glass electrodes to measured acidities in deuterium oxide. *J. Phys. Chem.* **64**, 188–190 (1960).

92. Selenko, P., Serber, Z., Gadea, B., Ruderman, J. & Wagner, G. Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. *Proc. Natl. Acad. Sci.* **103**, 11904–11909 (2006).

93. Cavallo, L. POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level. *Nucleic Acids Res.* **31**, 3364–3366 (2003).

94. Englander, S. W., Sosnick, T. R., Englander, J. J. & Mayne, L. Mechanisms and uses of hydrogen exchange. *Curr. Opin. Struct. Biol.* **6**, 18–23 (1996).

95. Bai, Y., Sosnick, T., Mayne, L. & Englander, S. Protein folding intermediates: native-state hydrogen exchange. *Science (80-. ).* **269**, 192–197 (1995).

96. Krantz, B. a *et al.* Understanding protein hydrogen bond formation with kinetic H/D amide isotope effects. *Nat. Struct. Biol.* **9**, 458–463 (2002).

97. Skinner, J. J., Lim, W. K., Bédard, S., Black, B. E. & Englander, S. W. Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **21**, 996–1005 (2012).

98. Becktel, W. J. & Schellman, J. A. Protein stability curves. *Biopolymers* **26**, 1859–1877 (1987).

99. Goehlert, V. a, Krupinska, E., Regan, L. & Stone, M. J. Analysis of side chain mobility among protein G B1 domain mutants with widely varying stabilities. *Protein Sci.* **13**, 3322–3330 (2004).

100. Ikeya, T. *et al.* Improved in-cell structure determination of proteins at near-physiological concentration. *Sci. Rep.* **6**, 38312 (2016).

101. Hwang, T. L., Van Zijl, P. C. M. & Mori, S. Accurate quantitation of water-amide proton exchange rates using the Phase-Modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR* **11**, 221–226 (1998).

102. Skinner, J. J., Lim, W. K., Bédard, S., Black, B. E. & Englander, S. W. Protein hydrogen exchange: Testing current models. *Protein Sci.* **21**, 987–995 (2012).

103. Zhang, Y. & Cremer, P. S. Chemistry of Hofmeister Anions and Osmolytes. *Annu. Rev. Phys. Chem.* **61**, 63–83 (2010).

104. Lo Nostro, P. & Ninham, B. W. Hofmeister Phenomena: An Update on Ion Specificity in Biology. *Chem. Rev.* **112**, 2286–2322 (2012).

105. Rembert, K. B. *et al.* Molecular Mechanisms of Ion-Specific Effects on Proteins. *J. Am. Chem. Soc.* **134**, 10039–10046 (2012).

106. Otting, G. NMR studies of water bound to biological molecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **31**, 259–285 (1997).

107. Halle, B. Cross-relaxation between macromolecular and solvent spins: The role of long-range dipole couplings. *J. Chem. Phys.* **119**, 12372–12385 (2003).

108. Nucci, N. V., Pometun, M. S. & Wand, A. J. Mapping the Hydration Dynamics of Ubiquitin. *J. Am. Chem. Soc.* **133**, 12326–12329 (2011).

109. Nucci, N. V., Pometun, M. S. & Wand, A. J. Site-resolved measurement of water-protein interactions by solution NMR. *Nat. Struct. Mol. Biol.* **18**, 245–250 (2011).

110. Clore, G. M. & Gronenborn, A. M. Localization of bound water in the solution structure of the immunoglobulin binding domain of streptococcal protein G. *J. Mol. Biol.* **223**, 853–856 (1992).

111. Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).

112. Seewald, M. J. *et al.* The role of backbone conformational heat capacity in protein stability: Temperature dependent dynamics of the B1 domain of Streptococcal protein G. *Protein Sci.* **9**, 1177–1193 (2000).

113. Tillett, M. L., Blackledge, M. J., Derrick, J. P., Lian, L.-Y. & Norwood, T. J. Overall rotational diffusion and internal mobility in domain II of protein G from Streptococcus determined from 15 N relaxation data. *Protein Sci.* **9**, 1210–1216 (2000).

114. Idiyatullin, D., Nesmelova, I., Daragan, V. A. & Mayo, K. H. Heat Capacities and a Snapshot of the Energy Landscape in Protein GB1 from the Pre-denaturation Temperature Dependence of Backbone NH Nanosecond Fluctuations. *J. Mol. Biol.* **325**, 149–162 (2003).

115. Idiyatullin, D., Daragan, V. a & Mayo, K. H. 15 NH Backbone Dynamics of Protein GB1: Comparison of Order Parameters and Correlation Times Derived Using Various "Model-Free" Approaches. *J. Phys. Chem. B* **107**, 2602–2609 (2003).

116. Tjandra, N., Feller, S. E., Pastor, R. W. & Bax, A. Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation. *J. Am. Chem. Soc.* **117**, 12562–12566 (1995).

117. Anglister, J., Grzesiek, S., Ren, H., Klee, C. & Bax, A. Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J. Biomol. NMR* **3**, 121–126 (1993).

118. Kestin, J., Sokolov, M. & Wakeham, W. A. Viscosity of liquid water in the range −8 °C to 150 °C. *J. Phys. Chem. Ref. Data* **7**, 941–948 (1978).

119. Ferrage, F., Zoonens, M., Warschawski, D. E., Popot, J.-L. L. & Bodenhausen, G. Slow diffusion of macromolecular

assemblies by a new pulsed field gradient NMR method. *J. Am. Chem. Soc.* **125**, 2541–2545 (2003).

120. Wilkins, D. K. *et al.* Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques †. *Biochemistry* **38**, 16424–16431 (1999).

121. Jones, J. A., Wilkins, D. K., Smith, L. J. & Dobson, C. M. Characterisation of protein unfolding by NMR diffusion measurements. *J. Biomol. NMR* **10**, 199–203 (1997).

122. Cabrita, E. J. & Berger, S. DOSY studies of hydrogen bond association: tetramethylsilane as a reference compound for diffusion studies. *Magn. Reson. Chem.* **39**, S142–S148 (2001).

123. Schultz, S. G. & Solomon, A. K. Determination of the Effective Hydrodynamic Radii of Small Molecules by Viscometry. *J. Gen. Physiol.* **44**, 1189–1199 (1961).

124. Sekhar, A., Vallurupalli, P. & Kaya, L. E. Defining a length scale for millisecond-timescale protein conformational exchange. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11391–11396 (2013).

125. Hayes, R., Warr, G. G. & Atkin, R. Structure and Nanostructure in Ionic Liquids. *Chem. Rev.* **115**, 6357–6426 (2015).

126. Holyst, R. *et al.* Scaling form of viscosity at all length-scales in poly(ethylene glycol) solutions studied by fluorescence correlation spectroscopy and capillary electrophoresis. *Phys. Chem. Chem. Phys.* **11**, 9025–9032 (2009).

127. Kalwarczyk, T. *et al.* Comparative analysis of viscosity of complex liquids and cytoplasm of mammalian cells at the nanoscale. *Nano Lett.* **11**, 2157–2163 (2011).

128. Li, C., Wang, Y. & Pielak, G. J. Translational and Rotational Diffusion of a Small Globular Protein under Crowded Conditions. *J. Phys. Chem. B* **113**, 13390–13392 (2009).

129. Wang, Y., Li, C. & Pielak, G. J. Effects of Proteins on Protein Diffusion. *J. Am. Chem. Soc.* **132**, 9392–9397 (2010).

130. Adamski, W. *et al.* A Unified Description of Intrinsically Disordered Protein Dynamics under Physiological Conditions Using NMR Spectroscopy. *J. Am. Chem. Soc.* **141**, 17817–17829 (2019).

131. Persson, E. & Halle, B. Cell water dynamics on multiple time scales. *Proc. Natl. Acad. Sci.* **105**, 6266–6271 (2008).

132. Kimmich, R. & Fatkullin, N. Self-diffusion studies by intra- and inter-molecular spin-lattice relaxometry using field-cycling: Liquids, plastic crystals, porous media, and polymer segments. *Prog. Nucl. Magn. Reson. Spectrosc.* **101**, 18–50 (2017).

133. Wang, H., Zhao, M., Ackerman, J. L. & Song, Y. Saturation-inversion-recovery: A method for T1 measurement. *J. Magn. Reson.* **274**, 137–143 (2017).

134. d'Auvergne, E. J. & Gooley, P. R. Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J. Biomol. NMR* **40**, 107–119 (2008).

135. d'Auvergne, E. J. & Gooley, P. R. Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. *J. Biomol. NMR* **40**, 121–133 (2008).

136. Bieri, M., D'Auvergne, E. J. & Gooley, P. R. relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and μs motion of proteins. *J. Biomol. NMR* **50**, 147–155 (2011).

137. Clore, G. M. *et al.* Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.* **112**, 4989–4991 (1990).

138. Mandel, A. M., Akke, M. & Palmer, III, A. G. Backbone Dynamics ofEscherichia coliRibonuclease HI: Correlations with Structure and Function in an Active Enzyme. *J. Mol. Biol.* **246**, 144–163 (1995).

139. Zeeb, M., Jacob, M. H., Schindler, T. & Balbach, J. 15N relaxation study of the cold shock protein CspB at various solvent viscosities. *J. Biomol. NMR* **27**, 221–234 (2003).

140. Akke, M., Brueschweiler, R. & Palmer, A. G. NMR order parameters and free energy: an analytical approach and its application to cooperative calcium(2+) binding by calbindin D9k. *J. Am. Chem. Soc.* **115**, 9832–9833 (1993).

141. Yang, D. & Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: Application to protein folding. *J. Mol. Biol.* **263**, 369–382 (1996).

142. Yang, D., Mok, Y.-K., Forman-Kay, J. D., Farrow, N. A. & Kay, L. E. Contributions to protein entropy and heat capacity from bond vector motions measured by NMR spin relaxation. *J. Mol. Biol.* **272**, 790–804 (1997).

143. Jarymowycz, V. A. & Stone, M. J. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. *Chem. Rev.* **106**, 1624–1671 (2006).

144. Stone, M. J., Gupta, S., Snyder, N. & Regan, L. Comparison of Protein Backbone Entropy and β-Sheet Stability: NMR-Derived Dynamics of Protein G B1 Domain Mutants. *J. Am. Chem. Soc.* **123**, 185–186 (2001).

145. Mayer, K. L. *et al.* Covariation of backbone motion throughout a small protein domain. *Nat. Struct. Mol. Biol.* **10**, 962–965 (2003).

146. Smith, C. K., Withka, J. M. & Regan, L. A Thermodynamic Scale for the beta-Sheet Forming Tendencies of the Amino Acids. *Biochemistry* **33**, 5510–5517 (1994).

147. De Santis, S. *et al.* Cholinium-amino acid based ionic liquids: a new method of synthesis and physico-chemical characterization. *Phys. Chem. Chem. Phys.* **17**, 20687–20698 (2015).

148. Privalov, P. L. & Makhatadze, G. I. Heat capacity of proteins. *J. Mol. Biol.* **213**, 385–391 (1990).

149. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).

150. Findeisen, M., Brand, T. & Berger, S. A 1H-NMR thermometer suitable for cryoprobes. *Magn. Reson. Chem.* **45**, 175–178 (2007).

151. Wishart, D. S. *et al.* 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135–140 (1995).

152. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 687–692 (1995).

153. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Bioinforma.* **59**, 687–696 (2005).

154. Gronenborn, A. M. & Clore, G. M. Rapid screening for structural integrity of expressed proteins by heteronuclear NMR spectroscopy. *Protein Sci.* **5**, 174–177 (1996).

155. Kay, L. E., Keifer, P. & Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665 (1992).

156. Takayama, Y., Castañeda, C. A., Chimenti, M., Garcĺa-Moreno, B. & Iwahara, J. Direct Evidence for Deprotonation of a Lysine Side Chain Buried in the Hydrophobic Core of a Protein. *J. Am. Chem. Soc.* **130**, 6714–6715 (2008).

157. Zandarashvili, L., Esadze, A. & Iwahara, J. NMR Studies on the Dynamics of Hydrogen Bonds and Ion Pairs Involving Lysine Side Chains of Proteins. in *Biomolecular Spectroscopy: Advances from Integrating Experiments and Theory* (ed. Intergovernmental Panel on Climate Change) 37–80 (Cambridge University Press, 2013).

158. Tollinger, M., Forman-Kay, J. D. & Kay, L. E. Measurement of Side-Chain Carboxyl p K a Values of Glutamate and Aspartate Residues in an Unfolded Protein by Multinuclear NMR Spectroscopy. *J. Am. Chem. Soc.* **124**, 5714–5717 (2002).

159. Hansen, A. L. & Kay, L. E. Quantifying millisecond time-scale exchange in proteins by CPMG relaxation dispersion NMR spectroscopy of side-chain carbonyl groups. *J. Biomol. NMR* **50**, 347–355 (2011).

160. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89**, 496–514 (1990).

161. Yamazaki, T., Yoshida, M. & Nagayama, K. Complete assignments of magnetic resonances of ribonuclease H from Escherichia coli by double- and triple-resonance 2D and 3D NMR spectroscopies. *Biochemistry* **32**, 5656–5669 (1993).

162. Wallerstein, J., Weininger, U., Khan, M. A. I., Linse, S. & Akke, M. Site-specific protonation kinetics of acidic side chains in proteins determined by pH-dependent carboxyl 13C NMR relaxation. *J. Am. Chem. Soc.* **137**, 3093–3101 (2015).

163. Ying, J., Delaglio, F., Torchia, D. A. & Bax, A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **68**, 101–118 (2017).

164. Skinner, S. P. *et al.* CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* **66**, 111–124 (2016).

165. Zhang, H., Neal, S. & Wishart, D. S. RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR* **25**, 173–195 (2003).

166. Wu, D. H., Chen, A. D. & Johnson, C. S. An Improved Diffusion-Ordered Spectroscopy Experiment Incorporating Bipolar-Gradient Pulses. *J. Magn. Reson. Ser. A* **115**, 260–264 (1995).

167. Stejskal, E. O. & Tanner, J. E. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J. Chem. Phys.* **42**, 288–292 (1965).

168. Ferrage, F., Zoonens, M., Warschawski, D. E., Popot, J. L. & Bodenhausen, G. Erratum: Slow Diffusion of Macromolecular Assemblies Measured by a New Pulsed Field Gradient NMR Method (Journal of the American Chemical Society (2003) 125 (2541-2545)). *J. Am. Chem. Soc.* **126**, 5654 (2004).

169. Mao, X., Guo, J. & Ye, C. Radiation damping effects of transverse relaxation time measurements. *Chem. Phys. Lett.* **227**, 65–68 (1994).

170. Mori, S., Abeygunawardana, C., Johnson, M. O. & Vanzijl, P. C. M. Improved Sensitivity of HSQC Spectra of Exchanging Protons at Short Interscan Delays Using a New Fast HSQC (FHSQC) Detection Scheme That Avoids Water Saturation. *J. Magn. Reson. Ser. B* **108**, 94–98 (1995).

171. Farrow, N. A. *et al.* Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15 N NMR Relaxation. *Biochemistry* **33**, 5984–6003 (1994).

172. Connelly, G. P., Bai, Y., Jeng, M.-F. & Englander, S. W. Isotope effects in peptide group hydrogen exchange. *Proteins Struct. Funct. Genet.* **17**, 87–92 (1993).

173. Orban, J., Alexander, P. & Bryan, P. Hydrogen-Deuterium Exchange in the Free and Immunoglobulin G-Bound Protein G B-Domain. *Biochemistry* **33**, 5702–5710 (1994).

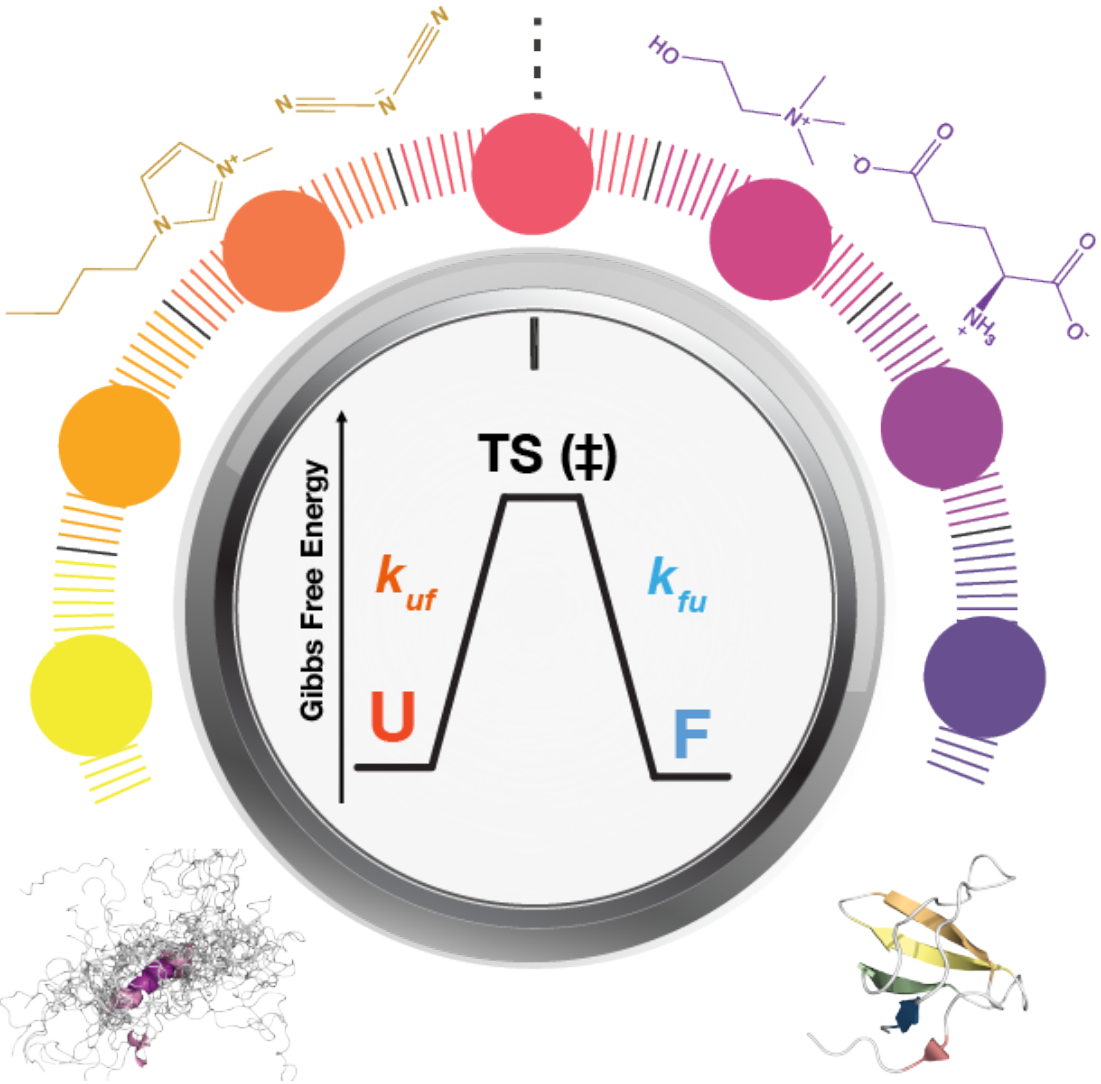
174. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. & Bax, A. Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J. Am. Chem. Soc.* **111**, 1515–1517 (1989).

175. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845–858 (1994).

176. Clore, G. M., Bax, A., Wingfield, P. T. & Gronenborn, A. M. Identification and localization of bound internal water in the solution structure of interleukin 1.beta. by heteronuclear three-dimensional proton rotating-fram Overhauser nitrogen-15-proton multiple quantum coherence NMR spectroscopy. *Biochemistry* **29**, 5671–5676 (1990).

3

# Chapter 3: Understanding the effects of ILs on drkN SH3 stability



## Abstract

Despite several years of intensive research, the topic of protein folding still moves a lot of interest, particularly the understanding of how salt ions and other cosolutes can modulate the folding landscape. To accomplish this, and to understand how the folding and unfolding paths can be tuned, a complete elucidation of the different accessible states that a protein can explore during this process is crucial. However, typical equilibrium experiments only provide information concerning the folded structure or about the transition between folded and unfolded states, and the effects involving the unfolded ensemble as well as the transition state remain ill-defined. Following chapter 2, which results were interpreted assuming only the GB1's native state, in this chapter, using the previous tailored ionic liquids (ILs) as cosolutes ([Ch][Glu] and [Bmim][dca] ILs), I was able to modulate the folding pathway of the N-terminal SH3 domain of the *Drosophila* Drk protein (drkN SH3). Using a combination of NMR techniques, I have studied the structure and conformational exchange of the drkN SH3 domain on each cosolute and provide a comprehensive understanding of the stabilising and destabilising effects, not only in the folded state but also in the unfolded ensemble. The data reveals that (de)stabilisation of SH3 occurs via electrostatic and hydrophobic IL-protein interactions, including a concomitant modification of the water structure around the protein. While [Ch][Glu] seems to interact preferentially with the folded state of SH3, [Bmim][dca] interacts preferentially with the unfolded state, introducing a non-native helical propensity in a hydrophobic patch. These interactions lead to different enthalpic and entropic changes that affect the equilibrium thermodynamics of SH3. Quantification of these changes shows that in both ILs the (de)stabilisation is entropically-driven and it is counteracted by an enthalpic (de)stabilisation. Characterization of the folding/unfolding interconversion rates by ZZ-exchange spectroscopy shows that [Bmim][dca] leads to SH3 destabilisation by significantly slowing the folding rate due to the increase of helical propensity in the ensemble, while [Ch][Glu] stabilises the folded state by slowing the unfolding rate due to ions preferential-folded state accumulation. Our data provides a thorough understanding of the IL-protein interactions and their effect on the equilibrium thermodynamics and, importantly, the NMR methodologies used represent a robust and effective way to study the elusive unfolded states of proteins.



Contributed to the publication:

**Shifting the equilibrium – the role of the unfolded state for protein stability in ionic liquids**

Micael S. Silva, Aldino Viegas, Philip O'Toole, Sara S. Felix, Angelo M. Figueiredo, Eurico J. Cabrita

*In preparation.*

## Introduction

A plethora of NMR-based experiments were used to probe GB1 stability in the presence of two ionic liquids with opposite effects. Because the protein exhibits high thermal and chemical stability, induced changes by the IL in its stability, structure, and dynamics were revealed in a small magnitude and are very difficult to interpret. While there is a strong justification for the denaturation of GB1 by direct and hydrophobic interactions with [Bmim][dca], the stabilising [Ch][Glu] effects are a balance between preferential hydration, preferential accumulation, and electrostatic interactions with specific sensible hotspots of the protein. These effects are still misunderstood. One of the main limitations of chapter 2 is that the results were interpreted assuming that ionic liquids (and salts) affect only the folded state. This may have provided only partial insights because the stability of a protein depends on the relationship between the Gibbs free energies of the folded and the unfolded states.

As observed in the introductory chapter of protein stability, unfolded states (including intrinsically disordered regions, IDRs, or entire intrinsically disordered proteins, IDPs) play vital roles in numerous diseases (e.g., neurodegenerative diseases) and cellular processes (e.g., regulation of transcription and translation, signal transduction, phosphorylation or storage of small molecules) and should not be overlooked[1,2]. Despite their apparent simplicity (in the sense they lack the organization and complexity of ordered proteins), due to their dynamic ensemble nature and a scarcity of experimental restraints as well as reduced informational content of their amino acid sequences, disordered states of proteins are difficult to characterize structurally[3,4]. Yet, despite these caveats, unfolded ensembles may not be completely random entities, since they evolved to have some adjustable, controllable, regulable, tuneable, and, oftentimes, very specific properties required for their biological functions[4]. Subtle changes in the environment or the protein sequence can easily modulate the flux between the pathways, potentially leading to a pathogenic outcome.

Here, I explore not only the folded state of a protein but also its unfolded state and the mechanism by which they interchange. In particular, the study of the well characterized and metastable drkN SH3 domain in the presence of aqueous - ionic liquids, namely, the stabilising choline glutamate ([Ch][Glu]) and the denaturing 1-buty-3-methyl-imidazolium dicyanamide ([Bmim][dca]) (Fig. 3.1), in terms of stabilisation, interaction, structural, kinetic, and thermodynamic changes.

The drkN SH3 is the isolated 59-residue (6.2 kDa) N-terminal Src homology 3 (SH3) domain of the *Drosophila* adapter protein Drk (Downstream of receptor kinase), involved in cellular signalling pathways, and it provides an excellent model system to study the unfolded state[3,5,6]. While this domain contains all residues necessary for its fold, it is relatively unstable. Folded ($F_{exch}$) and unfolded ($U_{exch}$) states exist in equilibrium, almost equally populated, and in slow exchange on NMR chemical shift timescale (Fig. 3.1), under aqueous buffer near-neutral pH[7,8]. The interconversion of the two states was shown to be a "two-state" folding mechanism with an average slow exchange, $k_{ex}$, of ~2 $s^{-1}$ (pH 6.0, 293 K)[5], where both states can be observed simultaneously using different NMR experiments (Fig. 3.1a). In particular, the indole NH sidechain of the single W36 residue is resolved in both folded and unfolded forms, and the area under each resonance is proportional to its population and therefore used to quantify protein stability[8]. The use of SH3 as a model system allows not

only the direct comparison between $F_{exch}$ and $U_{exch}$ (in equilibrium) under the same conditions, but also the comparison of $U_{exch}$ with denatured states that exist under conditions of chemical (e.g., denaturant agents) or thermal denaturation[9,10].

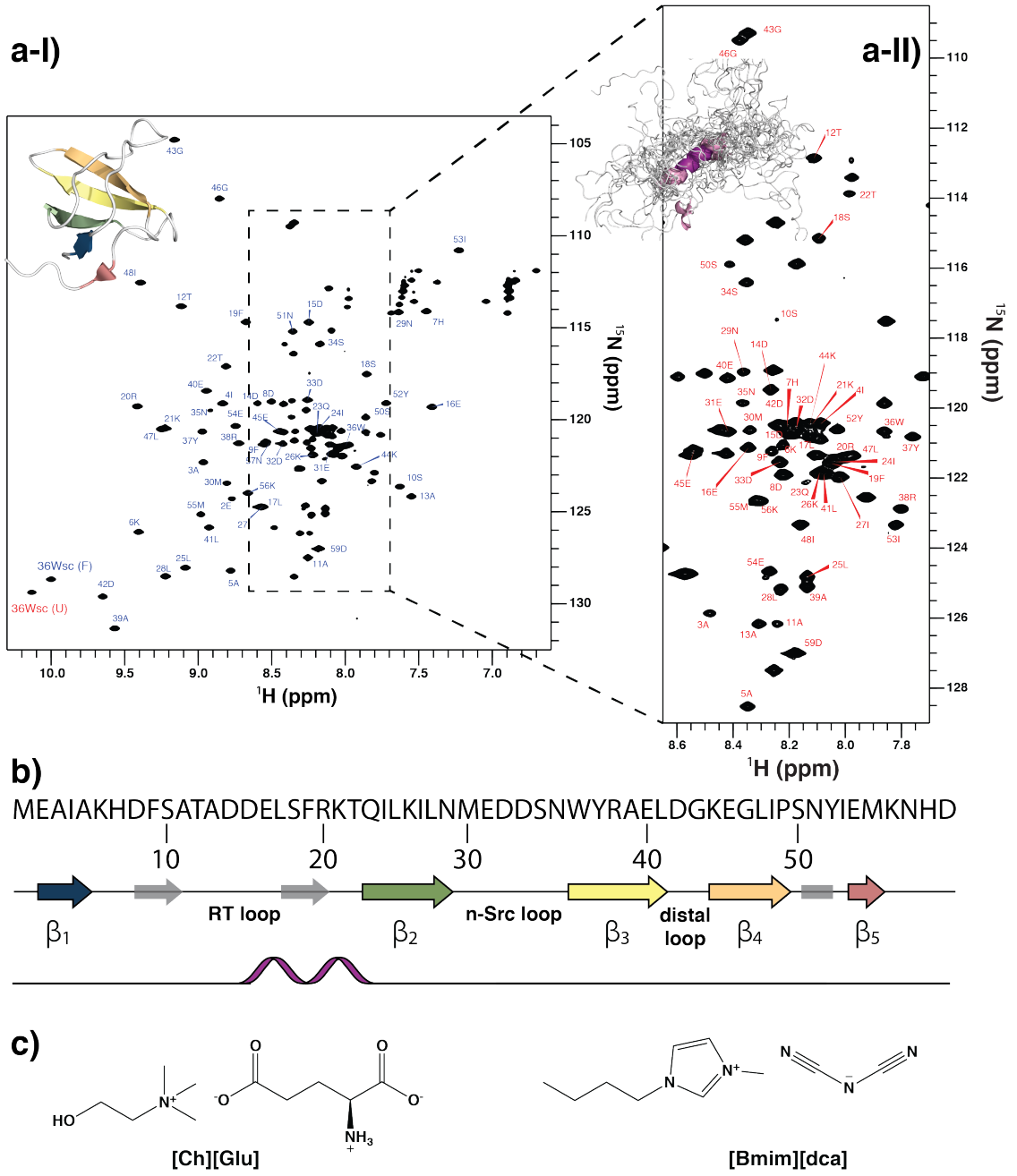



**Figure 3.1. The folded and unfolded ensembles of the drkN SH3 domain.**

**a**) [$^{1}$H,$^{15}$N]-HSQC of drkN SH3 showing the assignment of the folded (**I**) and unfolded (**II**) amide NH peaks in water, 298 K, pH=7.1. A representative of the 3D ensemble for the folded (PDB: 2A36[6]) and 30 representatives of the unfolded (PED-8AAC, kindly provided by Dr. Forman-Kay[3]) protein are depicted inside each corresponding spectrum. **b**) drkN SH3 primary sequence (top) and secondary structure diagrams of the folded (middle) and unfolded (bottom) states. β-sheets are represented as arrows and coloured according to the cartoon representation; the loops (RT, reverse turn; n-Src and distal loop) are represented as a line and the α-helix (bottom) is coloured purple with its variation in length is represented by a lighter tone and

dashed lines. Note that the RT loop is also constituted by two irregular β-hairpins (gray arrows) and strands β4 and β5 are connected by a short three residue $3_{10}$ helix (gray rectangle) **c**) Used ionic liquids: [Ch][Glu], on the left, and [Bmim][dca] on the right.

In the folded state ($F_{exch}$), the drkN SH3 domain folds as a β-sandwich (Fig. 3.1a-I, 1b, PDB:2A36[6]). The β-strands are arranged in two sheets and an irregular β-hairpin, with strands β1, β5 and part of β2 forming the first sheet and strands β2, β3, and β4 forming the second sheet. The irregular twisted β-hairpin is referred to as the RT loop. Strands β2 and β3, and β3 and β4 are connected by the N-Src loop, and the distal loop, respectively. Strands β4 and β5 are connected by a short three-residue $3_{10}$ helix typically seen in other SH3 domains. In the unfolded ensemble ($U_{exch}$), where structures are interconverting rapidly, a large pool of data and experimental restraints (e.g., chemical shifts, J-couplings, residual dipolar couplings, $^{15}N$ relaxation, nuclear Overhauser effect (NOE), paramagnetic relaxation enhancement (PRE), hydrodynamic radius measurements, small angle X-ray scattering, computational studies and singe molecule Förster resonance energy transfer data (FRET)), essentially made available by Forman-Kay group[11–14], provide evidence for fluctuating structure, including a non-native $\alpha$-helical secondary structure roughly between residues D15 and T22 (Fig. 3.1a-II, 1b, PED8-AAC[3] which has been updated to PED00022). The following residues from Q23 to L28 experience significant $^{15}N$ linewidth-broadening in the $U_{exch}$ ensemble due to conformational exchange on intermediate ($\mu s/ms$) timescale, indicating a local and non-native hydrophobic collapse[5,9,15]. Also, the W36 indole appears to be more buried in the $U_{exch}$ state the $F_{exch}$ state (it is located in the surface at the C-terminal of n-Src loop), which may be explained by the involvement in an non-native hydrophobic cluster[10,16,17]. Importantly, the observation that the unfolded ensemble of drkN SH3 undergoes conformational sampling while chaperone-bound[18,19] supports that such ensemble is representative of a physiologically relevant disordered state.

Since SH3 (pI ≈ 4.6) is fully stabilised upon binding to its positively charged proline-rich target peptide from the guanidine nucleotide exchange factor Sos[9], it seems logical that neutralization of its negative charge should increase protein stability. In light of this, NMR studies have demonstrated that adding inorganics salts, such as 0.4 M $Na_2SO_4$, leads to stabilisation of the protein[8]. However, studies in the presence of high concentrations of NaCl as well as the results of substitutions of negatively charged residues to neutral amino acids[15] demonstrated that decrease in electrostatic repulsion does not lead to complete stabilisation of the drkN SH3, suggesting that electrostatic interactions are not the major factor contributing to protein instability. This is supported by the fact that a single substitution of T22 in the diverging turn provides dramatic stabilisation in the protein, not only by the formation of a stable turn in the folded state but also by the formation of a non-native structure in the unfolded state[15]. The general stabilisation of the domain by $Na_2SO_4$ is thus explained by the effect of the sulphate anion, which is near the extreme end of the Hofmeister series of stabilising anions (as discussed in the introductory chapter: beyond the Hofmeister series: Ion-specific effects).

Recently, using $^{19}F$ labelling of the drkN SH3 sole tryptophan, Pielak's group has been explored its metastability with small and larger cosolutes[20–23]. Because SH3 is destabilised in

cells in response to osmotic stress, and this destabilisation is alleviated by adding an osmolyte such as glycine betaine[21], the authors suggest that crowding-induced charge-charge interactions slow folding due to preferential interactions with the unfolded ensemble, and reducing these interactions increases protein stability. Using ionic liquids and the same methodology of $^{19}F$ NMR, Li and coworkers[24] demonstrated that drkN SH3 is destabilised in the presence of [Bmim][Br] IL, suggesting a mechanism of entropic destabilisation offset by enthalpic stabilisation. While the method used is reliable to acquire thermodynamic data, there is no distinction between the IL effects on the folded and unfolded states neither in the nature of the IL-protein interactions.

In this chapter, NMR was used to quantify the changes in structural and folding-unfolding kinetics of drkN SH3 at a residue-level under stabiliser or destabilisers aqueous-IL conditions (i.e., the presence of [Ch][Glu] or [Bmim][dca], respectively). The protein folding landscape was also characterized with changes in the free-energy barriers, unfolding ($F \rightarrow TS^{\ddagger}$) and folding ($U \rightarrow TS^{\ddagger}$), discerning simultaneously the (de)stabilisation degree of folded and unfolded state by ILs.

## Results and Discussion

**Protein stability**

As I have seen before in the introductory chapter and discussed in the chapter 2, the effects of ILs on protein stability depend on different parameters, such as, IL composition, ion-pair strength, concentrations, protein surface charge and composition and specific ion-protein (electrostatics and hydrophobics) interactions. In general, neutral-to-stabiliser effects of [Ch][Glu] and denaturant effects of [Bmim][dca] ILs on the stability of the negatively charged GB1 are well correlated with its expected ion position in the Hofmeister series (chapter 2 – ILs and salt effects on protein stability). While the effect of the $[Ch]^+$ cation is overcompensated by the effect of a strongly stabilising oxyanion ($[Glu]^-$), the combination of both destabilisers $[Bmim]^+$ cation and $[dca]^-$ anion results in a strong denaturant.

To investigate the effect of [Ch][Glu] and [Bmim][dca] ILs on the equilibrium thermodynamics of drkN SH3, the protein was titrated with increasing amounts of these ILs and the intensities of backbone amides for both the folded and unfolded states were followed by acquiring an $[^1H,^{15}N]$-HSQC spectrum at each titration point (see Fig. 3.2). Addition of [Ch][Glu] caused a small decrease in intensity for the resonances of the folded state (due to viscosity effects that lead to slow tumbling) but also a progressive disappearance of the resonances of the unfolded state, suggesting that the equilibrium between $F_{exch}$ and $U_{exch}$ is shifted toward the folded state upon IL addition (Fig. 3.2a). On the other hand, addition of [Bmim][dca] had the opposite effect (Fig. 3.2b).

Since the signals of the indole NH side chain of W36 are well resolved in each state and it is located in an exposed and flexible region of the $F_{exch}$ state, the volume intensity of each resonance ($V_f$ and $V_u$) is proportional to its population[8], $p_f$ and $p_u$, respectively, accordingly to the relation,

$$p_{f,u} = \frac{V_{f,u}}{V_f + V_u}. \qquad [3.1]$$

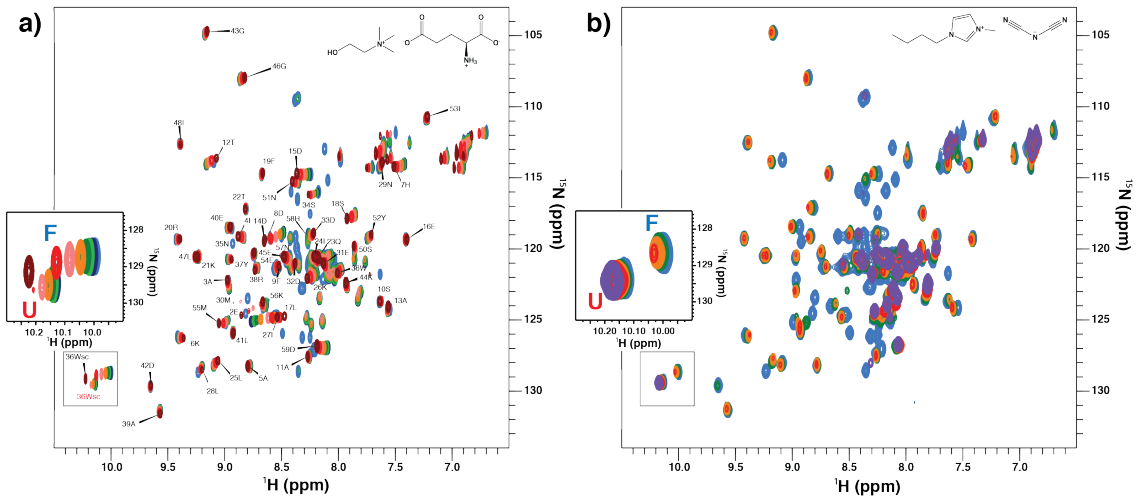



**Figure 3.2. [$^1$H,$^{15}$N]-HSQC titration with [Ch][Glu] and [Bmim][dca.**

**a)** Overlay of 2D $^1$H-$^{15}$N HSQC spectra of drkN SH3 acquired in [Ch][Glu] titration (0, 0.01, 0.025, 0.05, 0.1, 0.25, 0.5 and 1.0 M). The label shows the assignment for the folded state. **b)** Overlay of 2D $^1$H-$^{15}$N HSQC spectra of drkN SH3 acquired in [Bmim][dca] titration (0, 0.05, 0.1, 0.2 and 0.3 M). The inlay shows the indole W36 side chain peaks of the folded (F) and unfolded (U) states. All the spectra were acquired in water, to avoid interference from the buffer, solution pH ≈ 7.1, at 298.2 K and 600.13 MHz.

Assuming a two-state unfolding model and an equilibrium between a folded (F) and an unfolded state (U),

$$F \rightleftarrows U, \qquad [3.2]$$

The F/U populations were used to quantify protein stability via the modified standard Gibbs free energy of unfolding (F → U, $\Delta G_u^{0'}$)[25],

$$\Delta G_u^{0'} = -RTln\left(\frac{p_u}{p_f}\right), \qquad [3.3]$$

where $R$ is the gas constant and $T$ is the absolute temperature.

The obtained $\Delta G_u^{0'}$ values, at each titration point, were then fitted by a least square analysis using the linear extrapolation method[26] according to the Eq. 3.4 which differs in sign from that traditionally used but express whether the U ⇄ F transition is favourable or unfavorable[27].

$$\Delta G_u^{0'} = \Delta G_u^{H_2O} + m\,[cosolute], \qquad [3.4]$$

where $\Delta G_u^{H2O}$ is an estimate of the free energy of unfolding in water and the $m$-value (slope) is a measure of the efficacy of a given cosolute to either force proteins to fold or to unfold. The $m$-value is equal to the free energy difference at 1 M of cosolute, $\Delta\Delta G_u^{1M} = \Delta G_{u,cosolute}^{1M} -$

$\Delta G_u^{H2O}$. In other words, higher the *m*-value of a given cosolute, higher the stabilisation of the protein.

The equilibrium between $F_{exch}$ and $U_{exch}$ ($p_f \approx p_u \approx 0.5$, in water) is shifted to a fully folded or unfolded state ($p_f$ or $p_u = 1$) with addition of [Ch][Glu] or [Bmim][dca], respectively (Fig. 3.3). Obtained $\Delta G_u^{0'}$ values are linearly-dependent as a function of IL concentration and clearly indicate that [Ch][Glu] stabilises the protein with an *m*-value of 1.9 ± 0.1 kcal/mol/M, while [Bmim][dca] strongly destabilises the protein with an *m*-value of -7.7 ± 0.3 kcal/mol/M.

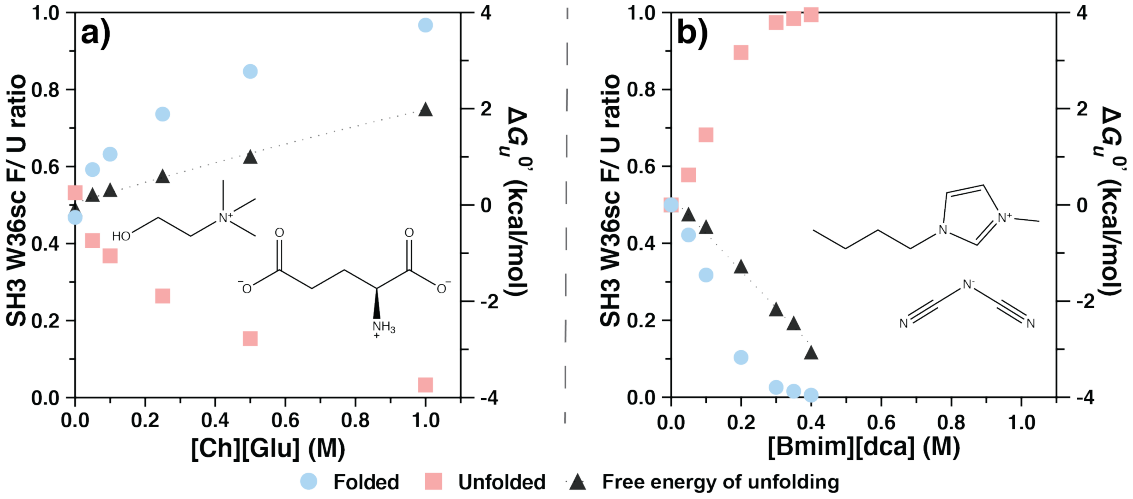



**Figure 3.3. Populations of folded and unfolded ensemble along IL titrations.**

Populations and Gibbs free energy of unfolding ($\Delta G_u^0$) were calculated as function of **a)** [Ch][Glu] and **b)** [Bmim][dca] concentration. Based on each $^1H$-$^{15}N$ HSQC spectrum, peak volume intensities of the exposed indole side chain of W36 were used to calculate the populations of folded state ($p_f$, blue) and unfolded ensemble ($p_u$, red). For each titration point, the Gibbs free energy of unfolding ($\Delta G_u^{0'}$, black) was calculated assuming a two-state unfolding model.

Having observed that [Ch][Glu] and [Bmim][dca] are able to tune the protein's folding landscape and given the significant electrostatic surface potential of the SH3 domain (pI ≈ 4.6), I needed to understand: (1) if the observed stability was caused by a simple ionic strength effect (via electrostatic shielding); (2) if some ion specific interactions should be considered; (3) and the existence of a specific ionic liquid effect (i.e., if the combination of the ions in the ILs resulted in an effect different from the sum of the individual ion contributions). To differentiate between a general ionic strength effect and a specific ionic liquid effect, I repeated the above experiments in the presence of cholinium ($[Ch]^+$), glutamate ($[Glu]^-$), 1-Butyl-3-methylimidazolium ($[Bmim]^+$) and dicyanamide ($[dca]^-$) salts, namely [Ch][Cl], Na[Glu], [Bmim]Cl, Na[dca] and, as a reference NaCl (Fig. 3.4).

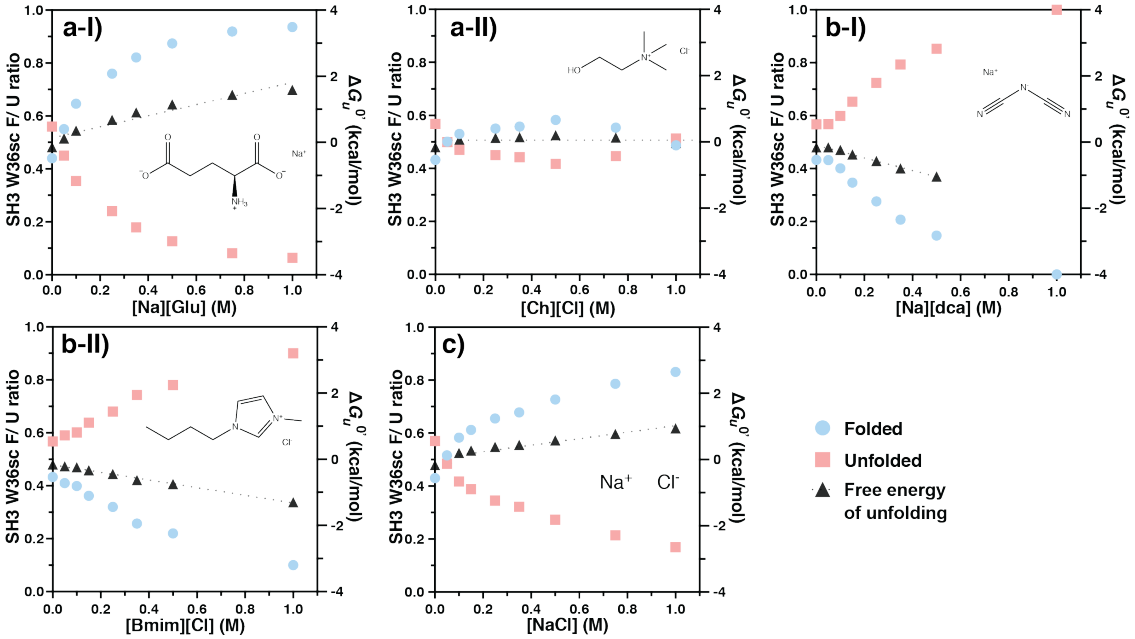



**Figure 3.4. Populations of folded and unfolded ensemble along salt titrations.**

Populations and Gibbs free energy of unfolding ($\Delta G_u^0$) were calculated as function of **a)** salts related with [Ch][Glu], **I**) Na[Glu] and **II**) [Ch]Cl; or salts related with [Bmim][dca], I) Na[dca] and II**)** [Bmim]Cl; and **c)** NaCl concentration as a control. Based on each $^1$H-$^{15}$N HSQC spectrum, peak volume intensities of the exposed indole side chain of W36 were used to calculate the populations of folded state ($p_f$, blue) and unfolded ensemble ($p_u$, red). For each titration point, the Gibbs free energy of unfolding ($\Delta G_u^{0'}$, black) was calculated assuming a two-state unfolding model.

In general, based on the measured $\Delta G_u^{0'}$ values and respective *m*-values (Fig. 3.5 and Table 3.1), I observe that the effects of the ILs cannot be accounted solely by the individual contributions of the individual ions. Also, the observed effects cannot be accounted for simply by shielding of unfavourable electrostatic interactions at higher ionic strength. In fact, previous NMR studies of the drkN SH3 in the presence of high concentrations of NaCl have shown that electrostatic shielding is not sufficient to completely fold this domain[15]. The data reveals that the stabilising effect observed for [Ch][Glu] derives mainly from the anion $[Glu]^-$ (as observed for the effects of this IL and $[Glu]^-$ salt on GB1's $T_m$ values in chapter 2). Therefore, despite the almost neutral effect of the $[Ch]^+$ cation the general stabilising effect of [Ch][Glu] must be driven by an overcompensating effect of the $[Glu]^-$ anion, which is near the extreme end of the Hofmeister series of stabilising anions[28].

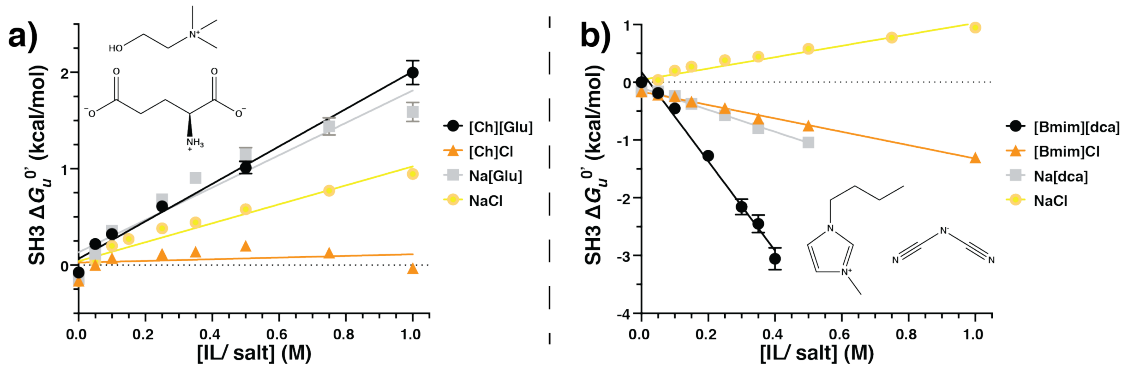



**Figure 3.5. Cosolute-induced (de)stabilisation of drkN SH3.**

Effect of **a**) [Ch][Glu] or **b**) [Bmim][dca] and related ionic salts in the $\Delta G_u$ of drkN SH3. The populations were estimated by calculating the ratio of volume intensity of indole side chain of Trp36 signals at F/U states in each HSQC spectrum. These populations in water are almost equally, $p_f \approx p_u$. $\Delta G_u$ and *m* values were determined as described in the main text, with the errors being propagated from peak volume error measurement.

**Table 3.1. *m*-Values of drkN SH3 in the presence of different cosolutes.**

| Co-solute | *m*-value (kcal/mol/M) |
|---|---|
| **[Ch][Glu]** | **1.9 ± 0.1** |
| Na[Glu] | 1.7 ± 0.2 |
| NaCl | 1.0 ± 0.1 |
| [Ch]Cl | 0.1 ± 0.1 |
| [Bmim][Cl] | -1.16 ± 0.04 |
| Na[dca] | -1.9 ± 0.1 |
| **[Bmim][dca]** | **-7.7 ± 0.3** |

*m*-Values of drkN SH3 were extracted from the slope of each titration at 298.2 K accordingly with $\Delta G_u^{0'} = \Delta G_u^{H2O} + m$ [cosolute], where $\Delta G_u^{H2O}$ is an estimate of the free energy of unfolding in water and the *m*-value is a measure of the efficacy of a given cosolute to either force proteins to fold or to unfold. The errors were derived from the fitting. Details can be found in Materials and Methods section.

On the other hand, as expected[29] and in agreement with the results of chapter 2 about GB1 destabilisation with [Bmim][dca] IL, both the cation $[Bmim]^+$ and the anion $[dca]^-$ seem to be equally destabilisers, with their combination resulting in a very strong denaturant. Overall, considering the (de)stabilisation of SH3 (Table 3.1), the tested cosolutes can be ranked as [Ch][Glu] ≈ Na[Glu] > NaCl > [Ch]Cl ‖ [Bmim]Cl > Na[dca] ≫ [Bmim][dca] (the double bar (‖) indicates the crossover from stabilising to destabilising behaviour).

**Ion specific interactions**

Having observed that different salts are able to modulate drkN SH3 stability, and since the simple shielding of unfavourable electrostatic interactions is not enough to account for the observed effects, I proceeded to investigate the mechanisms behind protein (de)stabilisation. Previous work[29,30] has shown that the interactions between proteins and ILs including electrostatic interactions and hydrophobic interactions are crucial to understanding the effects of ILs on the stability of proteins. Thus, I proceeded to identifying and mapping the IL-protein interactions. For this I used the backbone amide resonance chemical shift from [$^{1}$H,$^{15}$N]-HSQC spectra acquired in the presence of increasing amounts of [Ch][Glu] and [Bmim][dca] and followed the chemical shift perturbations (CSP) of the amide resonances along the titration. For clarity, and to be able discuss in detail the analysis of the interactions of the studied cosolutes with both the folded (F) and unfolded (U) states one condition is presented for each IL (Fig. 3.6).

For [Ch][Glu], the data at 0.35 M of concentration (Fig. 3.6a) corresponds to approximately 78% of folded SH3 ($F_{[Ch][Glu]}$) and 22% of unfolded ($U_{[Ch][Glu]}$). For [Bmim][dca], the data at 0.15 M is shown (Fig. 3.6b), which corresponds to approximately 14% of folded SH3 ($F_{[Bmim][dca]}$) and 86% of unfolded ($U_{[Bmim][dca]}$). To distinguish the affected from the non-affected residues, the CSPs were represented in terms of combined chemical shift[31] and the cut-off line was determined according to Schumann et al[32]. The CSPs were mapped onto the 3D structure of the folded protein (PDB:2A36[6]) and in one representative of the unfolded ensemble. This unfolded representative is the structure 576 from the PED8-AAC ensemble[3] (which has been updated to PED00022 and corresponds to 17 ensembles of 100 structures, where structures are number from 1 to 1700). Considering the ensemble that is more consistent with the large experimental dataset previously obtained[11–14], where the region from approximately residues 15-22 is consistently found to have a greater than 70% population in the $\alpha$-region and ~30% formation of $\alpha$ -helix[12], the structure 576 is representative of that ensemble in the sense that has significant non-native $\alpha$-helical secondary structure.

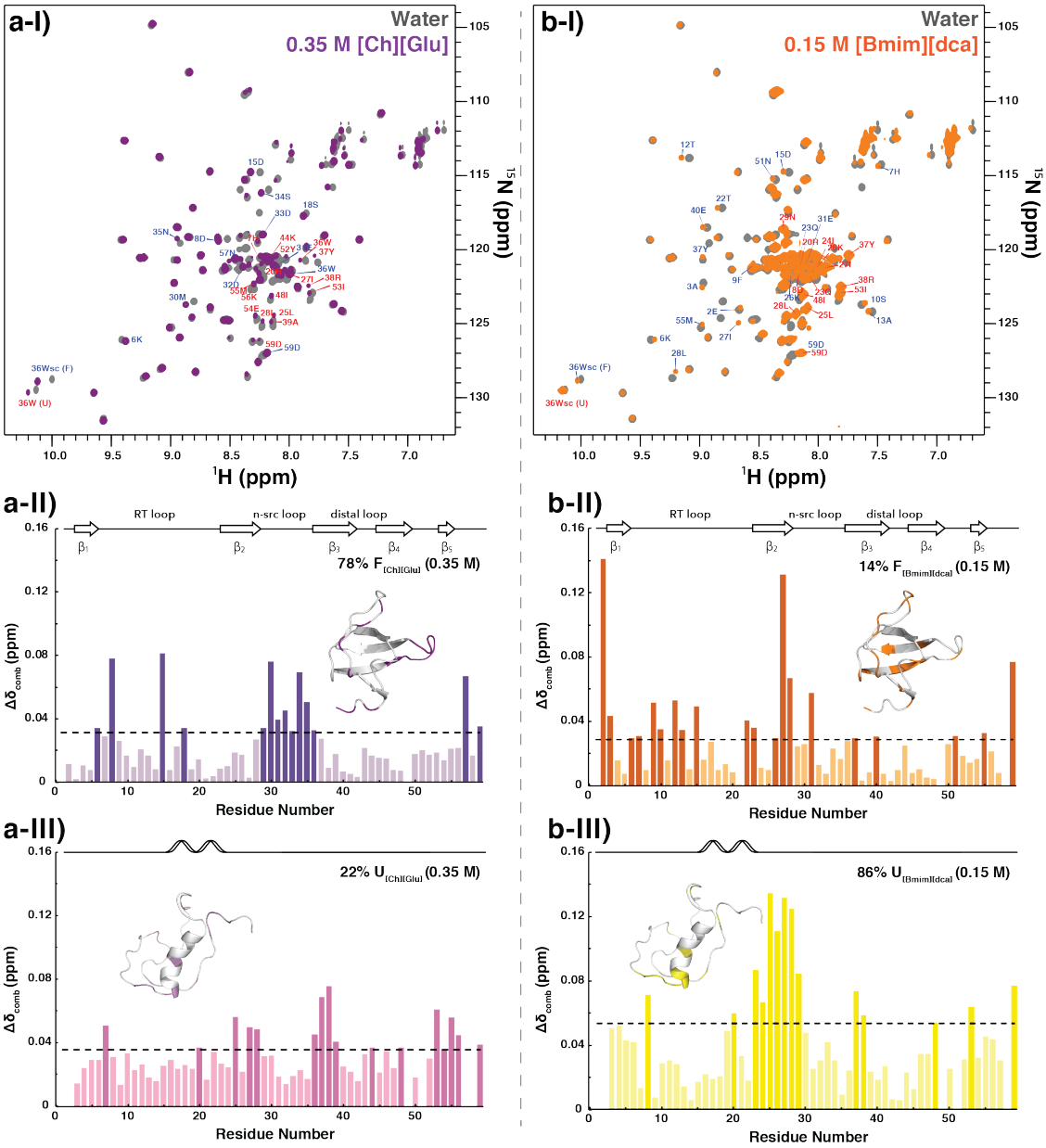



**Figure 3.6. Ionic liquid–protein interactions.**

**a**) [Ch][Glu] – SH3 interactions. a-**I**) Overlay of 2D $^{1}$H-$^{15}$N HSQC of drkN SH3 acquired in water (gray) and 0.35 M [Ch][Glu] (purple), where the labels show the assignments for the residues most affected (blue for folded and red for unfolded state); Combined chemical shift of drkN SH3 in the presence of 0.35 M of [Ch][Glu] for the a-**II**) folded (purple) and a-**III**) unfolded (pink) states, respectively. and in the presence of 0.15 M of [Bmim][dca] for the folded (orange) - and unfolded (yellow) states, respectively (the residues that show a combined chemical shift above the threshold (dashed line) are coloured in a darker tone). **b**) [Bmim][dca] – SH3 interactions. b-**I**) Overlay of 2D $^{1}$H-$^{15}$N HSQC of drkN SH3 acquired in water (gray) and 0.15 M [Bmim][dca] (orange); Combined chemical shift of drkN SH3 in the presence of 0.15 M of [Bmim][dca] for the b-**II**) folded (orange) and b-**III**) unfolded (yellow) states, respectively. Inside each plot the affected residues are mapped onto the respective 3D structure (folded: PDB: 2A36[6]); unfolded: representative of the unfolded ensemble PED-8AAC, #576[3]). Above each plot it is depicted the secondary structure of the protein according to the 3D structures. The combined chemical shifts were calculated against the folded and unfolded chemical shifts in water ($F_{water}$ and $U_{water}$, respectively).

Looking at the CSP plots for the folded and unfolded states of SH3 in [Ch][Glu] and [Bmim][dca] (Fig. 3.6), the first thing I notice is that both ILs interact with both states of the protein. However, a more careful analysis indicates that the specificity of this interaction and the affected residues are different. While [Ch][Glu] seems to interact in a non-specific manner with the folded state, (Fig. 6 a-II *vs* a-III), clearly, [Bmim][dca] seems to interact more specifically with the unfolded state (Fig. 6 b-II *vs* b-III). Interaction of [Ch][Glu] with the unfolded stated and [Bmim][dca] with the folded state seems to be more non-specific. Inspection of the location of the CSPs in the folded state in the presence of [Ch][Glu] revealed that they are predominately clustered at the n-Src loop between β2 and β3 sheets (Fig. 3.6 a-II), while in the presence of [Bmim][dca] (Fig. 3.6 b-II), with exception of E2 and I27 residues, they are spread throughout the protein sequence. For the unfolded state, in the presence of [Ch][Glu] (Fig. 3.6 a-III) the affected residues also do not seem to cluster in any specific region of the protein, in contrast to the ones affected by [Bmim][dca] (Fig. 3.6 b-III), which cluster in a region directly adjacent to the region prone to form a non-native $\alpha$-helical structure[3,8,9,14,15]. Interestingly, as previously observed for this region[33], residues I24 to L28 correspond, to a local minimum in a hydrophilicity plot.

In this section, I will go throughout each condition, first for [Ch][Glu] and latter for [Bmim][dca], as well as first for SH3's folded state and latter for the unfolded one.

[Ch][Glu]-SH3 interactions

For the $F_{[Ch][Glu]}$ state the CSPs have, in general, a rather low magnitude and, besides the n-Src loop, also some residues in the loop between β1 and β2 sheets are affected, with particular relevance for D8 and D15 residues. These locations correspond to flexible regions of the protein, and hence would be expected to be more susceptible to changes in the environment and to interact with dissolved ions. Analysis of the affected amino acid types (in terms of charge, polarity, or hydrophobicity) shows that the most affected are negatively charged (D8, D15, E31, D32, D33, D9) and polar, uncharged (S18, N29, S34, N35, N57). Since only one positively charged residue is affected (K6), the observed chemical shift perturbations are mainly consistent with the interaction of the $[Ch]^+$ cation with the protein. Nevertheless, the exclusion of water molecules at the protein surface for anion interaction may exist.

In fact, analysis of the CSP caused by NaCl, Na[Glu] and [Ch]Cl at 0.35M (Fig. 3.7), with the folded state that is negatively charged at the working pH (pI ≈ 4.6), precisely shows this, the overall CSP observed for [Ch][Glu] is more similar to that observed for [Ch][Cl] (Fig. 3.7a *vs* 3.7c) than for Na[Glu] or NaCl. Interestingly, a rather significant CSP of the residues E2, A3, I27 and L28 were also observed in NaCl, Na[Glu] or and [Ch]Cl, which are not seen in [Ch][Glu]. In the folded state, I27 is likely hydrogen bonding with the amide groups of E2 and/or A3, while L28, located at the C-terminus of strand β2 hydrogen bonds to the C-terminal residues of β3[6]. Given that these CSPs are independent of the salt, they should be a consequence of electrostatic screening due to the increase of the ionic strength. The fact that they are absent in [Ch][Glu] indicates that the ionic liquid as an entity has a different effect than the one that can be attributed to its constituent ions separately. Such differences might be related to the fact that, contrary to the common salts, in the IL there is also the possibility

of interactions between the protein and contact ion pairs established by the ions of the ionic liquid[29,34].

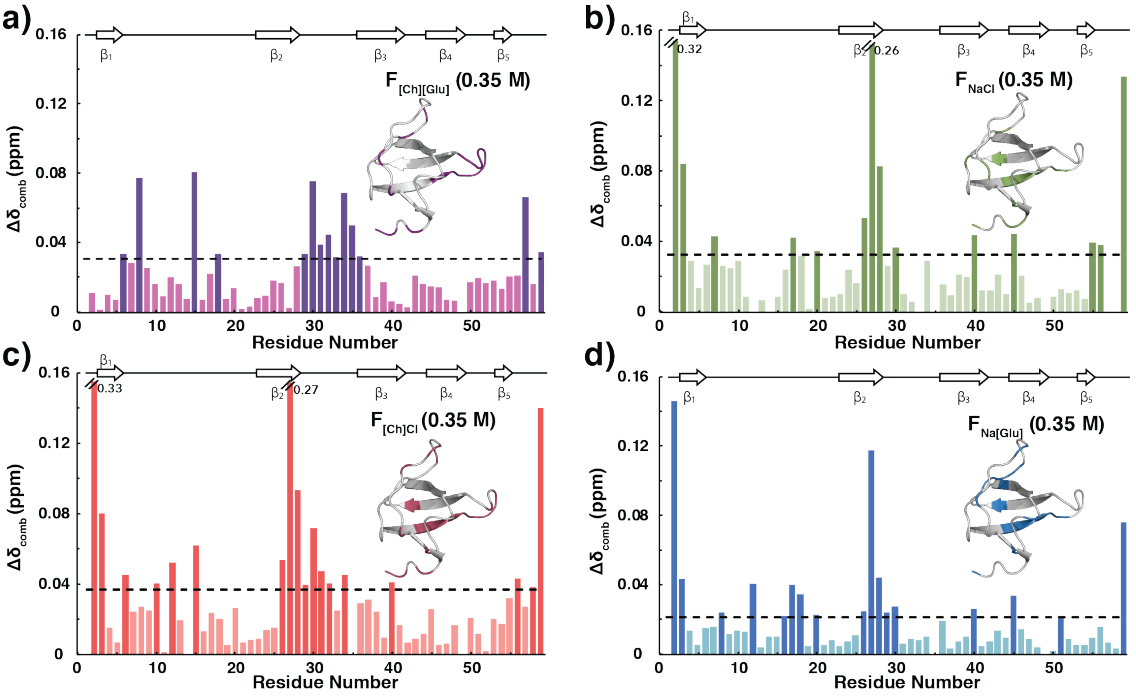



**Figure 3.7. [Ch][Glu] IL and salts–protein interactions in the folded state.**

**a)** [Ch][Glu], **b)** NaCl, **c)** [Ch]Cl, **d)** Na[Glu]. In **b**) and **c)** the magnitude of the CSPs of residues Q2 and L28 is higher than 0.16, therefore, to maintain the same plot scale (for easier comparison) the bars corresponding to these residues were cut and their magnitude value is written at their side. The residues that show a combined chemical shift above the threshold (dashed line) are coloured in a darker tone. Above each plot it is depicted the secondary structure of the protein. The combined chemical shifts were calculated in the presence of [Ch][Glu] or respective salts at 0.35 M of concentration and against the folded chemical shifts in water ($F_{water}$).

Inspection of the structure of drkN SH3 shows that there are two clusters of negatively charged amino acids in close proximity: D14, D15 and E16 and E31, D32 and D33. These clusters are known to be involved in electrostatic interactions between SH3 and other proteins[8,9] and the removal of these interactions leads to destabilisation of the SH3 domain by electrostatic repulsion. Negative electrostatic repulsion on the surface of the SH3 domain has also been implicated in the reduced stability of other SH3 domains[8,35]. Hence, the stabilisation of these negative electrostatic repulsions could lead to the observed overall stabilisation of the protein. Yet, from our data and as observed by Mok et al[15] for the stabilised protein in $Na_2SO_4$, ($F_{Na2SO4}$), in particular that none of a series of single and double mutants of negatively charged residues demonstrate complete stabilisation of the drkN SH3, this shielding, by itself, may not be the major factor contributing to protein stability. As seen above (Fig. 3.7), the $[Ch]^+$-SH3 interactions are not the reason SH3 is stabilised (for [Ch][Cl] the *m*-value is the lowest of the stabilising ones). Additionally, as discussed, the $[Ch]^+$ cation by itself is known to lead to destabilisation[36,37]. Thus, the observed overall stabilisation of SH3 must a

great extent to be attributed to the $[Glu]^-$ anion (similar *m*-values for [Ch][Glu] and Na[Glu], Table 3.1), but likely not from its direct interaction with the protein. Instead, the stability increase of drkN SH3 could be partially explained by the fact that $[Glu]^-$ anion is strongly hydrated and it has a kosmotropic behaviour, that is, with potential to modify the structure of water and/or solution surface tension, and hence the protein hydration[38], as discussed for the acetate anion in a cholinium-based IL that stabilises $\alpha$-chymotrypsin enzyme[39].

Concerning the CSP profile of the folded SH3 in the presence of [Ch][Glu] (Fig. 3.7a), I observe, for the residues above the cut-off value ≈ 0.03 ppm, that the changes in chemical shifts vary between linear and a typical binding isotherm, in line with a specific ion-protein interaction. For these (N29, M30, E31, D32, S34, and N35), which are located in the region between β2 and β3 sheets, I were able to determine an average equilibrium binding affinity ($K_D$) of approximately 0.1 M. Considering this affinity and knowing that: i) SH3 is destabilised in cells in response to osmotic stress, ii) and that is alleviated by adding the osmolyte betaine[21,23], iii) choline and glutamate are two of the most abundant osmolytes in cells[40] (in an exponentially growing *E. coli*, $[Glu]^-$ is found at ~0.1 M[41]), and iv) the cellular environment is crowded with other charged osmolytes; it is possible that, *in cell*, the equilibrium of the drkN SH3 domain is more shifted towards the folded state than that observed *in vitro*. By measuring the temperature dependence of $^{19}F$ resonances on drkN SH3, Smith et al., demonstrate that stability of SH3 in living *E. coli* cells decreases or it is unchanged as compared with buffer[20] because stabilising hard-core repulsions can be completely offset by attractive (charge-charge) interactions. Under crowding, preferential interactions are facilitated with the unfolded ensemble (slow folding) which result in the decrease of protein stability. On the other hand, the fact that $Ch^+/Glu^-$ ions seem to interact more with the folded state, *in vitro*, should result in slower unfolding and consequently in overall protein stabilisation. The relevance *in cell* remains to be studied.

Interestingly, comparing the CSP for the folded states of drkN SH3 in [Ch][Glu] and $Na_2SO_4$ ($F_{[Ch][Glu]}$ and $F_{[Na2SO4]}$, respectively) reveal distinct patterns (Fig. 3.8).

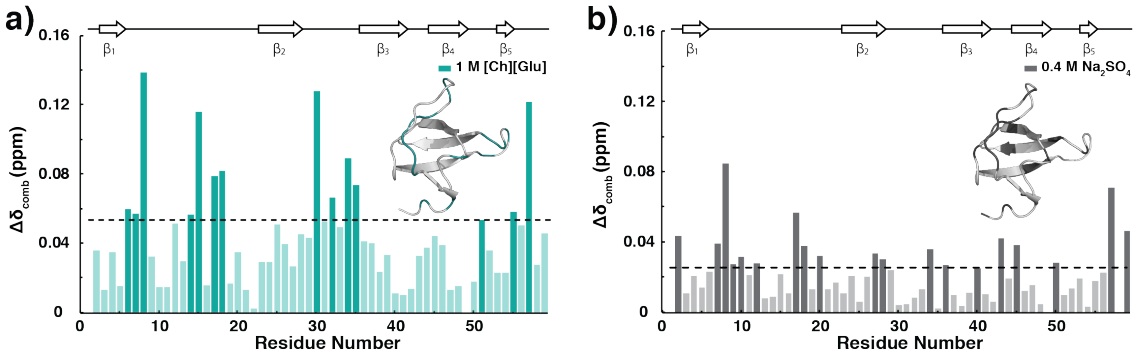



**Figure 3.8. Combined chemical shift of the folded drkN SH3 in the presence of stabiliser IL or salt.**

Combined chemical shift of the folded drkN SH3 in the presence of 1.0 M of [Ch][Glu] (purple) and 0.40 M of $[Na_2SO_4]$ (grey). The dashed lines correspond to the cut-off value determined for both cosolutes and are coloured accordingly. Above the plot it is depicted the secondary structure of the protein. The combined chemical shifts were calculated against the folded chemical shifts in water ($F_{water}$).

While with 1 M of [Ch][Glu] (which leads to ~100% folded protein) I still observe a cluster of CSP of specific residues, particularly located at the loop between β2 and β3 (as described above for 0.35 M), in $Na_2SO_4$, at approximately the same ~100% of folded protein (with 0.4 M of $Na_2SO_4$), I observe a less specific interaction, with the affected residues being spread throughout the protein sequence, with no apparent bias towards the type of residue. This observation hints at a possible different stabilisation mechanism. Besides, while the low magnitude and location of CSP observed suggests the $F_{[Ch][Glu]}$ state should be quite similar to $F_{[Na2SO4]}$ and $F_{water}$ states, the possibility that they may be slightly different may not be ruled out (*vide infra*).

Because the chosen concentration of [Ch][Glu] (0.35 M) to compare with their relative salts (Fig. 3.9) does not lead to a 100% folded sample, I could also examine the effects of this IL on the unfolded structure ($U_{[Ch][Glu]}$). As discussed above, analysis of the CSP of $U_{[Ch][Glu]}$ (Fig. 3-6a-III and reproduced in Fig. 3.9a) shows that some residues of the unfolded state are also affected by [Ch][Glu].

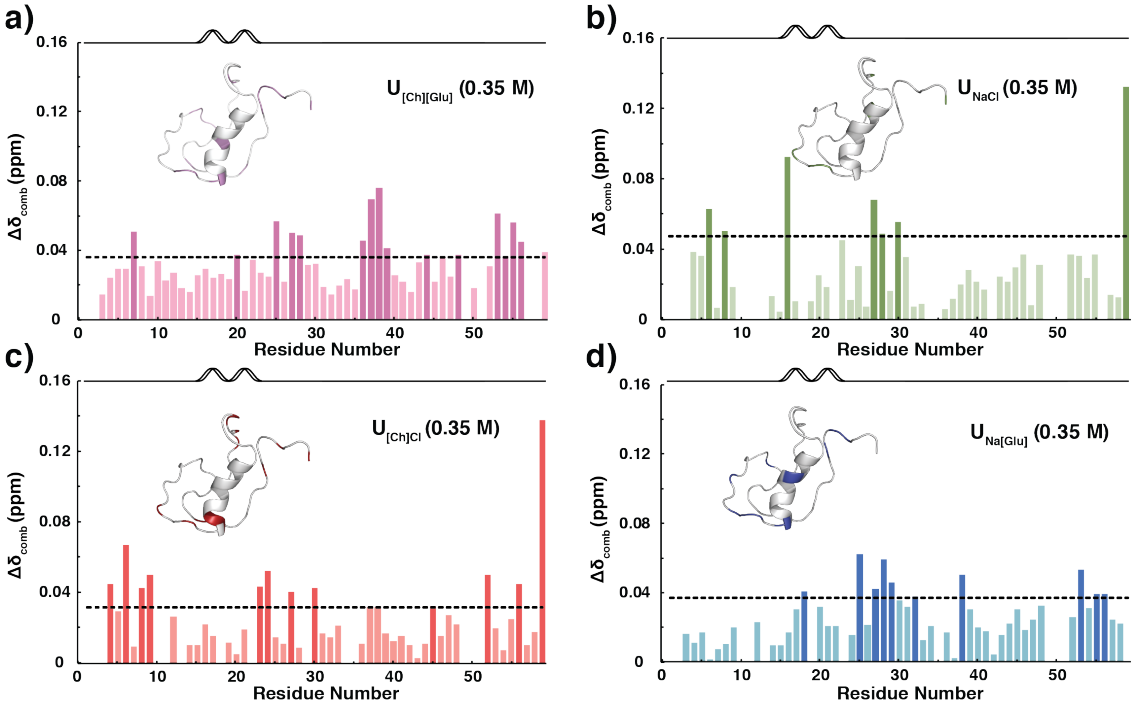



**Figure 3.9. [Ch][Glu]/ionic salts–protein interactions in the unfolded state.**

**a)** [Ch][Glu], **b)** [NaCl], **c)** [Ch][Cl], **d)** [Na][Glu]. The residues that show a combined chemical shift above the threshold (dashed line) are coloured in a darker tone. Above the plot it is depicted the predicted secondary structure of the protein. The combined chemical shifts were calculated in the presence of 0.35 M of [Ch][Glu] and respective salts and against the unfolded chemical shifts in water ($U_{water}$).

Interestingly, when comparing the affected residues in the $F_{[Ch][Glu]}$ and $U_{[Ch][Glu]}$ states (Fig. 3.6a), I observe that they do not overlap (with the exception of W36 and the C-terminal D59), indicating that the IL-protein interactions observed in the folded state are not only a consequence of the amino acid type, but of the structure as well. The affected residues in the

$U_{[Ch][Glu]}$ do not seem to cluster in any specific region of the protein and most of them are hydrophobic in nature (9 out of 16: H6, L25, I27, L28, W36, Y37, A39, I53 and M55) (see Table B1 of Appendix B). This interaction with the hydrophobic residues is expected and a consequence of their exposure to the solvent as a result of protein unfolding. As with the folded state, the low magnitude CSP observed suggests the $U_{[Ch][Glu]}$ state should be similar to that in water ($U_{water}$). Analysis of the CSP caused by the salts NaCl, Na[Glu] and [Ch]Cl (Fig. 3.9b-d) does not allow to identify a preferential interaction from neither the $[Ch]^+$ cation or the $[Glu]^-$ anion, with the observed CSP pattern seeming to be different. Nevertheless, the contribution of the ionic strength (e.g, NaCl, Fig. 3.9b) as well as the result of the Na[Glu]/[Ch]Cl combination (Fig. 3.9c, 9d) cannot be ruled out.

[Bmim][dca]-SH3 interactions

Similar to [Ch][Glu], as previously noted, also the destabilising IL [Bmim][dca] seems to interact with both the folded ($F_{[Bmim][dca]}$) and unfolded ($U_{[Bmim][dca]}$) states of SH3 (Fig. 3.6b-II and b-III). Inspection of the interaction of [Bmim][dca] through the $F_{[Bmim][dca]}$ state (Fig. 3.6b-II and reproduced in Fig. 3.10a) shows that a large number of residues are affected (20 out of 59). These residues are more or less distributed throughout the sequence and there seems to exist no particular bias towards their chemistry (five negatively charged, five polar, uncharged, seven hydrophobic and three positively charged) (see Table B1 of Appendix B for a complete list of the affected residues). Still, most of the affected residues appear to be located at the protein surface.

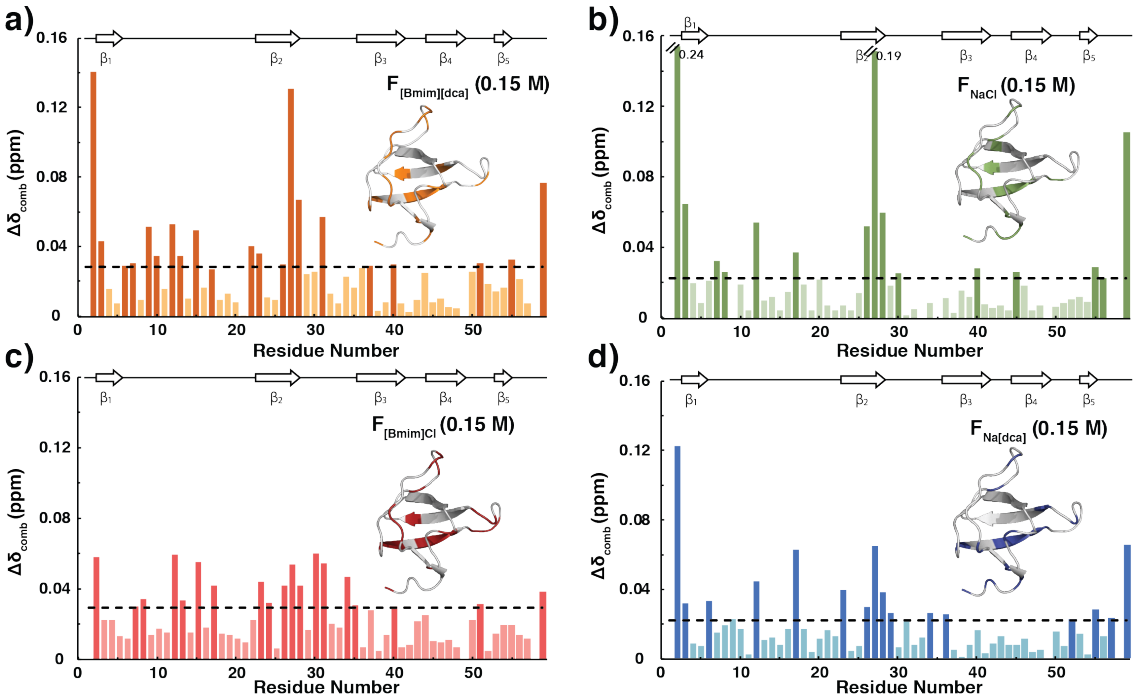



**Figure 3.10. [Bmim][dca]/ionic salts–protein interactions in the folded state.**

**a)** [Bmim][dca], **b)** [NaCl], **c)** [Bmim][Cl], **d)** [Na][dca]. In b) the magnitude of the CSPs of residues Gln2 and Leu28 is higher than 0.16, therefore, to maintain the same plot scale (for easier comparison) the bars corresponding to these residues were cut and their magnitude value is written at their side. The residues that

show a combined chemical shift above the threshold (dashed line) are coloured in a darker tone. Above each plot it is depicted the secondary structure of the protein. The combined chemical shifts were calculated in the presence of 0.15 M of [Bmim][dca] and respective salts and against the folded chemical shifts in water ($F_{water}$).

Noteworthy, comparing the affected residues in the $F_{[Bmim][dca]}$ and $F_{[Ch][Glu]}$ (Fig. 3.6a-II *vs* Fig. b-II or Fig. 3.7a *vs* Fig. 3.10a) reveals that only three are commonly affected (K6, D15 and E31), all located in unstructured loops. Despite not being the same amino acids, the number of negative and polar, uncharged residues perturbed is similar for both [Ch][Glu] and [Bmim][dca], the major differences are observed for the hydrophobic residues, which are in a much higher number for the [Bmim][dca]. These results are consistent with those of Nordwald et al[42], which observed the decrease in the stability of lipase in [Bmim]Cl is due to the direct binding of the $[Bmim]^+$ cation with the surface of the protein. Similarly, we have previously observed[29] that the interaction of the $[Bmim]^+$ cation with hydrophobic groups, typically protected from the solvent as they are located in the core of the protein Im7, leads to its denaturation due to the cation ability to expose them to the solvent. This is comparable to what happens with [Gdm]Cl denaturant agent[43], because $[Gdm]^+$ and $[Bmim]^+$ have similar properties (i.e., flat structure and diffusively distributed positive charge on its ring), the observed destabilisation effect is in line with a preferential binding of the cation to the protein surface. Analysis of the CSP of caused by NaCl, [Bmim]Cl IL and Na[dca] salt (at 0.15 M) (Fig. 3.10b-d) shows that, in fact, that $[Bmim]^+$ cation should be the main responsible for most of the observed perturbations. The remaining CSPs can be mostly attributed to electrostatic interactions, as observed with NaCl (Fig. 3.10b). Thus, while it has been observed that the strong denaturing power of [Bmim][dca] can be related to the high H-bond basicity of the anion[29,44,45], our data shows that the overall CSP of the folded state of SH3 seems to be a combined effect of specific interactions of the $[Bmim]^+$ cation with the hydrophobic residues of the protein, together with electrostatic interactions.

The CSPs of the $U_{[Bmim][dca]}$ reveal a much different scenario (Fig. 3.6b-II and reproduced in Fig. 3.11a), indicating that the unfolded state in [Bmim][dca] should present significant conformational differences from that in water ($U_{water}$). Here, the great majority of the observed CSPs occurs in a particular region of the polypeptide chain, namely between residues Q23 and N29 (see Table B1 of Appendix B). This cluster of residues is particularly interesting since it is located directly next to the region prone to form a non-native $\alpha$-helical structure in water (roughly residues 15-23)[3,8,9,14,15]. The affected region is highly hydrophobic (I24, L25, I27 and L28) and, in the folded state, is part of β2 sheet. The fact that in the presence of [Bmim]Cl and NaCl (Fig. 3.11b, 3.11c) these residues are not significantly affected and the fact that the same region shows a similar CSP pattern in [Na][dca] (Fig. 3.11a *vs* 11d) indicates that the observed effect should be caused by the $[dca]^-$ anion.

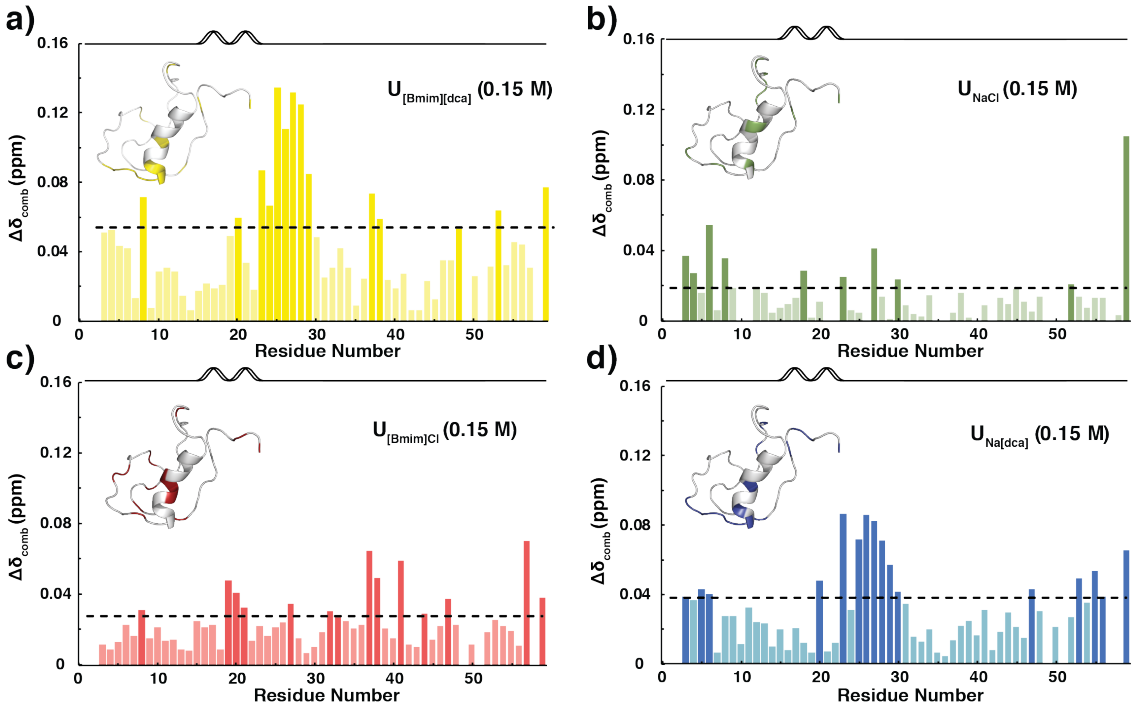



**Figure 3.11. [Bmim][dca]/ionic salts–protein interactions in the unfolded state.**

**a)** [Bmim][dca], **b)** [NaCl], **c)** [Bmim][Cl], **d)** [Na][dca]. The residues that show a combined chemical shift above the threshold (dashed line) are coloured in a darker tone. Above the plot it is depicted the predicted secondary structure of the protein. The combined chemical shifts were calculated in the presence of 0.15 M of [Bmim][dca] and respective salts and against the unfolded chemical shifts in water ($U_{water}$).

In previous work[29,30] I showed that the binding of weakly hydrated anions (such as $[dca]^-$) to positively charged or polar residues leads to the partial dehydration of the protein's backbone groups, and is critical to control stability, explaining why $[dca]^-$ is more denaturing than $[Cl]^-$. Looking at the affected residues I see that there is a positively charged residue, i.e., K26, right in the middle of this area (see below the CSP as function of IL concentration). Thus, our data suggests this residue will be the one interacting with the $[dca]^-$ anion and this interaction then drives the observed CSP of the adjacent residues. At the same time, I also see that the $[Bmim]^+$ cation is the one responsible for the CSP of residues Y37 and R38 (Fig. 3.11c). This may be due to a possible hydrophobic interaction with the aromatic side chain of Y37. Although the difference in the magnitude of CSP comparing [Bmim][dca] and Na[dca] (Fig. 3.11a *vs* 11d) is small, it supports the result of a more hydrophobic cation-anion pair than the anion alone.

Since unfolded protein has significant amounts of exposed hydrophobic surface which can interact irreversibly at sufficiently high concentrations such as those normally utilized in NMR samples, precipitation can be a sign of protein instability (aggregation), especially if it occurs over time. Thus, the absence of precipitation observed in the presence of [Bmim][dca] might reflect a stabilisation of the unfolded state rather than the destabilisation of the folded state[29].

To probe the possible stabilising effects of the shielding of the positive charge at K26 I have estimated the non-native helical structure of different SH3 mutants (K26A, K26G, K26M and K26Q) using the software AGADIR[15,46–48]. In this case replacing K26 by neutral residues with polar or hydrophobic side chain to decrease electrostatic repulsion. The amino acid sequences of the wild-type and mutants were input to AGADIR for analysis (http://agadir.crg.es) and a temperature of 298 K and ionic strength of 0.15 M were used for the calculations. The predicted $\alpha$-helical contents as a function of residue are shown in Fig. 3.12. Note that, as expected for the unfolded state, the helix propensities are all relatively low (< 7%). However, it is striking that when replacing the positively charged K26 by a neutral and hydrophobic residue (through alanine, K26A, or methionine, K26M), the helical content increases substantially (Fig. 3.12a), particularly around the mutation, and in almost perfect agreement with the observed CSP, thus corroborating our conclusions.

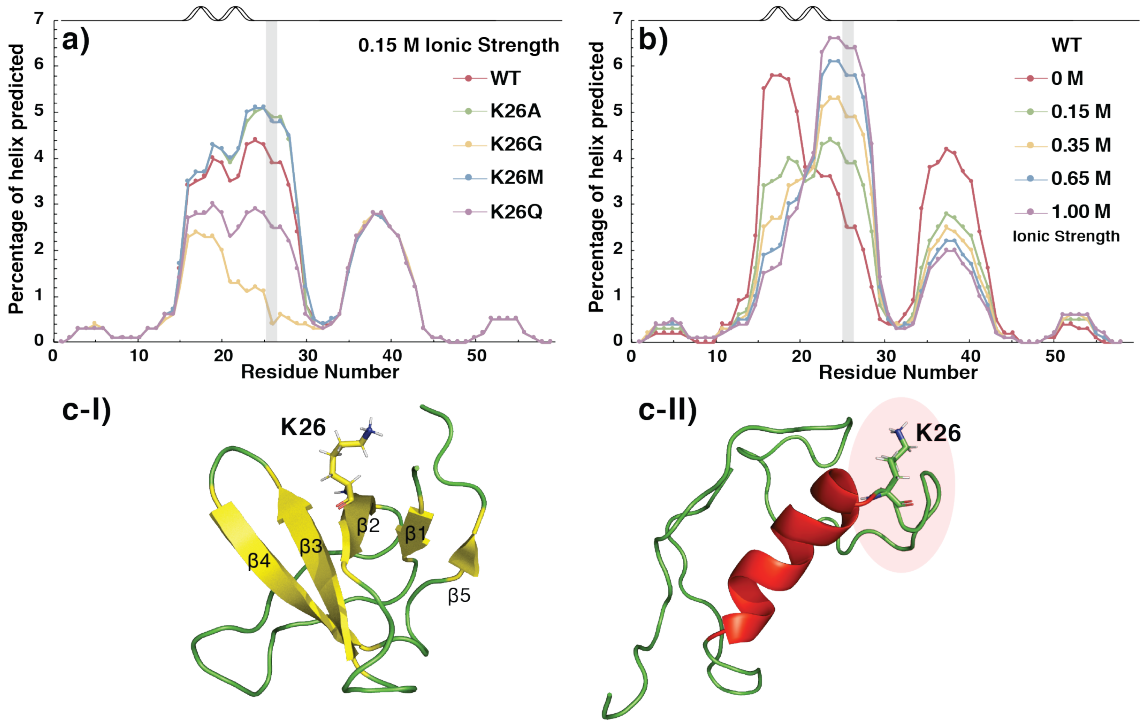



**Figure 3.12. AGADIR prediction of the fractional $\alpha$-helical population of drkN SH3 domain as a function of residue.**

**a)** Wild-type (WT - red) versus mutants (K26A – green; K26G – yellow; K26M – blue and K26Q – purple); ionic strength 0.15 M. **b)** Wild-type at different ionic strengths (0 M – red; 0.15 M– green; 0.35 M – yellow; 0.65 M – blue and 1.0 M – purple). The predictions were performed at 298 K, pH = 7. Above the plot it is depicted the predicted secondary structure of the protein. Residue 26 is highlighted by a grey box. **c**) 3D structure for the **I**) folded state (PDB: 2A36[6]) and **II**) representative of the unfolded ensemble in water (PED-8AAC, #576[3]), the K26 is highlighted by sticks.

To test the effect of increasing the ionic strength on the helical propensity of SH3 (Fig. 3.12b) I used AGADIR, as before. Remarkably, upon addition of salt, the percentage of helix predicted for the region immediately after the residue Q23 (the same that experiences

significant CSP) increases significantly (see Fig. 3.12c for representative structure of WT in water), and in a concentration-depended manner, in complete agreement with our experimental observations. The region around Y37 is also important to highlight since it retains some helical tendency[10,15,17] that is not affected by mutation of K26 (Fig. 3.12a) but is severely diminished by an increase in the ionic strength (Fig. 3.12b). This can indicate that this hydrophobic patch becomes even more solvent-exposed and available to interact with the $[Bmim]^+$ cation.

In summary, the observed stronger destabilising properties of this IL as compared to that of its salts (Fig. 3.5b and Table 3.1) is in line with a cooperative effect of the $[Bmim]^+$ cation with the $[dca]^-$ anion where the first destabilises the folded structure via the direct binding to the protein surface while the latter seems to stabilise the unfolded structure via a stabilisation of a particular hydrophobic patch, located in a region directly adjacent to the non-native $\alpha$-helical structure observed in water. The stabilisation of the unfolded state may by further increased by the hydrophobic contacts with the ion-pair which stabilise a particular hydrophobic patch. The existence of a residual structure in the [Bmim][dca]-unfolded ensemble is validated using secondary structure propensities (discussed below in the section of structure perturbations).

It is interesting to notice that, albeit only 0.15 M of [Bmim][dca] are sufficient to shift the equilibrium to 86% unfolded (Fig. 3.3b), however a dependence of the chemical shifts of the unfolded state with the concentration of the IL is still observed (Fig. 3.13). Although the dependence of CSP for the segment most affected (Q23-M30) on IL concentration resembles a binding isotherm curve (see Fig. 3.13b), the derived average $K_D$ value (residues I24, L25, K26, I27, L28) for the interaction is ~ 0.55 M, which is consistent with weak, non-specific binding with the protein.

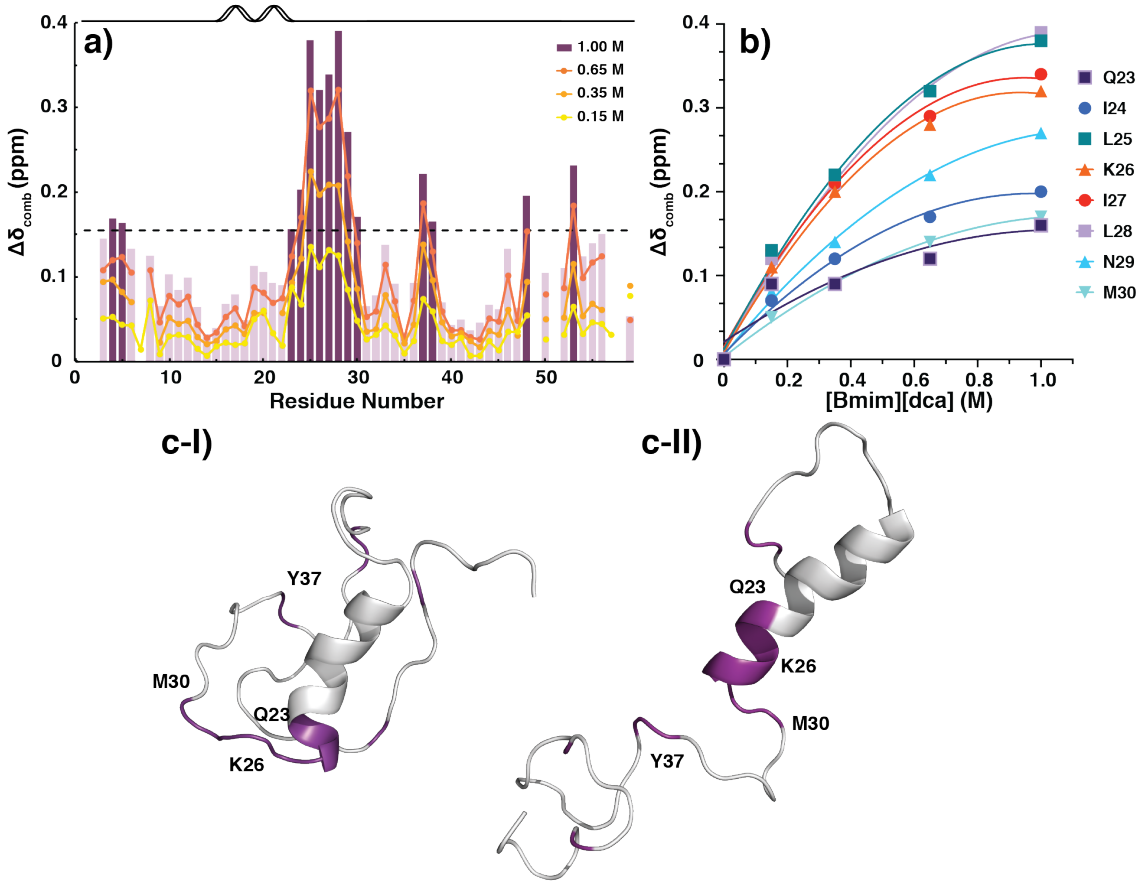



**Figure 3.13. Effect of [Bmim][dca] on the backbone amides of the unfolded drkN SH3.**

Overlay of the combined chemical shift of the unfolded drkN SH3 in the presence of increasing concentrations of [Bmim][dca] (0.15 M (yellow), 0.35 M (light orange), 0.65 M (dark orange) and 1.0 M (dark purple bars)). The dashed line corresponds to the cut-off value determined for 1.0 M IL. Above the plot it is depicted the predicted secondary structure of the protein. **b**) Plot of $\Delta\delta_{comb}$ as function of [Bmim]dca], concentration for Q23-M30 residues. The lines are indicative of a non-linear fitting. **c**) 3D representative structures of the unfolded ensemble in water: **I**) structure 576 and **II**) structure 1195 from PED-8AAC ensemble[3]). The affected residues (purple) are mapped onto both structures.

Considering the various structures within the unfolded ensemble[3,12], I realise that the segment most affected by the IL could be an extension of the non-native helical segment of $U_{exch}$ (Fig. 3.13c-I *vs* c-II) which led me to perform additional experiments to validate that (see below: temperature dependence and structure perturbations sections). It is clear that that the observed effects are not a simple consequence of the ionic strength, in agreement with previous observations from Mok et al[15]. The authors showed that, albeit electrostatic shielding could play a role in the stability as well as in the definition of the unfolded state of the SH3 domain, mutations in highly conserved residues (e.g., on T22 that corresponds to an otherwise conserved glycine residue in the diverging β-turn) are much more relevant than mutations of negatively surface charged residues.

In order to determine if the action of [Bmim][dca] has some similarities with other commonly used denaturants, I compared the unfolded states of drkN SH3 upon addition of

1 M [Bmim][dca] or 2 M [Gdm]Cl ($U_{[Bmim][dca]}$ or $U_{[GdmCl]}$, respectively) which reveal very distinct CSP patterns (Fig. 3.14). In contrast to the sharp and intense resonances of unfolded SH3 in IL, under [Gdm]Cl I found that a number of resonances experience severe line broadening which could imply some restriction of conformational sampling within the ensemble, preventing also unambiguously assignment of all residues. Even though the concentration of [Gdm]Cl is twice that of [Bmim][dca], the magnitude of the observed chemical shifts for the latter is much higher, indicative of a stronger interaction. Also, the magnitude of the affected *vs* non-affected residues by the IL is much higher than that of [GdmCl].

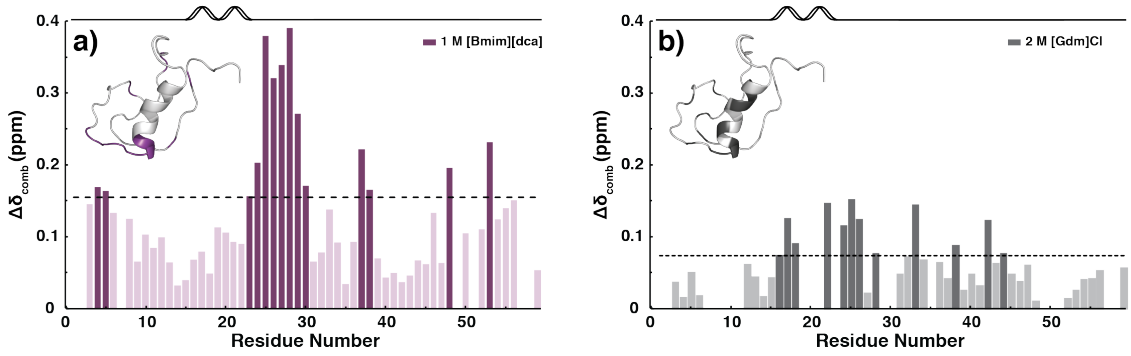



**Figure 3.14. CSP of the unfolded drkN SH3 in the presence of [Bmim][dca] and [Gdm]Cl.**

Combined chemical shift of the unfolded drkN SH3 in the presence of 1.0 M [Bmim][dca] (purple) and 2.0 M of [GdmCl] (dark purple). The dashed lines correspond to the cut-off value determined for both cosolutes and are coloured accordingly. Above the plot it is depicted the predicted secondary structure of the protein. The combined chemical shifts were calculated against the unfolded chemical shifts in water ($U_{water}$).

In fact, the unfolded state of drkN SH3 in a high-concentrated [Gdm]Cl solution has been extensively studied[5,6,8,9,15,33,49]. It has been noticed, using variable-temperature experiments, that upon addition of [Gdm]Cl the region encompassing residues Q23-L28 in the unfolded state is destabilised[9] and leads to the loss of the preferential structure in this region ($\alpha$-helix propensity) to a random-coil like structure as compared with the $U_{exch}$ state. This seems to be in contrast with the observations made in [Bmim][dca] and it indicates that the unfolded state in both conditions may be significantly different (*vide infra*).

**Temperature dependence**

As analysed for protein GB1 in chapter 2, an amide proton temperature coefficient ($\Delta\sigma_{HN}/\Delta T$) corresponds to the slope of a plot of $^1H^N$ upfield shift against increasing temperature, as an indicator of intra-protein hydrogen bonding[50,51]. An upfield shift of the amide proton resonance is, in general, due to a weakening of the hydrogen bond associated to a larger thermal motion and it is more pronounced for intermolecular interactions. Therefore, chemical shifts of protons involved in intermolecular H-bonds show stronger temperature dependence (larger coefficient, < -7 ppb/K) than those of protons involved in intramolecular hydrogen bonds (smaller/less negative coefficient > -4.6 ppb/K). To probe IL-induced changes in the amide H-bond network of both folded and unfolded states, associated to

secondary structural changes, I have first determined amide temperature coefficients for SH3 in water and in the 0.35 M [Ch][Glu] solution, and latter for 0.15 M [Bmim][dca], by varying the temperature between 278 and 313 K.

Amide coefficients in the presence of [Ch][Glu]

The HSQC spectra of SH3 acquired in water and 0.35 M [Ch][Glu] from 278 to 313 K of temperature as well as the corresponding plot of $^{1}H^{N}$ chemical shift versus temperature for each residue of SH3 at IL can be found in the Fig. 3.15.

Examination of the resultant coefficients in water and 0.35 M [Ch][Glu] (subsequent Fig. 3.16), shows no significant modifications in the H-bond network due to the minor changes in $\Delta\sigma_{HN}/\Delta T$ values along the protein sequence. Nevertheless, the amide coefficients of I24 and W36 residues reveal significantly larger or smaller (less negative) values, respectively, in the presence of IL which are associated to a decrease or to an increase, respectively, in the probability to participate in an intra-protein hydrogen bond. Analysis of the 3D structure of SH3 shows that I24 is located at the N-terminal of β2, with its amine group making a possible hydrogen bond with the sidechain carbonyl of Q23 (Fig. 3.16c-I). However, analysing all the structures in the NMR ensemble, I see that this only happens in 3 of the 10 structures. Therefore, the $F_{[Ch][Glu]}$ structure is possible stabilised in such a way that it hinders the formation of this hydrogen bond, with no penalty for the overall stability of the protein. Concerning the amide group of W36, this may participate in a hydrogen bond with the carbonyl group of D33 (Fig. 3.16c-II) and it may help to stabilise the loop between β2 and β3 sheets, contributing to the observed overall stability of the protein.

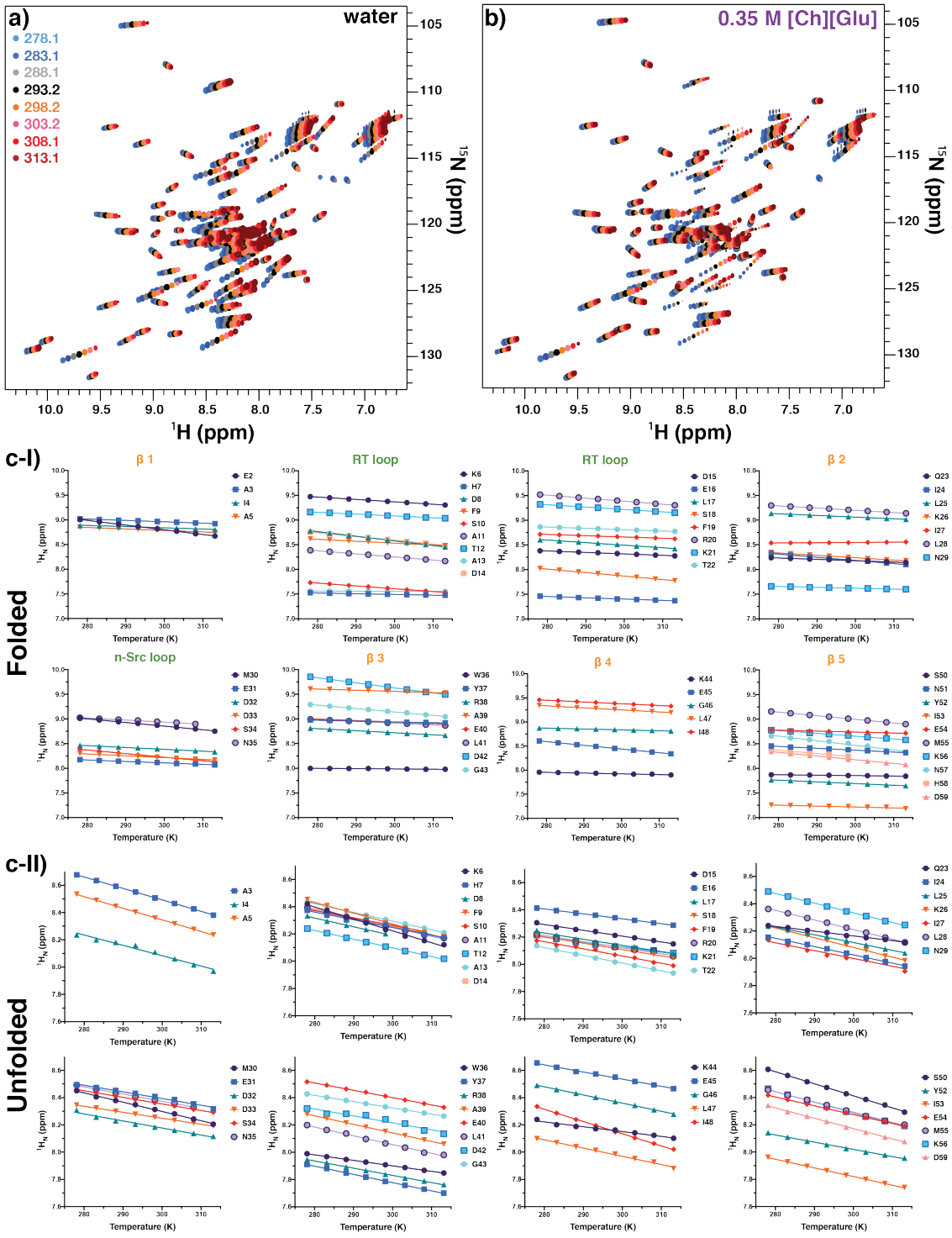



**Figure 3.15. Temperature dependence of SH3 in water and [Ch][Glu].**

Overlay of 2D $^{1}$H-$^{15}$N HSQC spectra of drkN SH3 acquired at different temperatures (278.1, 283.1, 288.1, 293.2, 298.2, 303.2, 308.1 and 313.1 K) - from blue to red color – in **a)** water, **b)** 0.35 M [Ch][Glu]. The change in the upfield $^{1}H_N$ chemical shifts as function of temperature increase[51] corresponds to the amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$). **c**) For each backbone amide at 0.35 M [Ch][Glu], **I**) folded or **II**) unfolded, a linear least-squares fit (lines) of $^{1}H_N$ chemical shifts at 8 temperatures were performed. For each residue, the δ $^{1}H_N$ was extracted from the peak in the 2D $^{1}$H-$^{15}$N HSQC measured at a given temperature. The slope (δ $^{1}H_N$ versus temperature) corresponds to the amide proton temperature coefficient ($\Delta\sigma_{HN}/\Delta T$)

and its standard deviation represents uncertainty. Each plot corresponds to the group of residues that are part of that secondary structure.

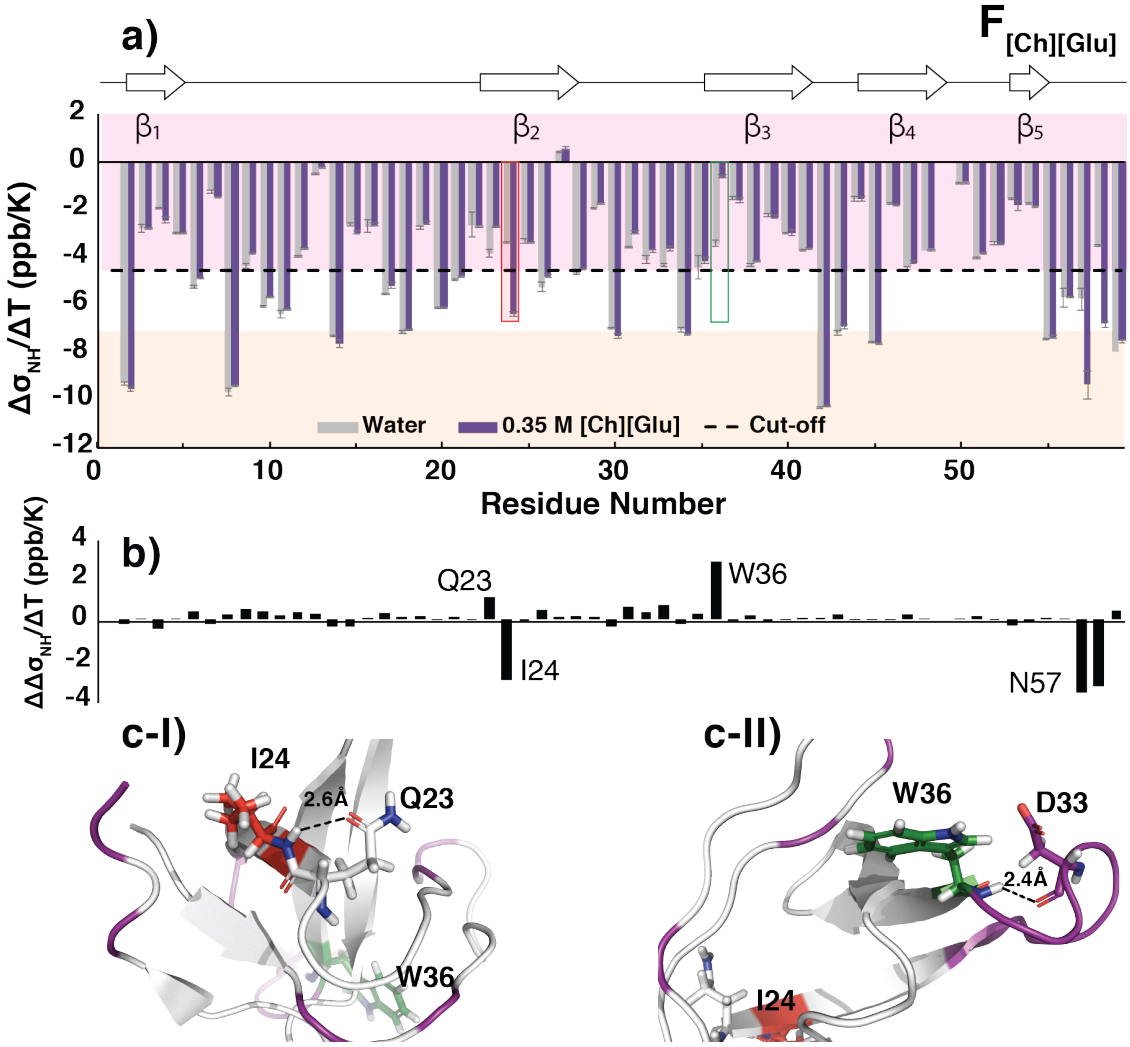



**Figure 3.16. Amide proton temperature coefficients for the folded state of SH3 in [Ch][Glu].**

**a)** Amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) for the folded state of SH3 in [Ch][Glu]. Grey and purple bars correspond to the coefficients in water and [Ch][Glu], respectively. The horizontal dashed line corresponds $\Delta\sigma_{HN}/\Delta T = -4.6$ ppb/K. Bars ending in the pink box have a ≥85% probability of participating in an intramolecular hydrogen bond. Bars ending in an orange box have a ≤20% probability of participating in an intramolecular hydrogen bond. Values were determined using a linear least-squares fit of $^{1}H_N$ chemical shifts from 278 to 313 K in 5 K increments. Uncertainties represent the slope standard deviation. Above the plot it is depicted the secondary structure of the protein. Red and green rectangles highlight residues I24 and W36, respectively. **b)** difference between the amide proton temperature coefficients in water and in [Ch][Glu] ($\Delta\Delta\sigma_{HN}/\Delta T$). **c)** Residues showing the largest change (with exception of the N-terminal) are highlighted in white (Q23), red (I24) and green (W36) and mapped on the structure of SH3. The potential H-bonds are represented by dashed lines. The residues that suffer significant CSP are highlighted in purple.

Overall, from the results I hypothesise that the observed perturbation in chemical shift of n-Src loop (Fig. 3.6a) is associated to the stabilisation of the negative electrostatic

repulsions (caused by the $[Ch]^+$ cation), and that could be allied with a modification of the water structure (preferential hydration?) around the protein, caused by the $[Glu]^-$ anion, as the changes in H-bond network are very small (Fig. 3.16b), leading to the global observed protein stabilisation.

On the other hand, inspection of the temperature coefficients for the unfolded state in the presence of [Ch][Glu] (Fig. 3.17) shows that a considerable number of residues is affected, but most of them insignificantly. Interpretation for the coefficients for the unfolded state must be taken carefully since temperature gradients are poor predictors of hydrogen bonding in unstructured peptides and in proteins undergoing conformational exchange, both of which apply to unfolded SH3[22].

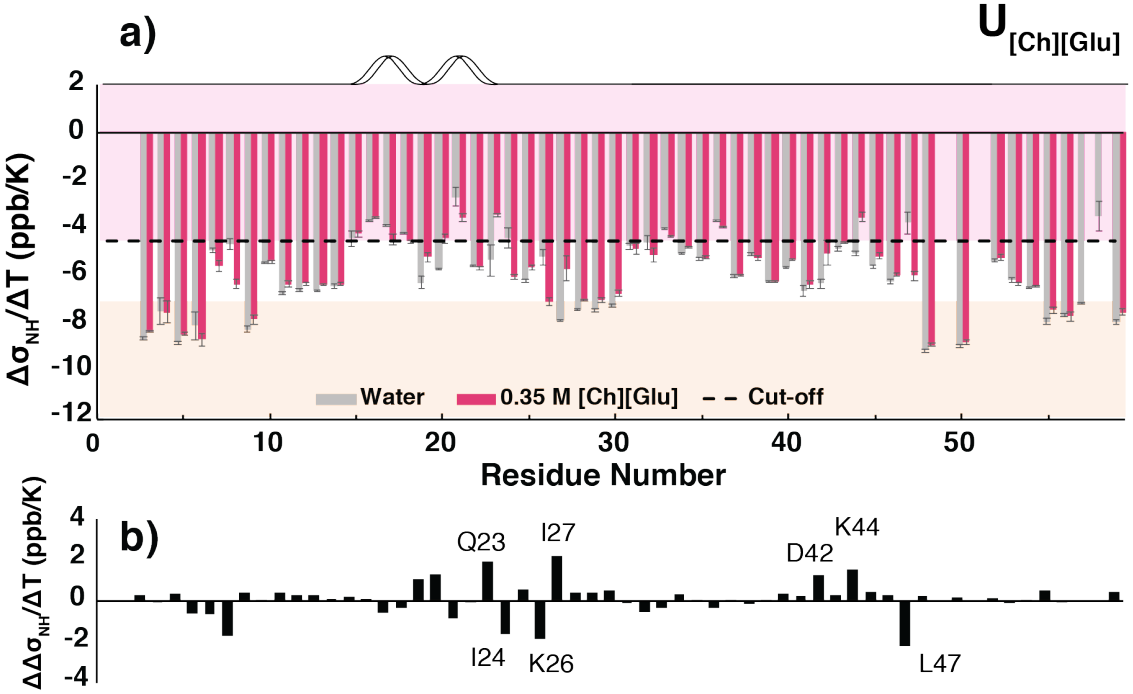



**Figure 3.17. Amide proton temperature coefficients for the unfolded state of SH3 in [Ch][Glu].**

**a)** Amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) for the unfolded state of SH3 in [Ch][Glu]. Grey and pink bars correspond to the coefficients in water and [Ch][Glu], respectively. The horizontal dashed line corresponds $\Delta\sigma_{HN}/\Delta T = -4.6$ ppb/K. Bars ending in the pink box have a ≥85% probability of participating in an intramolecular hydrogen bond. Bars ending in an orange box have a ≤20% probability of participating in an intramolecular hydrogen bond. Values were determined using a linear least-squares fit of $^1H_N$ chemical shifts from 278 to 313 K in 5 K increments. Uncertainties represent the slope standard deviation. Above the plot it is depicted the predicted secondary structure of the protein. **b)** difference between the amide proton temperature coefficients in water and in [Ch][Glu] ($\Delta\Delta\sigma_{HN}/\Delta T$).

As noted in the introductory section of this chapter, under aqueous-buffer conditions, the unfolded state of SH3 has been previously characterized[5,6,8,9,15,49,52] and shown to have a propensity for the formation of a non-native $\alpha$-helix in the region between residues D15 to Q23[3,8,9,14,15]. Furthermore, based on the presence of sequential amide-amide NOE cross-peaks

for resonances of the unfolded state, Zhang et al[8] have speculated that the formation of preferential turn-like structure in residues N35 to I48 may nucleate the β-sheet containing strands β3 and β4, thus allowing long-range interactions between residues of β3 and β4 and catalysing the formation of the β-sheet hydrogen bonding interactions. Thus, while the variations in the temperature coefficients in the unfolded state are difficult to interpret, it is interesting to notice that most affected amide coefficients by [Ch][Glu] (Fig. 3.17b) are located in the region prone to form the non-native *α*-helical structure (F19, R20, Q23, I24, K26, and I27), where most of them have larger coefficients with IL which are associated to a breaking of H-bond, and the region with preferential turn-like structure (D42, K44 and L47), considered to be important in the stabilisation of the $U_{exch}$ and nucleation of the $F_{exch}$ states, respectively[3,8]. In face of the overall increase in the population of the folded state (stabilisation of the protein), I hypothesise that for the unfolded state, a modification of the H-bond network at these locations could, in the one hand hinder the formation of the non-native helical structure (due to the large coefficients found in this segment with IL) and, at the other hand, facilitate the formation of the turn-like structure (due to the smaller coefficients of D42 and K44 with IL), thus helping the nucleation of the β-sheet structure.

Amide coefficients in the presence of [Bmim][dca]

The HSQC spectra of SH3 acquired in water and 0.15 M [Bmim][dca] from 278 to 313 K of temperature as well as the corresponding plot of $^{1}H^{N}$ chemical shift versus temperature for each residue of SH3 at IL can be found in the Fig. 3.17.

Analysis of the resultant temperature coefficients (Fig. 3.18) shows that the H-bond network of the folded state is not significantly affected which is in line with a similar overall conformation of the $F_{water}$ and $F_{[Bmim][dca]}$ states. However, some key residues experience a decrease in their potential to participate in intra-protein H-bonds. In particular residues Q23 and I24 (larger/ more negative values with IL), which are located at the N-terminal of β2, E31, located in the n-Src loop, and W36, located N-terminal residue of β3. Surprisingly, this behaviour is the opposite of that observed in the $F_{[Ch][Glu]}$ (Fig. 3.16). An increase in the coefficient for I27 is also observed. This residue is located in β2 and it interacts with the protein's N-terminal (E2 and A3) and, in fact, I do observe that the amide chemical shifts of these residues are among the most affected (Fig. 3.6b). Following the same line of though as above, it is possible that the presence of [Bmim][dca] leads to conformational modifications, hindering H-bonding, at the level of the n-Src loop and the β-sheet structure which ultimately leads to a destabilisation of the folded state.

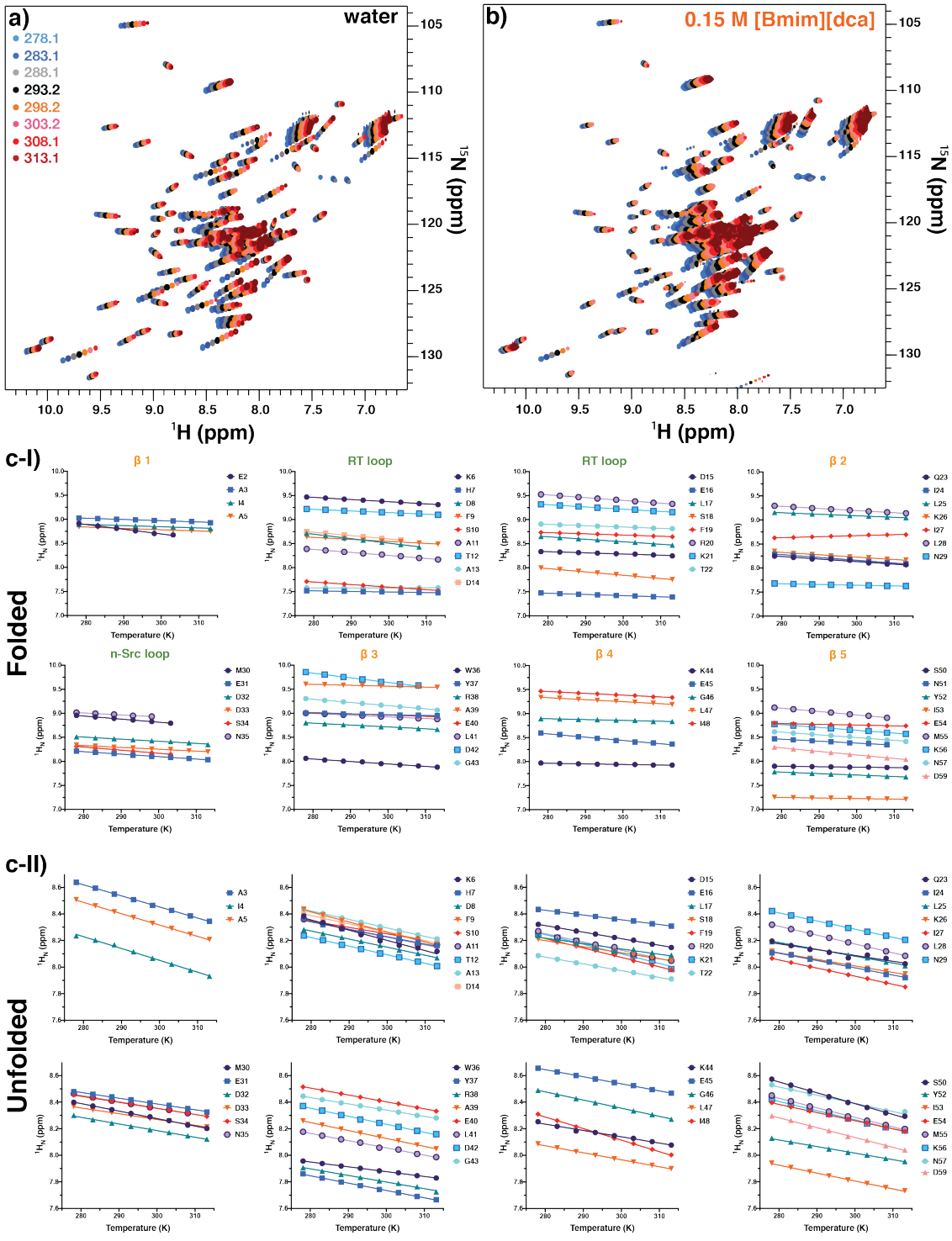



**Figure 3.18. Temperature dependence of SH3 in water and [Bmim][dca].**

Overlay of 2D $^{1}$H-$^{15}$N HSQC spectra of drkN SH3 acquired at different temperatures (278.1, 283.1, 288.1, 293.2, 298.2, 303.2, 308.1 and 313.1 K) - from blue to red color – in **a)** water and **b)** 0.15 M [Bmim][dca]. The change in the upfield $^{1}H_N$ chemical shifts as function of temperature increase[51] corresponds to the amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$). **c**) For each backbone amide at 0.15 M [Bmim][dca], I) folded or II) unfolded, a linear least-squares fit (lines) of $^{1}H_N$ chemical shifts at 8 temperatures were performed. For each residue, the δ $^{1}H_N$ was extracted from the peak in the 2D $^{1}$H-$^{15}$N HSQC measured at a given temperature. The slope (δ $^{1}H_N$ versus temperature) corresponds to the amide proton temperature coefficient

($\Delta\sigma_{HN}/\Delta T$) and its standard deviation represents uncertainty. Each plot corresponds to the group of residues that are part of that secondary structure.

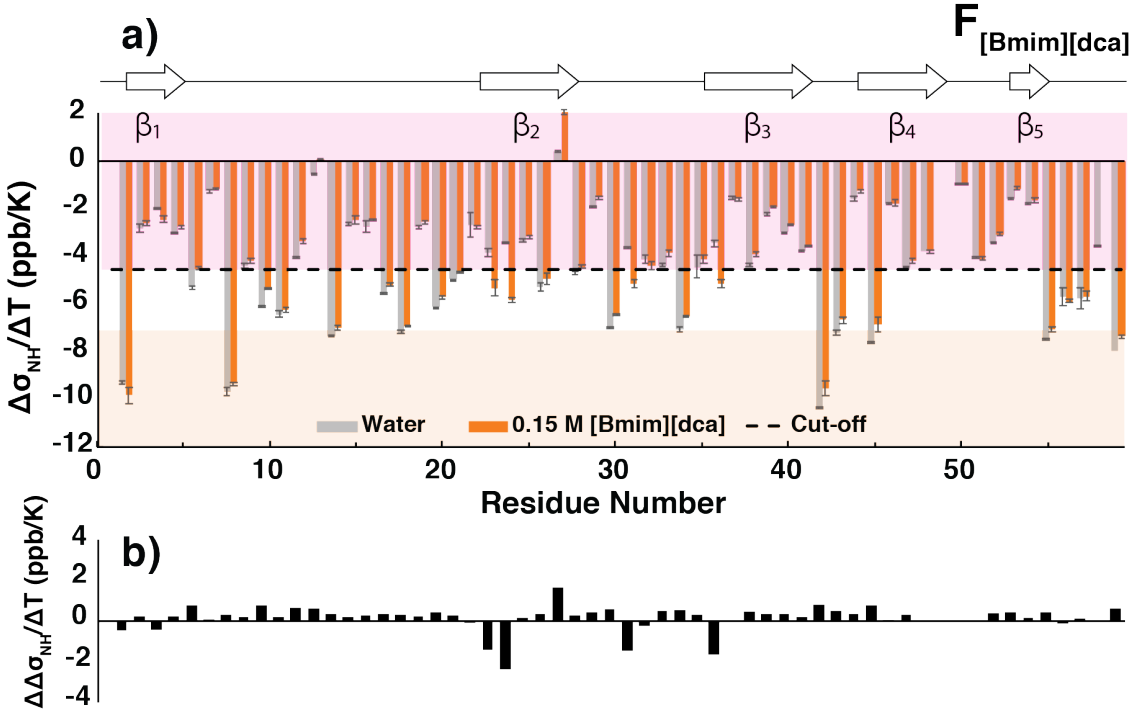



**Figure 3.19. Amide proton temperature coefficients for the folded state of SH3 in [Bmim][dca]**

**a)** Amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) for the folded state of SH3 in [Bmim][dca]. Grey and orange bars correspond to the coefficients in water and [Bmim][dca], respectively. The horizontal dashed line corresponds $\Delta\sigma_{HN}/\Delta T = -4.6$ ppb/K. Bars ending in the pink box have a ≥85% probability of participating in an intramolecular hydrogen bond. Bars ending in an orange box have a ≤20% probability of participating in an intramolecular hydrogen bond. Values were determined using a linear least-squares fit of $^1H_N$ chemical shifts from 278 to 313 K in 5 K increments. Uncertainties represent the slope standard deviation. Above the plot it is depicted the secondary structure of the protein. **b)** difference between the amide proton temperature coefficients in water and in [Bmim][dca] ($\Delta\Delta\sigma_{HN}/\Delta T$).

Concerning the unfolded state, an increase of the overall temperature coefficients (Fig. 3.20) is observed, in particular, the residues Q23 to N29 have significant larger amide coefficients which are associated to an increase of probability to participate in an intramolecular hydrogen bonding (pink region of Fig. 3.14). This is corroborated with the observed perturbation of the chemical shifts for that segment (Fig. 3.14). These observations hint at a possible structural modification of the unfolded structure, such as an residual and extended $\alpha$-helical structure (Fig. 3.14c), as compared to that observed in water.

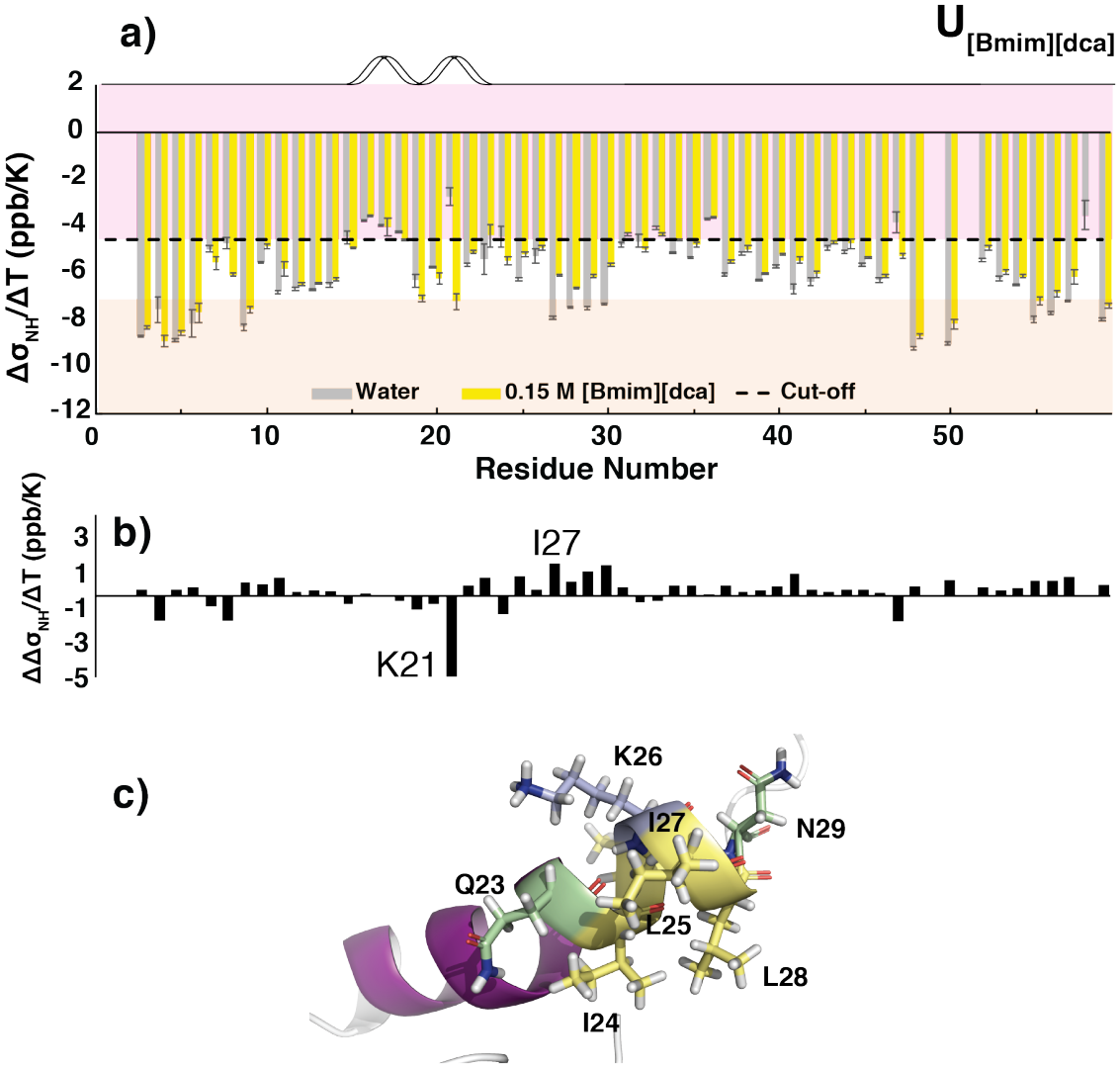



**Figure 3.20. Amide proton temperature coefficients for the unfolded state of SH3 in [Bmim][dca].**

**a)** Amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) for the unfolded state of SH3 in [Bmim][dca]. Grey and yellow bars correspond to the coefficients in water and [Bmim][dca], respectively. The horizontal dashed line corresponds $\Delta\sigma_{HN}/\Delta T = -4.6$ ppb/K. Bars ending in the pink box have a ≥85% probability of participating in an intramolecular hydrogen bond. Bars ending in an orange box have a ≤20% probability of participating in an intramolecular hydrogen bond. Values were determined using a linear least-squares fit of $^1H_N$ chemical shifts from 278 to 313 K in 5 K increments. Uncertainties represent the slope standard deviation. Above the plot it is depicted the predicted secondary structure of the protein. **b)** Difference between the amide proton temperature coefficients in water and in [Ch][Glu] ($\Delta\Delta\sigma_{HN}/\Delta T$). **c)** 3D representative structures of the unfolded ensemble (PED-8AAC, #1195[3]) where residues Q23-N29 are represented by sticks.

### Structure perturbations

As seen above, the data from the drkN SH3 domain chemical shift dependence on the cosolutes indicates that there is a possibility that the structures of the $F_{[Ch][Glu]}$ and $U_{[Bmim][dca]}$ could diverge with lesser or greater extend from those in water ($F_{water}$ and $U_{water}$).

To determine the impact of the different ILs on the structure of the protein, using triple-resonance NMR experiments, I assigned the backbone and aliphatic $H^N$, $N^H$, Cα, Cβ and CO chemical shifts for the folded and unfolded states in water ($F_{water}$ and $U_{water}$, respectively), as well as those of the fully stabilised folded and unfolded states ($F_{[Ch][Glu]}$ and $U_{[Bmim][dca]}$, respectively). Backbone chemical shifts of fully stabilised folded and unfolded drkN SH3 states, in 1 M [Ch][Glu] and 1 M [Bmim][dca], respectively, can be found in Table B2 of Appendix B. Fig. 3.21 shows the assignment of NH amides of backbone in water (21a), 1 M [Ch][Glu] (21b) and different concentrations of [Bmim][dca] (0.35, 0.65 and 1.0 M) (21c). The Cα and Cβ chemical shifts were then analysed to determine whether these changes corresponded to a change in secondary structure propensity (SSP) using the SSP program[53].

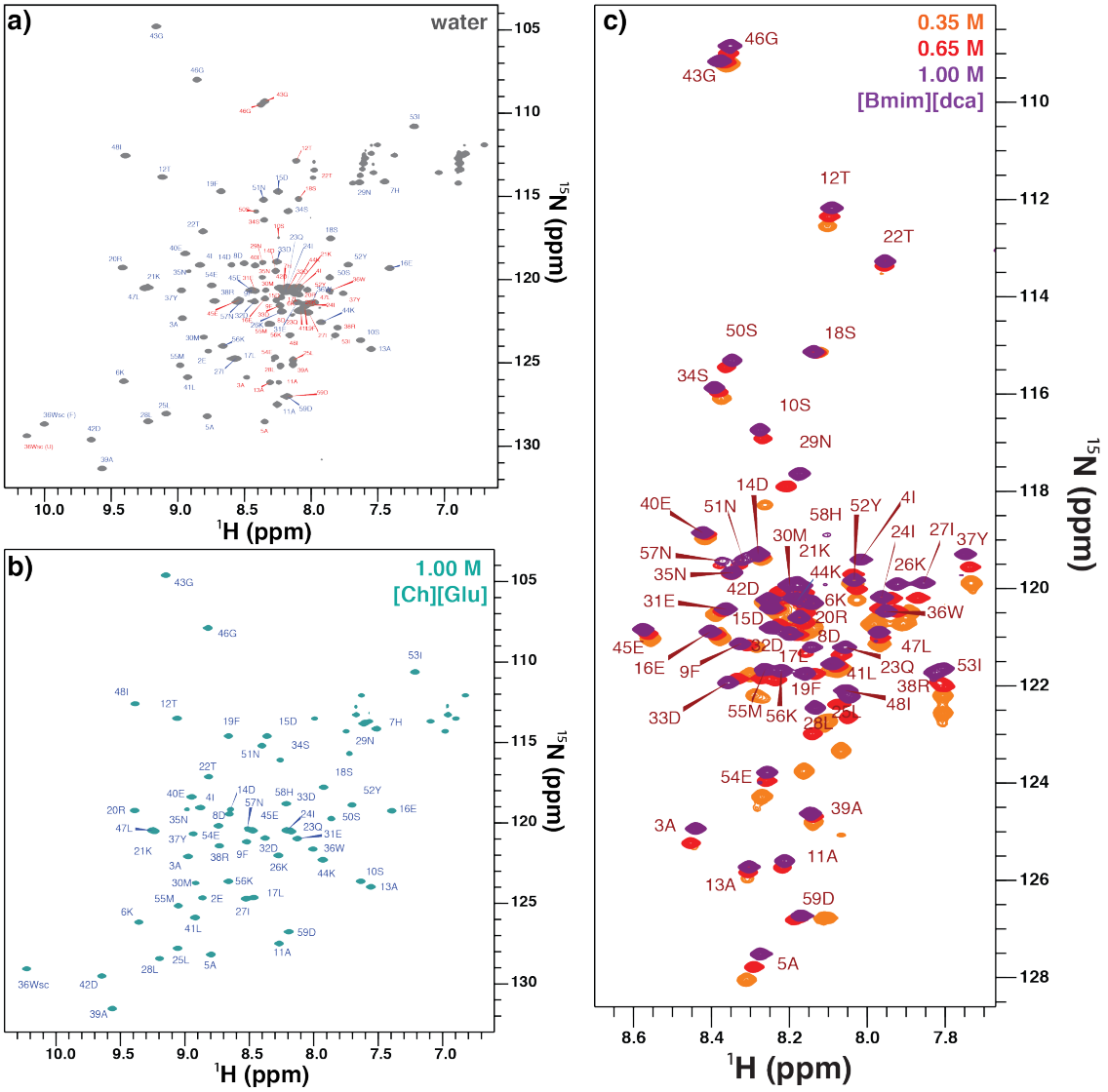



**Figure 3.21. Folded and Unfolded ensemble NH backbone assignments in water and IL conditions.**

**a)** $^1$H-$^{15}$N Spectrum of drkN SH3 in water (gray) with label of folded (blue) and unfolded (red) assignments. **b)** $^1$H-$^{15}$N Spectrum of drkN SH3 in 1.0 M [Ch][Glu] (green) – fully folded stabilised state - with label of folded (blue) assignments. **c)** Overlay of $^1$H-$^{15}$N spectra of drkN SH3 in 0.35 M (orange), 0.65 M (red) and 1.0

M [Bmim][dca] (purple) – fully unfolded stabilised state - with label of unfolded (red) assignments for the final concentration.

Figure 3.22 shows the predicted SSP scores, representing the expected fraction of $\alpha$- or $\beta$-structure (positive and negative values, respectively), for the folded and unfolded states of SH3 in the presence of water ($F_{water}$ and $U_{water}$, respectively), [Ch][Glu] (Fig. 3.22a) and [Bmim][dca] (Fig. 3.22b).

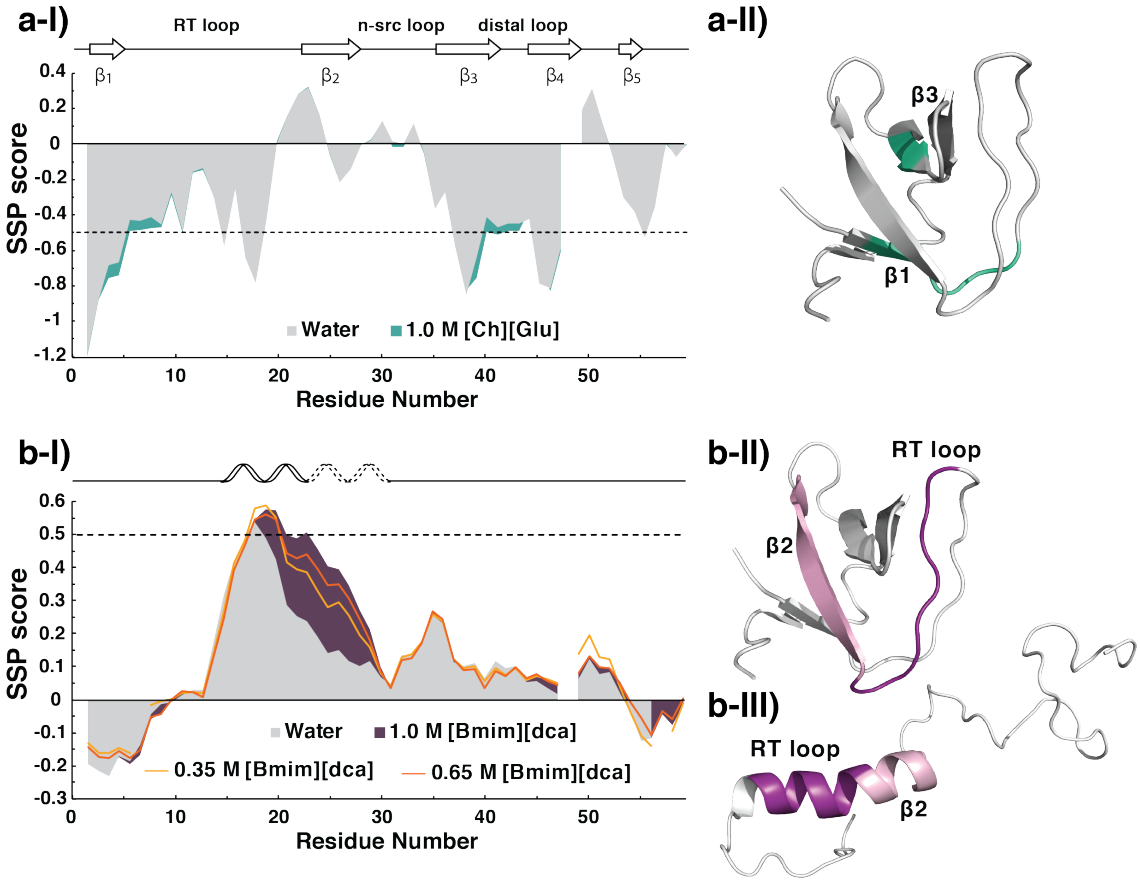



**Figure 3.22. Secondary structure propensity (SSP) of drkN SH3.**

Comparison of the SSP scores for SH3 for **a**) stabilising ([Ch][Glu]) and **c)** destabilising ([Bmim][dca]) conditions. **b)** 3D structure of SH3 in the folded (PDB: 2A36[6]) highlighting the affected regions (green). **d)** 3D structures of SH3 in the folded (left) and unfolded (right) (representative of the unfolded ensemble PED-8AAC, #1195[3]) states highlighting the affected regions of β2 (pink) and RT loop (purple). The SSP scores were calculated on the basis of Cα and Cβ chemical shifts, and a score at a given residue of 1 or -1 reflects fully formed $\alpha$- or β-structure, respectively, while a score of 0.5 or -0.5 (horizontal dashed lines) indicates that 50% of the conformers in the state ensemble are in $\alpha$- or β-structure, respectively, at that position[53]. The data was acquired at 298 K, pH=7.1 in water (grey) and in the presence of 1.0 M [Ch][Glu] (green) and in the presence of increasing amounts [Bmim][dca] (light orange line: 0.35 M; dark orange line: 0.65 M and dark purple; 1.0 M). Above each plot it is depicted the secondary structure of the protein according to the 3D structures. Note that the representation in **c)** includes the potential extension of the $\alpha$-helix (dashed lines).

For the $F_{[Ch][Glu]}$ state (Fig. 3.22a), only a small difference is observed the protein in water and [Ch][Glu], in agreement with the small-magnitude CSP observed (Fig. 3.6a). Still, while small, the observed difference goes in the direction of an increase in the β-structure propensity, arguing in favour of a small stabilisation of this structure. Nevertheless, the similarity between the SSP scores in water and [Ch][Glu] indicates that the folded structures should be essentially identical. This observation is in line with that of Zhang and Forman-Kay[8] for the $F_{[Na2SO4]}$ state, which, on the basis of similarities in NOE patterns and chemical shifts, showed that its structure should not differ from the $F_{water}$ state. Thus, despite the observed small CSPs in the presence of [Ch][Glu] and [$Na_2SO_4$], and the differences between the CSPs for these two cosolutes, I can conclude that the structures of the $F_{water}$, $F_{[Ch][Glu]}$ and $F_{[Na2SO4]}$ should be essentially the same.

For the $U_{[Bmim][dca]}$ state, a significant increase in $\alpha$-helical propensity for the residues between residues R20 and N29 is observed upon addition of [Bmim][dca]. Furthermore, the $\alpha$-helical propensity increases proportionally to the concentration of IL, despite the proteins being fully unfolded at 0.35 M. These observations are in full agreement with the CSP (Fig. 3.6b and Fig. 3.14), AGADIR predictions (Fig. 3.12b), and temperature dependence (Fig. 3.20), and suggest, in fact, [Bmim][dca] not only stabilises the already existent $\alpha$-helix, but also extends it further. Sequential amide-amide NOEs experiments should be performed to validate this $\alpha$-helical segment. Yet, Fig. 3.22b-III shows the cartoon representation of the one representative structure of the unfolded ensemble (structure 1195 from PED-8AAC[3]) where that extended helix is clearly visible. Using this structure, I can see that upon addition of 1.0 M of [Bmim][dca] IL the whole β2 (pink segment in the folded structure) is converted to an $\alpha$-helix structure. This stabilisation of a non-native structure should dramatically slow down the folding process, thus driving the observed shift towards the unfolded state of SH3.

In summary, the stabilisation of the folded state of SH3 caused by [Ch][Glu] seem not to be a consequence of conformational changes in its structure. Inversely, for the observed destabilisation of SH3 in [Bmim][dca] the data indicates that it is not mainly driven by a destabilisation of the folded state, but by a stabilisation of a particular conformation of the unfolded state.

**Thermodynamic fingerprint: entropy-enthalpy analysis**

As previously noted in the introductory section of kinetics of protein folding, the folded and unfolded states are almost balanced in energy, which undergo a reversible two-state process characterised by the equilibrium constant,

$$F \rightleftarrows U,$$

$$K_{eq} = \frac{p_u}{p_f} = exp\left(\frac{-\Delta G_u^{0\prime}}{RT}\right), \qquad [3.5]$$

where $p_f$ and $p_u$ are the folded and unfolded populations, respectively, R is the gas constant and $\Delta G_u^{0\prime}$ the Gibbs free energy of unfolding (F → U) in the standard state (index "0"). A positive value of $\Delta G_u^{0\prime}$ values indicates an excess of folded over unfolded protein and it is

founded in the mutual compensation of large enthalpic and entropic contributions, $\Delta H_u^{0'}$ and $\Delta S_u^{0'}$, respectively, known as enthalpy-entropy compensation[54,55],

$$\Delta G_u^{0'} = \Delta H_u^{0'} - T\Delta S_u^{0'}, \quad [3.6]$$

where increasing $\Delta H_u^{0'}$ and/or decreasing $\Delta S_u^{0'}$ increases protein stability by increasing $\Delta G_u^{0'}$. Instead of focusing on a single state, where $\Delta G_u^{0'} = 0$ (at melting temperature, $T_m$, where $p_f = p_u = 0.5$), the understanding of ion-specific effects is largely increased by determining the unfolding free energy change ($\Delta G_u^{0'}$) over a wide range of temperatures[56], thus obtaining the so-called "protein stability curve'', $\Delta G_u^{0'}(T)$. This temperature-dependent measurement allows assessment of enthalpic ($\Delta H_u^{0'}(T)$) and entropic ($T\Delta S_u^{0'}(T)$) components[55].

In general, for a globular protein, the curve dependence of $\Delta G_u^{0'}(T)$ resembles a parabola with a maximum at $T_s$ (temperature of maximum stability) and it intersects the zero line at high and low temperatures, defining the melting temperature ($T_m$) and the cold denaturation temperature ($T_c$). At the maximum, $\Delta S_u^{0'}$ is zero and all the stability arises from $\Delta H_u^{0'}$. At high temperatures, $\Delta G_u^{0'}$ is dominated by $\Delta S_u^{0'}$, while at lower temperatures, $\Delta H_u^{0'}$ is more influential[57]. This temperature dependence (curvature) is a generic effect of the heat-capacity increase upon unfolding ($\Delta C_{p,u}^{0'}$), that is assumed to be constant over the temperature range studied since calorimetric data have established that over a reasonably small temperature range (here from 278 K to 313 K), the change in $C_p$ is rather small[58], and that considering $\Delta C_{p,u}^{0'}(T_m) = \Delta C_{p,u}^{0'}(T)$ leads to meaningful results for $\Delta G_u^{0'}(T)$. This assumption was recently used, through $^{19}$F resonances, for drkN SH3 domain, in different reports from Pielak lab[20,59,60].

Expressions for the temperature dependence of $\Delta H_u^{0'}(T)$ and $\Delta S_u^{0'}(T)$ are obtained by using Kirchhoff's relations[61],

$$\Delta H_u^{0'}(T) = \Delta H_u^{0'}(T_{ref}) + \int_{T_{ref}}^{T} \Delta C_{p,u}^{0'} dT, \quad [3.7]$$

$$\Delta S_u^{0'}(T) = \Delta S_u^{0'}(T_{ref}) + \int_{T_{ref}}^{T} \frac{\Delta C_{p,u}^{0'}}{T} dT, \quad [3.8]$$

where $T_{ref}$ is a reference temperature, $\Delta H_u^{0'}(T_{ref})$ and $\Delta S_u^{0'}(T_{ref})$ are the enthalpy and entropy of unfolding at $T_{ref}$.

Inserting temperature-dependent relations (Eqs. 3.7 and 3.8) into Eq. 3.6 and integrating, gives the integrated Gibbs-Helmholtz equation[55],

$$\Delta G_u^{0'}(T) = \Delta H_u^{0'}(T_{ref}) - T\Delta S_u^{0'}(T_{ref}) + \Delta C_{p,u}^{0'}\left[(T - T_{ref}) - T\ ln\left(\frac{T}{T_{ref}}\right)\right], \quad [3.9]$$

where $T_{ref}$ is either the melting temperature, $T_m$ (where $p_u = p_f$), $\Delta G_u^{0'}(T) = 0$ and $\Delta S_u^{0'}(T_m) = \Delta H_u^{0'}(T_m)/T_m$; or the temperature of maximum stability, $T_s$, where $\Delta S_u^{0'}(T) = 0$, which results in Eqs. 3.10 and 3.11, respectively[55],

$$\Delta G_u^{0'}(T) = \Delta H_u^{0'}(T_m)\left(1 - \frac{T}{T_m}\right) + \Delta C_{p,u}^{0'}\left[(T - T_m) - T\ ln\left(\frac{T}{T_m}\right)\right], \quad [3.10]$$

$$\Delta G_u^{0'}(T) = \Delta H_u^{0'}(T_s) + \Delta C_{p,u}^{0'}\left[(T - T_s) - T\ ln\left(\frac{T}{T_s}\right)\right]. \quad [3.11]$$

In the temperature range where the protein is stable, the stability curve is essentially determined by the values of $\Delta C_{p,u}^{0'}$ and $\Delta H_u^{0'}$. The $\Delta C_{p,u}^{0'}$ is mostly related to the hydrophobic effect and is due to an increase in the solvent exposed non-polar surface area (broadly arising from water molecules solvating hydrophobic residues in the unfolded state that are not exposed in the folded state)[58]. When the protein unfolds, $\Delta C_{p,u}^{0'}$ provides a useful measure of the increase in solvent-accessible surface area of the F to U transition[62]. Because the hydration grows "stronger" at lower temperatures, the larger surface area of U promotes cold unfolding and curved $\Delta G_u^{0'}(T)$ profiles[63]. Using the intensities of indole NH side chain of W36 resonances in both folded and unfolded states in a series of $^1$H,$^{15}$N-HSQC spectra acquired in a wide range of temperatures (from 278 to 313 K in 5 K increments), that allow quantitative determination of $p_f$ and $p_u$ (as previously discussed using Eq. 3.1), I measured the temperature dependence of $\Delta G_u^{0'}$ (using Eq. 3.3 or 3.5) in water, buffer 0.05 M sodium phosphate (pH 7.2), 0.10 and 0.35 M [Ch][Glu], and 0.15 M [Bmim][dca] (Table 3.2 and Fig. 3.23).

**Table 3.2. $\Delta G_u^{0'}$ (T) for drkN SH3 in water and ILs.**

| | $\Delta G_u^{0'}$ (kcal/mol) | | | | |
|---|---|---|---|---|---|
| **Temperature (K)** | **Water** | **Buffer** | **0.10 M [Ch][Glu]** | **0.35 M [Ch][Glu]** | **0.15 M [Bmim][dca]** |
| **278.1** | -0.16 ± 0.01 | 0.11 ± 0.01 | 0.16 ± 0.01 | 0.70 ± 0.04 | -0.95 ± 0.06 |
| **283.1** | -0.030 ± 0.002 | 0.22 ± 0.01 | 0.26 ± 0.02 | 0.77 ± 0.05 | -0.94 ± 0.06 |
| **288.1** | 0.035 ± 0.002 | 0.26 ± 0.02 | 0.31 ± 0.02 | 0.82 ± 0.05 | -0.98 ± 0.06 |
| **293.2** | -0.0012 ± 0.0001 | 0.23 ± 0.01 | 0.30 ± 0.02 | 0.83 ± 0.05 | -1.04 ± 0.06 |
| **298.2** | -0.12 ± 0.01 | 0.12 ± 0.01 | 0.21 ± 0.01 | 0.75 ± 0.05 | -1.26 ± 0.08 |
| **303.2** | -0.30 ± 0.02 | -0.05 ± 0.003 | 0.028 ± 0.002 | 0.65 ± 0.04 | -1.5 ± 0.1 |
| **308.1** | -0.61 ± 0.04 | -0.34 ± 0.02 | -0.27 ± 0.02 | 0.41 ± 0.03 | -2.0 ± 0.1 |
| **313.1** | -1.12 ± 0.07 | -0.81 ± 0.05 | -0.72 ± 0.04 | 0.047 ± 0.003 | -2.7 ± 0.2 |

$\Delta G_u^{0'}(T)$ of drkN SH3 were calculated with Eq. 3.5 as described in main text. The values are based on peak volume (simultaneously for folded and unfolded states) of indole side chain of W36 drkN SH3 in water, buffer 0.05 M sodium phosphate pH 7.2, 0.10 and 0.35 M [Ch][Glu], and 0.15 M [Bmim][dca], from 278 K to 313 K. Error represent the propagation of uncertainty in peak volume estimated by PINT[64]. Details can be found in Materials and Methods section.

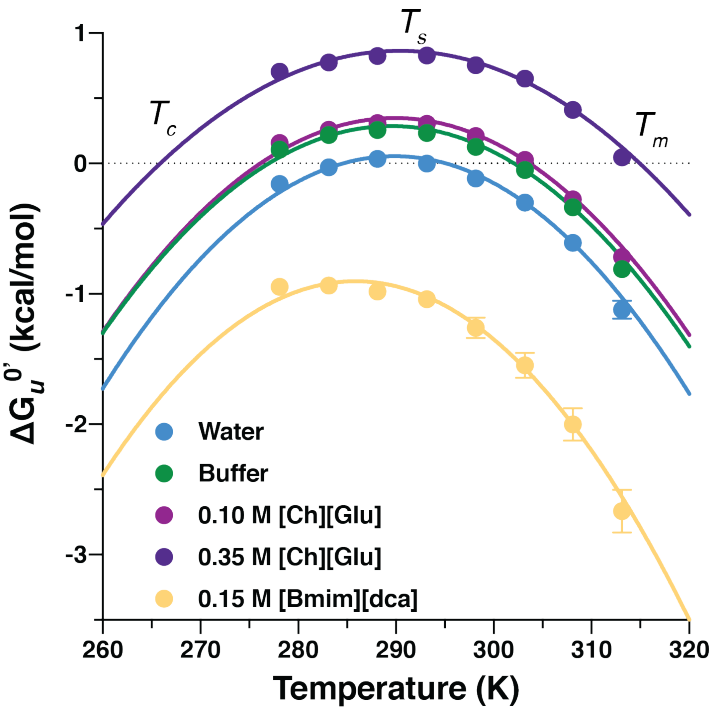



**Figure 3.23. Stability curves for SH3 in the presence of [Ch][Glu] and [Bmim][dca] ILs.**

Experimental data points for drkN SH3 in water, buffer 0.05 M sodium phosphate pH 7.2, 0.10 and 0.35 M [Ch][Glu], and 0.15 M [Bmim][dca], measured from 278 K to 313 K in 5 K increment. Error bars for almost all conditions are smaller than the labels and represent the propagation of uncertainty in peak volume estimated by PINT[64]. The lines correspond to the fitting curves as calculated from Eq. 3.10 or 3.11. The curve crosses the abscissa at the temperature of cold denaturation, $T_c$ and the melting temperature, $T_m$, and $T_s$ is the temperature of maximum stability. Details can be found in Materials and Methods section.

By fitting Eq. 3.10 or Eq. 3.11[55] (when $T_m$ cannot be defined) to the data (Table 3.2), I obtain $T_m$, $\Delta H_u^{0'}(T_m)$ and/ or $T_s$, $\Delta H_u^{0'}(T_s)$, and $\Delta C_{p,u}^{0'}$ (Table 3.3). Then, $\Delta H_u^{0'}$ ($T_{m,water}$) and $T\Delta S_u^{0'}(T_{m,water})$ at the $T_m$ of SH3 in water ($T_{m,water}$ = 294 K and $\Delta G_{u,water}^{0'} = 0$) were extrapolated by Kirchhoff's temperature-dependent relations (Eq. 3.7 and 3.8)[61]. The results indicate that the buffer (0.05 M sodium phosphate) slightly stabilises SH3 compared to water, which are consistent with reported values[8,20]. The results in 0.1 M [Ch][Glu] are similar to those in buffer but indicate the same tendency of stabilisation increase as 0.35 M [Ch][Glu] data.

**Table 3.3. Thermodynamic parameters for SH3 in the presence of ILs.**

| | $T_m$, K | $\Delta H_u^{0'}$ ($T_m$), kcal/mol | $T_s$, K | $\Delta H_u^{0'}$ ($T_s$), kcal/mol | $\Delta C_{p,u}^{0'}$, kcal/mol/K | $\Delta G_u^{0'}$ ($T_{m,water}$), kcal/mol | $\Delta H_u^{0'}$ ($T_{m,water}$), kcal/mol | $T\Delta S_u^{0'}$ ($T_{m,water}$), kcal/mol |
|---|---|---|---|---|---|---|---|---|
| **Water** | **294 ± 1** | 6 ± 1 | **289.0 ± 0.5** | 0.06 ± 0.02 | 1.16 ± 0.07 | **0** | **6.2 ± 0.5** | **6.2 ± 0.5** |
| **Buffer** (0.05 M Sodium Phosphate) | **301.4 ± 0.4** | 13.6 ± 0.8 | **288.7 ± 0.6** | 0.29 ± 0.02 | 1.05 ± 0.07 | **0.23 ± 0.01** | **6.2 ± 0.2** | **6.0 ± 0.2** |
| **0.10 M [Ch][Glu]** | **303.0 ± 0.6** | 15.0 ± 0.8 | **289.1 ± 0.5** | 0.35 ± 0.02 | 1.06 ± 0.07 | **0.30 ± 0.02** | **5.9 ± 0.1** | **5.6 ± 0.1** |
| **0.35 M [Ch][Glu]** | **314.5 ± 0.6** | 21 ± 1 | **289.6 ± 0.7** | 0.86 ± 0.02 | 0.83 ± 0.07 | **0.83 ± 0.05** | **4.8 ± 0.1** | **4.0 ± 0.1** |
| **0.15 M [Bmim][dca]** | NA | NA | **285 ± 1** | -0.89 ± 0.03 | 1.3 ± 0.1 | **-1.08 ± 0.06** | **11 ± 2** | **12 ± 2** |

$\Delta G_u^{0'}$ ($T$) were calculated with Eq. 3.5. $T_m$, $T_s$, $\Delta H_u^{0'}$ ($T_m$), $\Delta H_u^{0'}$ ($T_s$), and $\Delta C_{p,u}^{0'}$ were obtained by fitting of Eq. 3.10 or 3.11 when $T_m$ cannot be defined. Uncertainties represent the standard deviation from the fitting. The extrapolation of $\Delta H_u^{0'}$ ($T_{m,water}$) and $\Delta S_u^{0'}$ ($T_{m,water}$) and their uncertainties to $T_{m,water} \approx 294$ K ($T_m$ in cosolute-free solution), using $\Delta H_u^{0'}$($T_m$) or $\Delta H_u^{0'}$($T_s$) when $T_m$ cannot be defined, were obtained by Kirchhoff's temperature dependent relations[61]. See Materials and Methods for details. NA, not applicable.

The drkN SH3 is stabilised with [Ch][Glu] (Fig. 3.19) as reflected by the higher $T_m$ and overall higher $\Delta G_u^{0'}$ when compared with water ($\Delta G_u^{0'}$ (at $T_{m,water}$) = 0.83 ± 0.05 kcal/mol). This increase in $\Delta G_u^{0'}$ is the result of a large increase in $\Delta H_u^{0'}$ and a decrease in $\Delta C_{p,u}^{0'}$. A significant change in $\Delta C_{p,u}^{0'}$ was observed for 0.35 M [Ch][Glu] ($\Delta\Delta C_{p,u}^{0'} = \Delta C_{p,u,0.35MChGlu}^{0'} - \Delta C_{p,u,water}^{0'}$ = -0.3 ± 0.1 kcal/mol/K). This decrease which corresponds to a shallower curve (broadening the width of the thermal unfolding transitions) is attributed to a decreased solvent-accessible surface area of the unfolded ensemble[62] compared to water conditions. Similar mechanism was found for glucose-induced stabilisation of drkN SH3[20]. The effect could be assigned to stabilising excluded-volume contributions, in opposition to in-cell studies that show destabilising weak attractive interactions often dominate stabilising hard-core excluded volume effect[20,65,66]. [Bmim][dca] destabilises SH3 relative to water ($\Delta G_u^{0'}$ (at $T_{m,water}$) = -1.08 ± 0.06 kcal/mol) with a notorious increase of enthalpic and entropic components. The heat capacity of unfolding ($\Delta C_{p,u}^{0'}$) in 0.15 M [Bmim][dca] is very similar to its value in dilute solution, in agreement with previous findings for other proteins and cosolutes[56,67].

The co-solute induced changes to the protein unfolding free energy (F → U), relative to the cosolute-free (water), are then quantified by the excess functions,

$$\Delta\Delta G_u^{0'} = \Delta G_{u,cosolute}^{0'} - \Delta G_{u,water}^{0'} = \Delta\Delta H_u^{0'} - T\Delta\Delta S_u^{0'}, \quad [3.12]$$

where, $\Delta\Delta G_u^{0'}$ is the excess Gibbs free energy and $\Delta\Delta H_u^{0'}$ and $T\Delta\Delta S_u^{0'}$ are the enthalpic and entropic contributions, respectively. These changes, here calculated at $T_m$ in water ($T_{m,water}$), are indicated in Table 3.4 and provide a thermodynamic fingerprint useful to elucidate

different mechanisms of co-solute action, allowing to validate molecular level theories of ion-specific behaviour[56,68,69].

**Table 3.4. Change in thermodynamic parameters for SH3 in the presence of ILs.**

| | $\Delta T_m$, K | $\Delta T_s$, K | $\Delta\Delta C_{p,u}^{0'}$, kcal/mol | $\Delta\Delta G_u^{0'}$ ($T_{m,water}$), kcal/mol | $\Delta\Delta H_u^{0'}$ ($T_{m,water}$), kcal/mol | $T\Delta\Delta S_u^{0'}$ ($T_{m,water}$), kcal/mol |
|---|---|---|---|---|---|---|
| **Buffer (0.05 M Sodium Phosphate)** | 7 ± 1 | -0.3 ± 0.7 | -0.1 ± 0.1 | **0.23 ± 0.01** | **0.0 ± 0.5** | **-0.2 ± 0.5** |
| **0.10 M [Ch][Glu]** | 9 ± 1 | 0.1 ± 0.7 | -0.1 ± 0.1 | **0.30 ± 0.02** | **-0.3 ± 0.5** | **-0.6 ± 0.5** |
| **0.35 M [Ch][Glu]** | 20 ± 1 | 0.6 ± 0.9 | -0.3 ± 0.1 | **0.83 ± 0.05** | **-1.4 ± 0.5** | **-2.2 ± 0.5** |
| **0.15 M [Bmim][dca]** | NA | -4 ± 1 | 0.1 ± 0.1 | **-1.04 ± 0.06** | **5 ± 2** | **6 ± 2** |

Excess changes using as reference water for the thermodynamic parameters of drkN SH3. Uncertainties represent the error propagation from the initial standard fitting error. NA, not applicable.

The $\Delta\Delta H_u^{0'}$ and $T\Delta\Delta S_u^{0'}$ contributions of [Ch][Glu] and [Bmim][dca] ILs on SH3 unfolding can then be conveniently mapped out in an entropy-enthalpy compensation plot (Fig. 3.24). As discussed before in the introductory chapter of IL-protein interactions, this stability diagram characterises the mutual compensation of large enthalpic and entropic contributions which tend to cancel in $\Delta G_u^{0'}$ (Eq. 3.6), and their possible relative contributions (Eq. 3.12) delineate eight regions (marked by different colours in Fig. 3.24). The full enthalpic-entropic compensation, represented by the diagonal line in Fig. 3.24, separates destabilising cosolutes that lie above it (orange regions: V to I, where $\Delta\Delta G_u^{0'} < 0$) from those that cause stabilisation lying beneath it (blue/purple regions: II to V, where $\Delta\Delta G_u^{0'} > 0$). Regions II and III include cosolutes showing an enthalpically driven stabilisation ($\Delta\Delta G_u^{0'} > 0$ and $\Delta\Delta H_u^{0'} > 0$), while cosolutes that fall within regions IV and V induce increased stability that is primarily entropically driven ($\Delta\Delta G_u^{0'} > 0$ and $T\Delta\Delta S_u^{0'} < 0$). Similarly, regions VI and VII indicate an enthalpic destabilisation mechanism ($\Delta\Delta G_u^{0'} < 0$ and $\Delta\Delta H_u^{0'} < 0$), while regions VIII and I include cosolutes that are entropic destabilisers ($\Delta\Delta G_u^{0'} < 0$ and $T\Delta\Delta S_u^{0'} > 0$). Note that for cosolutes that populate near the borderline ($\Delta\Delta H_u^{0'} \approx T\Delta\Delta S_u^{0'}$) between sections I and II, or sections V and VI, where both enthalpic and entropic contributions are positive, or negative, it depends on the magnitudes of $\Delta\Delta H_u^{0'}$ and $T\Delta\Delta S_u^{0'}$ to obtain stabilising ($\Delta\Delta G_u^{0'} > 0$) or destabilising cosolutes ($\Delta\Delta G_u^{0'} < 0$). For the negatively charged drkN SH3 domain I find that [Ch][Glu] (at 0.1 and 0.35 M) and [Bmim][dca] (at 0.15 M) are entropically driven and located in the expected fields, according to Senske et al[56].

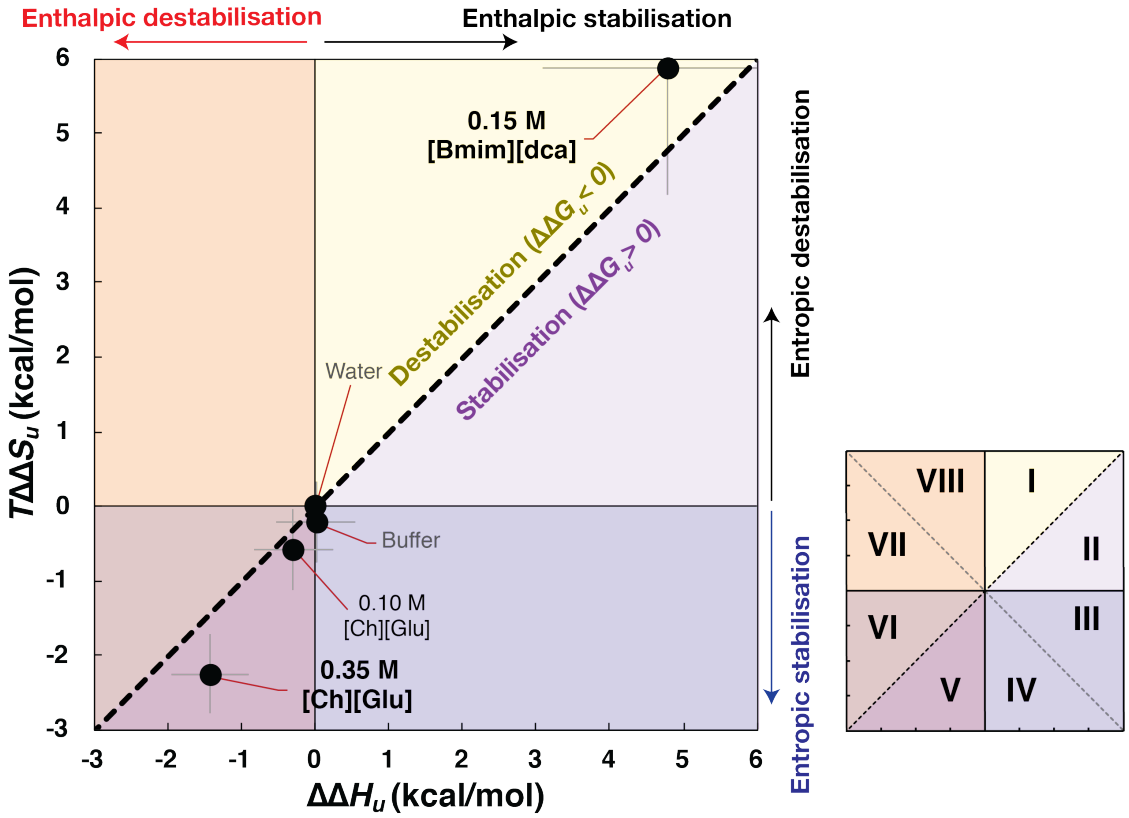



**Figure 3.24. Enthalpy-entropy compensation plot.**

The different segments correspond to different contributions of $\Delta\Delta G_u^{0'}$, $\Delta\Delta H_u^{0'}$, $T\Delta\Delta S_u^{0'}$, here calculated at $T_{m,\,water} \approx 294$ K. Plot adapted from Senske et al[56], the different signs of the excess functions $\Delta\Delta G_u^{0'}$, $\Delta\Delta H_u^{0'}$, $T\Delta\Delta S_u^{0'}$ define eight different fields. The blue diagonal corresponds to a complete enthalpy-entropy compensation and separates the protein destabilising region ($\Delta\Delta G_u^{0'} < 0$) from the stabilising region ($\Delta\Delta G_u^{0'} > 0$). Positive (negative) values of $\Delta\Delta H_u^{0'}$ imply stabilisation (destabilisation) by the cosolute. The entropy term acts in the opposite direction: a positive $T\Delta\Delta S_u^{0'}$ supports destabilisation.

[Ch][Glu], as most of stabiliser hydrophilic salts (e.g., NaCl, KCl, $K_2SO_4$, and sorbitol)[56], leads to protein stabilisation ($\Delta\Delta G_u^{0'} > 0$) *via* both negative enthalpic and entropic contributions, where the magnitude of $T\Delta\Delta S_u^{0'}$ (entropic stabilisation) dominates over $\Delta\Delta H_u^{0'}$ (enthalpic destabilisation). The decrease in $T\Delta\Delta S_u^{0'}$ is in line with stabilising hard-core repulsions[70,71], in which, the presence of a cosolute with a large excluded-volume lead to a loss of configurational entropy of a protein as the thermodynamic driving force. Because the entropic loss is larger for the protein's unfolded state compared to the more compact folded state, the protein is stabilised. Nevertheless, the presence of unspecific attractive interactions with the protein lead to negative $\Delta\Delta H_u^{0'}$ [67,71] that counteract the decrease in entropy. This is in agreement with the observations found for SH3 unfolding in the presence of sugars (glucose, sucrose, dextran)[20]. In addition, the observations that the trend in $\Delta\Delta G_u^{0'}$ (for 0.1 and 0.35 M, the values are 0.30 ± 0.02 and 0.83 ± 0.05 kcal/mol, respectively) is parallel to that of $\Delta T_m$ during unfolding of SH3 (for 0.1 and 0.35 M, the values are 9 ± 1 and 20 ± 1 K, respectively), and the small shift in $T_s$ to higher temperatures compared to water (Table 3.4),

support the entropic stabilisation pathway by [Ch][Glu], even with the destabiliser enthalpic contribution.

As [Ch][Glu] molecules seem to be strongly partitioned to the protein surface (based on chemical shift perturbation data, Fig. 3.6a), an alternative explanation for SH3 stabilisation could be due to a preferential exclusion of its ions and/or ion-pair from the protein surface (preferential hydration)[72,73]. Because the transfer free energy ($\Delta G_{tr}$) of the peptide backbone from water to a stabilising osmolyte solution is unfavourable[74,75]. The stabilisation by preferential hydration arises because the unfolded state exposes more peptide bonds than the native state. This results in an increase of $\Delta H_u^{0'}$ (enthalpic stabilisation) that is partially offset by an increase in $\Delta S_u^{0'}$ [76,77]. However, this mechanism is not consistent with our thermodynamic data (Table 3.4 and Fig. 3.24). Nevertheless, salts that are similar to [Ch][Glu], K[dhp] and [Ch][dhp], revealed a sharp transition from predominant enthalpic to entropic stabilisation near 0.25 or 0.5 M, respectively[56], which is associated to a maximum of $\Delta\Delta H_u^{0'}$ and $T\Delta\Delta S_u^{0'}$. This is not observed in SH3 unfolding due to the limited IL concentration data, but it is certain related to the non-monotonous behaviour of GB1's $T_m$ along [Ch][Glu] concentration (as discussed in chapter 2).

In the segment where both enthalpic and entropic components are negative, if the enthalpic component is raised compared to the entropic one, the cosolute shifts from segment V to VI (where $\Delta\Delta G_u^{0'} < 0$ and $\Delta\Delta H_u^{0'} < 0$, Fig. 3.24). Protein denaturants, such as urea and [Gdm]Cl, are commonly found in this segment where a strong enthalpic denaturing effect is observed[56,67,68], probably due to appreciable binding to proteins that reduce $\Delta H_u^{0'}$. In contrast, destabiliser ionic liquids are mainly located in the opposite field (positive $\Delta\Delta H_u^{0'}$ and $T\Delta\Delta S_u^{0'}$)[56] indicating a different molecular mechanism compared to typical protein denaturant. [Bmim][dca], as other hydrophobic imidazolium-based ILs (e.g., [Bmim]Br, [Bmpyrr]Br, [Hmim]Br)[56,78] leads to a destabilisation of SH3 *via* a positive enthalpic contribution but an even more positive entropic contribution, leading to entropic destabilisation offset by enthalpic stabilisation. The lower $T_s$ ($\Delta T_s$ = -4 ± 1 K) found for [Bmim][dca] also supports the entropic destabilisation effect. The large positive value of $\Delta\Delta H_u^{0'}$ found for [Bmim][dca] indicate a molecular mechanism of unfavourable interactions with the protein backbone, possibly through water-mediated hydrophobic contacts[79], originated by the hydrophobicity of the cation and/ or ion-pair. Since the entropy of water is already reduced due to the presence of [Bmim][dca], according to Senske et al[56], it is a reasonable explanation that the experimentally observed $\Delta\Delta S_u^{0'} > 0$ in the presence of solvated $[Bmim]^+$ and $[dca]^-$ ions is due to a diminished loss of water entropy with the solvation of the hydrophobic groups of drkN SH3 which get solvent exposed upon unfolding. This is in line with the observation that the more hydrophobic the cosolute, the stronger the reduction of $T_m$ and $T_s$, and the larger the destabilising entropic and the counteracting stabilising enthalpic contribution. The hydrophobicity of the butyl chain of $[Bmim]^+$ cation is further increased by $[dca]^-$ anion which contributes for the denaturing power of [Bmim][dca] ion-pair. A direct mechanism (as discussed above for urea and [Gdm]Cl), which could also explain the reduction of $T_m$, it is not likely because the direct interactions between protein and denaturant should lead to a negative $\Delta\Delta H_u^{0'}$. Instead, an indirect mechanism is supported by the stabilisation of the unfolded state through increase of helix propensity of the non-native hydrophobic segment of unfolded ensemble.

**ZZ-exchange: folding kinetics study**

Although the analysis of thermodynamic data suggest opposite mechanisms of [Ch][Glu] and [Bmim][dca] ILs on SH3 unfolding (entropic stabilisation offset by enthalpic destabilisation *versus* entropic destabilisation offset by enthalpic stabilisation, respectively), it is not possible to conclude if and how which protein's state (folded or unfolded) is affected. Taking advantage from the slow exchange process between folded and unfolded states in the drkN SH3 domain (exchange rate constant $k_{ex} \approx 2.2\ s^{-1}$, 0.05 M sodium phosphate pH 6.0, 293 K)[5] a longitudinal nitrogen magnetization exchange experiment ($N_z$)[80], referred as magnetization exchange, or ZZex, can be used to obtain both thermodynamic (folded/unfolded populations) and kinetic (exchange rate constants) information as well as structural information in the form of chemical shift between the two states, $\delta\omega_{UF} = \omega_U - \omega_F$, on a per-residue basis and without perturbation of equilibrium conditions. Here, I investigated the dependence of unfolding and folding interconversion rates ($k_{fu}$ and $k_{uf}$) of drkN SH3 in different solvent conditions to further characterise the effects of ILs on both $F_{exch}$ and $U_{exch}$ states.

As explored in the introductory chapter (methods for ms-s timescale dynamics), briefly, in the case of two-state exchanging ($F_{exch}/U_{exch}$) sample,

$$F \underset{k_{uf}}{\overset{k_{fu}}{\rightleftarrows}} U, \qquad [3.13]$$

with a slow timescale ($0.1\ s^{-1} < k_{ex} < 10\ s^{-1}$, where $k_{ex} = k_{fu} + k_{uf}$), a pair of auto-peaks (*ff* and *uu*) is observed for each residue as well as a pair of exchange cross-peaks (*fu* and *uf*), which result from the transfer of magnetization between the $F_{exch}$ and $U_{exch}$ states during the mixing time (due to a delay included between the $^{15}N$-$^{1}H$ frequency-recording periods).

The equations, for a two-state exchange model, describing the intensity decay due to longitudinal relaxation for auto-peaks (*ff*, *uu*), and the emergence due to chemical exchange and longitudinal relaxation of cross-peaks (*fu*, *uf*) peaks, on the variable mixing period, *t*, have been described before[5,80–82] and are given in Materials and Methods. The time evolution of the resulting four peaks is simultaneously fitted using the longitudinal Bloch-McConnell equations[83] for the measurement of chemical exchange ($k_{ex}$) along with longitudinal $^{15}N$ decay ($R_{1f}$ and $R_{1u}$) rates that allow the extraction of $k_{fu}$ and $k_{uf}$ interconversion rates.

Because the folded and unfolded populations are approximately equal for SH3 in water at 293 K ($p_f \approx p_u \approx 0.5$), which is approximately $T_{m,water}$, the concentrations of 0.35 M [ChGlu] and 0.15 M [Bmim][dca] were chosen to be investigate, since at 293 K, SH3 yield a F/U ratio of ≈ 4 ($p_f \approx 0.8$, $p_u \approx 0.2$) in the first and its inverse in the latter ($p_f \approx 0.2$, $p_u \approx 0.8$). See ZZex spectra for these conditions in Fig. 3.25 and Fig. 3.26.

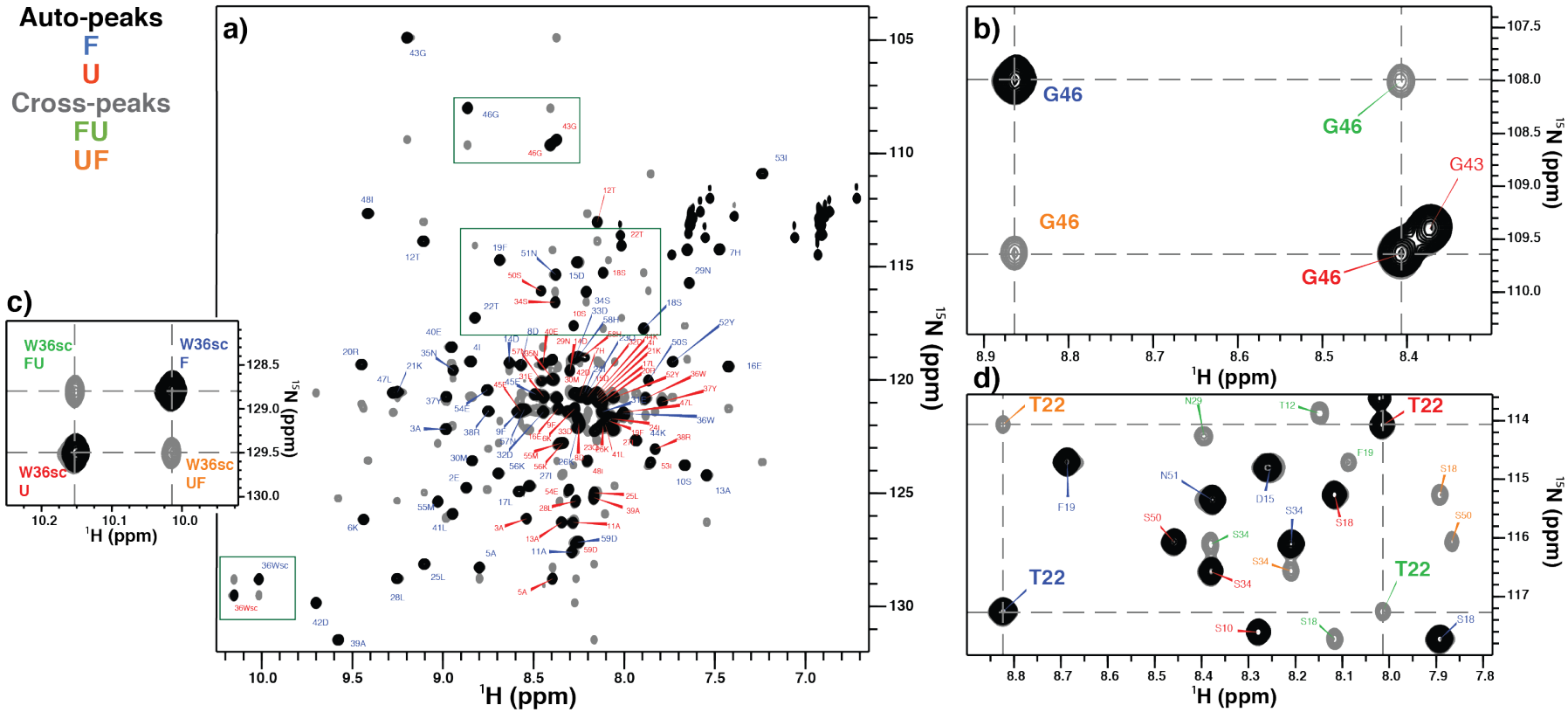



**Figure 3.25. $^{1}$H-$^{15}$N ZZ-exchange spectrum of drkN SH3 in water.**

The spectrum was acquired with 320 ms of mixing time at 600.13 MHz, 293.2 K for 1.1 mM of $^{15}$N labelled drkN SH3 in water. **a)** full spectrum. The different inlays show **b)** Gly46, **c)** side chain of Trp36, and **d)** Thr22 NH auto- and cross-peaks. Peaks for the folded and unfolded forms of the residue (black) are labelled with F (blue) and U (red), respectively, and dotted lines indicate the exchange cross-peaks (gray), FU (green) and UF (orange).

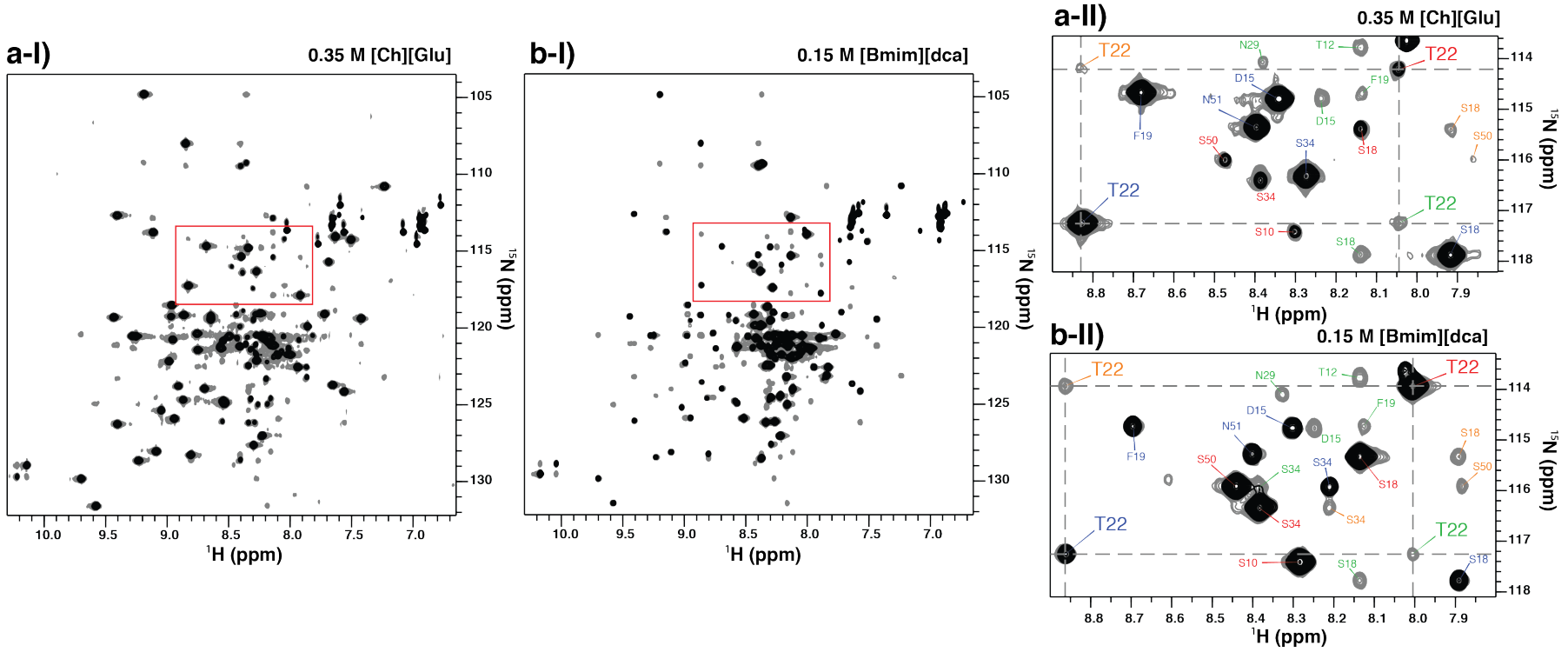



**Figure 3.26. $^{1}$H-$^{15}$N ZZ-exchange spectra of drkN SH3 in 0.35 M [Ch][Glu] and 0.15 M [Bmim][dca].**

The spectra (**I**) were acquired with 320 ms of mixing time at 600.13 MHz, 293.2 K for 1.1 mM of $^{15}$N labelled drkN SH3 in **a)** 0.35 M [Ch][Glu] and **b)** 0.15 M [Bmim][dca] IL. The inlay (**II**) in right shows T22 NH auto- and cross-peaks. Peaks for the folded and unfolded forms of the residue (black) are labelled with F (blue) and U (red), respectively, and dotted lines indicate the exchange cross-peaks (gray), FU (green) and UF (orange).

While there is an almost-perfect correlation between the $\delta\omega_{UF}$ in water and those in [Ch][Glu] and [Bmim][dca] (Fig. 3.27), some residues still deviate, with this behaviour more evident in [Bmim][dca] (I27 and R38 for [Ch][Glu] and I24 to L28 for [Bmim][dca]). These residues are also the ones previously shown to have their chemical shifts affected (Fig. 3.6) and validate a conformational change.

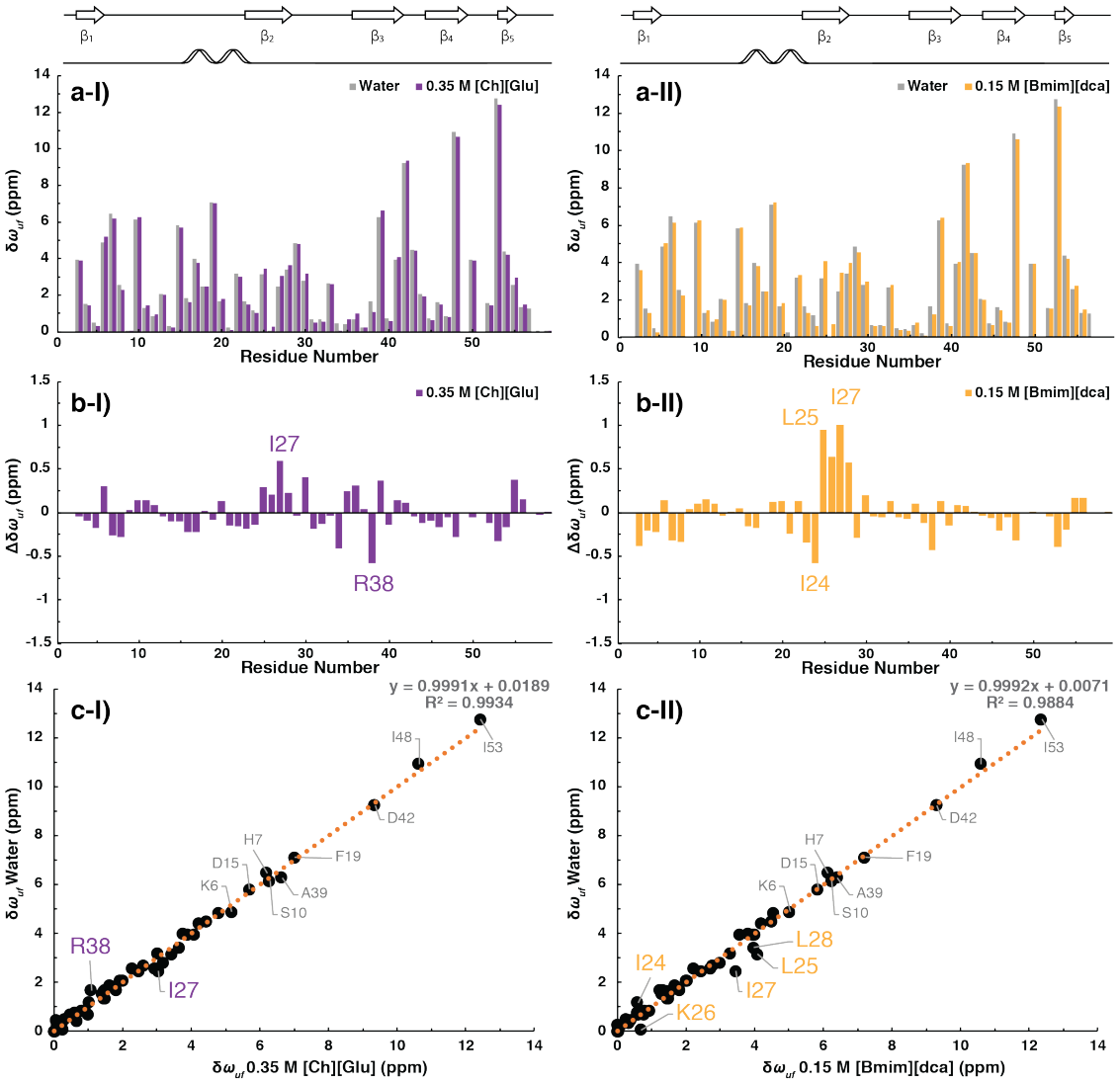



**Figure 3.27. Correlation of $\delta\omega_{UF}$ for water, [Ch][Glu] and [Bmim][dca] ILs.**

**a**) $^{15}N$ chemical shift difference between $F_{exch}$ and $U_{exch}$, $\delta\omega_{UF}$, extracted directly from ZZex-spectra for water (gray) and **I)** 0.35 M [Ch][Glu] or **II)** [Bmim][dca] (black). **b**) The difference, $\Delta\delta\omega_{UF} = \delta\omega_{UF,\ cosolute} - \delta\omega_{UF,\ water}$, as calculated for **I)** [Ch][Glu] and **II)** [Bmim][dca]. **c**) Correlation between $^{15}N$ chemical shift difference between $F_{exch}$ and $U_{exch}$, $\delta\omega_{UF}$, obtained directly from ZZex-spectra for water and **I**) 0.35 M [Ch][Glu] or **II**) [Bmim][dca]. The gray label indicates $\delta\omega_{UF}$ residues > 5 ppm and purple/orange label indicates the deviations ($|\Delta\delta\omega_{UF} = \delta\omega_{UF}(IL) - \delta\omega_{UF}\ (water)| > 0.5$ ppm).

The extracted parameters ($p_f$, $p_u$, $R_{1f}$, $R_{1u}$, $k_{fu}$ and $k_{uf}$) from the ZZex data fitting are compiled in Table B3 of Appendix B. As representative example, the time profile of the intensities of the auto and exchange peaks for residue T22 are shown for different conditions in Figure 3.28. In total, I was able to measure both interconversion rates ($k_{fu}$ and $k_{uf}$) for 15 of the 59 residues of SH3 (S10, T12, A13, E16, S18, T22, L25, L28, N28, A39, D42, G46, I53, K56 and W36 indole side chain) at the different studied conditions (Table 3.B4 of Appendix B).

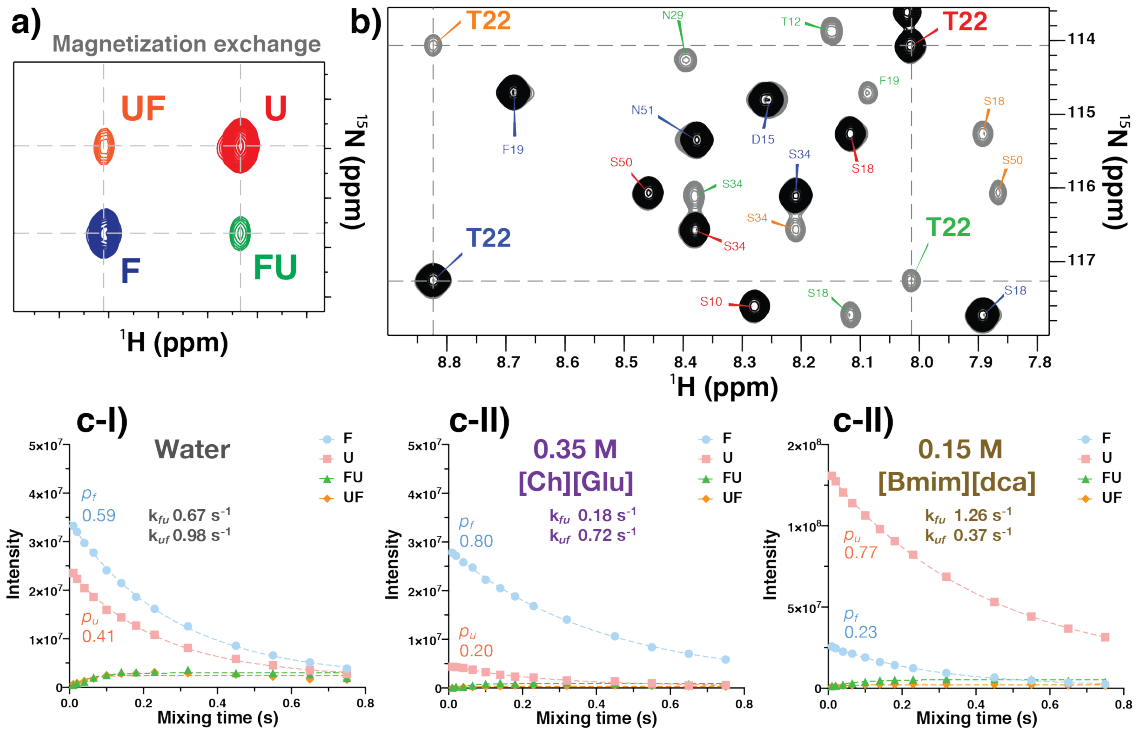



**Figure 3.28. Experimental data and fitted curves for Thr22 ZZex.**

**a**) ZZex typical spectra for a $^1$H/$^{15}$N spin pair undergoing slow two-site exchange with a mixing period *t* applied. Blue and red peaks correspond to the auto-peaks of states F and U, respectively, and green and orange peaks to the cross-peaks (FU and UF, respectively) which are produced by conformational exchange during the delay. **b**) Selected region of $^{15}$N-$^1$H ZZex-HSQC spectra containing peaks of T22 from the drkN SH3 in water, 293 K, 600.13 MHz. The spectra show T22 auto-peaks (black) and cross-peaks (grey) for the 320 ms mixing time. **c**) Time profiles of the intensities of the auto (*uu* and *ff*) and exchange (*uf* and *fu*) peaks for T22 in **I**) water, **II**) 0.35 M [Ch][Glu] and **III**) 0.15 M [Bmim][dca]. Dashed lines are the best-fit curves for magnetization originating in F or U state.

Protein folding necessarily involves diffusional events, as the expanded and hydrated unfolded state polypeptide chain collapses toward a more compact and conformationally restricted folded protein (water is expelled from the interior of the protein). Instead to use the original transition state formalism, Kramers' rate theory accounts that protein folding is a diffusional process and the energy barrier-crossing process is viscosity-dependent[84]. In this case, based on the Stokes law, it is expected that the (un)folding rate constants are inversely proportional to the solvent viscosity $(1/\eta)$ [85], whereas rate constants will be independent of solvent viscosity if the rate-limiting step involves only rearrangements that are

limited by the internal friction of the protein. Because viscosity should affect both unfolding and folding to the same extent, it does not change the equilibrium of the folding reaction (i.e., protein stability). Thus, to separate the viscosity effects from those that result from increased stability, the addition of 0.35 M [Ch][Glu] or 0.15 M [Bmim][dca] in the bulk viscosity of a solution need to be considered. Assuming that unfolding and folding rates of the drkN SH3 domain are indeed linearly-dependent on solvent viscosity[86], the experimental values of unfolding/folding rates ($k_{fu}$ and $k_{uf}$) can be adjusted,

$$k_{0\,(fu,uf)} = k_{fu,uf}\left(\frac{\eta_{cosolute}}{\eta_{water}}\right) = k_{fu,uf}\eta_{rel}, \qquad [3.14]$$

where $k_0$ is the viscosity-adjusted rate, $k$ is the folding/unfolding rate before viscosity correction, $\eta_{cosolute}$ is the viscosity of the solution in the presence of a given cosolute, $\eta_{water}$ is the viscosity of pure water at 293.2 K (1.002 cP)[87], and $\eta_{rel}$ is the viscosity adjusted to water at 293.2 K. Similar viscosity-correction approaches were used by other groups[22,88]. IL-solution viscosity measurements and extrapolations are described in Material and Methods and viscosity-adjusted rates are also listed in Table B4 of Appendix B.

The 15 residues accurately followed by ZZex are well distributed along the protein sequence (Fig. 3.29), and its (un)folding rates were adjusted accordingly to the relative viscosity. Table 3.5 compile the average values for the different extracted parameters from ZZex data analysis.

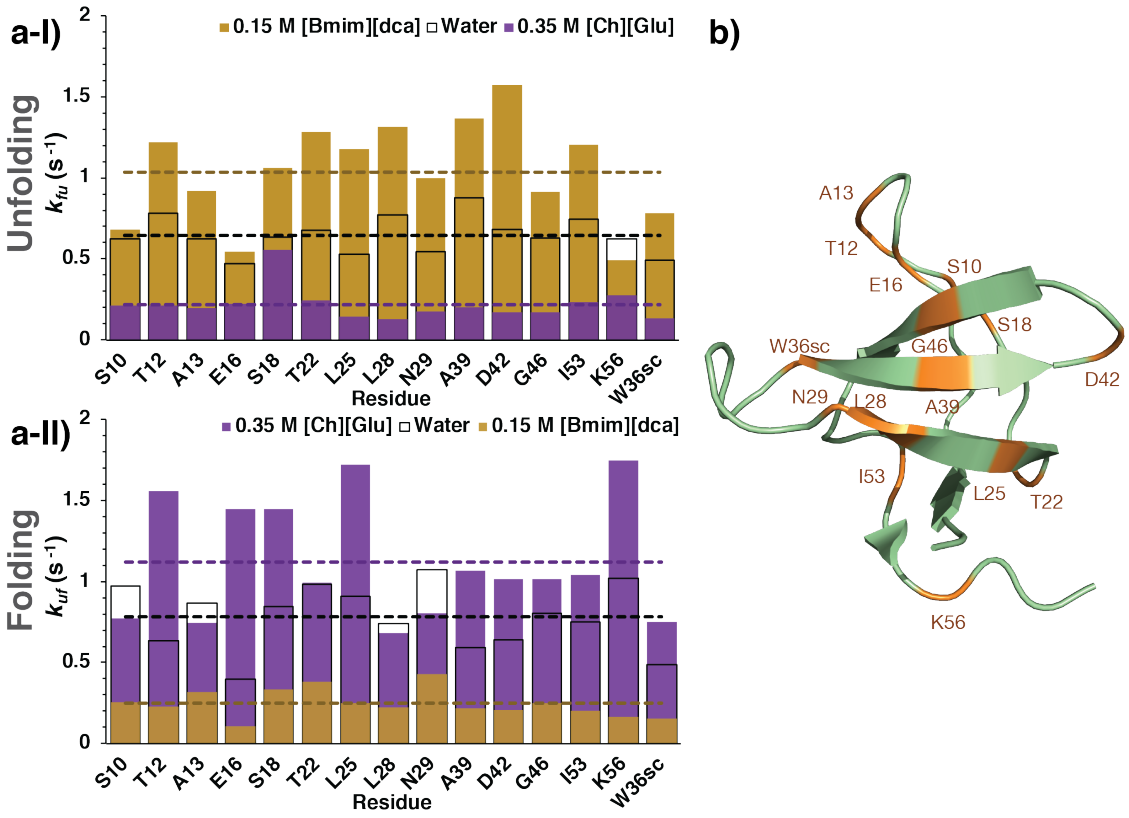



**Figure 3.29. Viscosity-adjusted rates for (un)folding interconversion of SH3 in water and ILs.**

**a)** Interconversion rates for **I**) unfolding and **II**) folding ($k_{fu}$ and $k_{uf}$, respectively) as determined from the fitting of experimental ZZex data in water, 0.35 M [Ch][Glu] and 0.15 M [Bmim][dca]. These rates were corrected accordingly with their relative viscosities ($\eta_{rel}$ used in 0.35 M [ChGlu], and 0.15 M [Bmim][dca] are 1.39, and 1.02, respectively.) Black, purple, and yellow bars correspond to the viscosity-adjusted rate in water, [Ch][Glu], and [Bmim][dca], respectively. The horizontal dashed lines correspond to their 15-residues average. **b)** 3D structure of SH3 in the folded state (PDB: 2A36[6]) highlighting the followed residues (orange). Only the residues accurately followed are shown (i.e., (un)folding interconversion rates simultaneously measured).

**Table 3.5: Average parameters for the drkN SH3 interconversion extracted from ZZex.**

| | $p_f$ | $p_u$ | $R_{1f}$ ($s^{-1}$) | $R_{1u}$ ($s^{-1}$) | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) | $k_{ex}$ ($s^{-1}$) |
|---|---|---|---|---|---|---|---|
| **Water** | 0.54 ± 0.08 | 0.46 ± 0.08 | 2.5 ± 0.2 | 2.3 ± 0.3 | **0.6 ± 0.1** | **0.8 ± 0.2** | **1.4 ± 0.2** |
| **0.35 M** | 0.83 ± 0.05 | 0.17 ± 0.05 | 2.1 ± 0.1 | 2.2 ± 0.3 | 0.16 ± 0.07 | 0.8 ± 0.3 | 1.0 ± 0.3 |
| **[Ch][Glu]** | | | | | **0.2 ± 0.1[a]** | **1.1 ± 0.4[a]** | **1.3 ± 0.3[a]** |
| **0.15 M** | 0.20 ± 0.07 | 0.80 ± 0.06 | 2.6 ± 0.5 | 2.3 ± 0.2 | 1.0 ± 0.3 | 0.24 ± 0.09 | 1.3 ± 0.3 |
| **[Bmim][dca]** | | | | | **1.0 ± 0.3[a]** | **0.25 ± 0.09[a]** | **1.3 ± 0.3[a]** |

Average from 15 residues values extracted from fitting in MATLAB for SH3 in water, 0.35 M [Ch][Glu], and 0.15 M [Bmim][dca] at 293.2 K, 600.13 MHz. Uncertainties represent the standard deviation from the average. [a] Viscosity-adjusted rates, the relative viscosity used for 0.35 M [Ch][Glu] and 0.15 M [Bmim][dca] was 1.39, and 1.02, respectively.

For two-state folding proteins, interconversion between the folded and unfolded ensemble is limited by the formation of a transition state (TS‡), representing the ensemble of conformations of highest free energy along the protein folding pathway[57]. Together with structural characterization of the folded and unfolded state, the characterization of TS‡ ensemble is vital to understand the free energy landscape of protein folding, however this state is difficult to analyse since TS‡ is only transiently populated. Using the folding and unfolding rates extracted from ZZex data, I defined the activation free energies required to reach the transition state (TS‡) from the unfolded ensemble (folding, $U \rightarrow TS‡$), $\Delta G_{U \rightarrow TS‡}^{0'‡}$ and folded state (unfolding, $F \rightarrow TS‡$), $\Delta G_{F \rightarrow TS‡}^{0'}$. The transition state theory[89] yields simple mono-exponential rate laws for a reaction across an energy barrier of the form[90,91]

$$k_{F,U \rightarrow TS‡} = \kappa \frac{k_B T}{h} exp\left(\frac{-\Delta G_{F,U \rightarrow TS‡}^{0'‡}}{RT}\right), \quad [3.15]$$

where $k_{F,U \rightarrow TS‡}$ is the folding or unfolding rate constant at absolute temperature $T$, $\Delta G_{F,U \rightarrow TS‡}^{0'‡}$ is the mean difference in energy between the conformations at the saddle point of the reaction and the ground state (F or U state) at $T$, $R$ is the gas constant, $h$ is Planck's constant, $k_B$ is the Boltzmann constant, and κ is a transmission coefficient, related to the probability that the reaction proceeds to product from the transition state. For the estimation of barrier heights, κ is assumed as 1.0. For clarity, $k_{fu} = k_{F \rightarrow TS‡}$, $k_{uf} = k_{U \rightarrow TS‡}$, and $\Delta G_u^{0'} = \Delta G_{F \rightarrow U}^{0'}$.

Together, $\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$ and $\Delta G_{UI\rightarrow TS\ddagger}^{0'\ddagger}$ describe the equilibrium between the thermodynamic states and the transition state,

$$\Delta G_{F\rightarrow U}^{0\prime} = -RTln\left(\frac{k_{F\rightarrow TS\ddagger}}{k_{U\rightarrow TS}}\right) = \Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger} - \Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}. \quad [3.16]$$

These activation parameters are listed, per residue, for the different conditions in the following Table 3.6. The viscosity-adjustment rate lowers the values of $\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$ and $\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}$ in 0.35 M [ChGlu], but in 0.15 M [Bmim][dca] the changes are insignificant. Overall, the interpretations are the same to those without viscosity correction.

**Table 3.6. Activation parameters for SH3 folding and unfolding.**

| | water | | | 0.35 M [Ch][Glu] IL | | | 0.15 M [Bmim][dca] IL | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\Delta G$ F→ TS‡ kcal/ mol | $\Delta G$ U→TS‡ kcal/ mol | $\Delta G_u$ kcal/ mol | $\Delta G$ F→ TS‡ kcal/ mol | $\Delta G$ U→TS‡ kcal/ mol | $\Delta G_u$ kcal/ mol | $\Delta G$ F→ TS‡ kcal/ mol | $\Delta G$ U→TS‡ kcal/ mol | $\Delta G_u$ kcal/ mol |
| **S10** | 17.43 | 17.17 | 0.26 | 18.25<br>18.06 [a] | 17.49<br>17.30 [a] | 0.76 | 17.38<br>17.37 [a] | 17.95<br>17.94 [a] | -0.57 |
| **T12** | 17.29 | 17.42 | -0.12 | 18.24<br>18.05 [a] | 17.08<br>16.89 [a] | 1.16 | 17.04<br>17.03 [a] | 18.04<br>18.02 [a] | -0.98 |
| **A13** | 17.43 | 17.23 | 0.19 | 18.29<br>18.10 [a] | 17.51<br>17.32 [a] | 0.78 | 17.21<br>17.20 [a] | 17.83<br>17.82 [a] | -0.62 |
| **E16** | 17.59 | 17.69 | -0.10 | 18.21<br>18.02 [a] | 17.12<br>16.93 [a] | 1.09 | 17.51<br>17.50 [a] | 18.46<br>18.45 [a] | -0.95 |
| **S18** | 17.42 | 17.25 | 0.17 | 17.68<br>17.49 [a] | 17.13<br>16.94 [a] | 0.56 | 17.12<br>17.11 [a] | 17.80<br>17.79 [a] | -0.68 |
| **T22** | 17.38 | 17.16 | 0.22 | 18.16<br>17.97 [a] | 17.34<br>17.15 [a] | 0.82 | 17.01<br>17.01 [a] | 17.72<br>17.71 [a] | -0.71 |
| **L25** | 17.52 | 17.21 | 0.31 | 18.47<br>18.28 [a] | 17.02<br>16.83 [a] | 1.45 | 17.06<br>17.05 [a] | 17.97<br>17.97 [a] | -0.91 |
| **L28** | 17.30 | 17.33 | -0.03 | 18.55<br>18.36 [a] | 17.56<br>17.37 [a] | 0.99 | 17.00<br>16.99 [a] | 18.03<br>18.02 [a] | -1.03 |
| **N29** | 17.51 | 17.11 | 0.40 | 18.35<br>18.16 [a] | 17.47<br>17.28 [a] | 0.88 | 17.16<br>17.15 [a] | 17.65<br>17.64 [a] | -0.49 |
| **A39** | 17.23 | 17.46 | -0.23 | 18.28<br>18.09 [a] | 17.30<br>17.11 [a] | 0.98 | 16.98<br>16.97 [a] | 18.05<br>18.04 [a] | -1.08 |
| **D42** | 17.37 | 17.41 | -0.04 | 18.37<br>18.18 [a] | 17.33<br>17.14 [a] | 1.03 | 16.90<br>16.89 [a] | 18.08<br>18.07 [a] | -1.18 |
| **G46** | 17.42 | 17.28 | 0.14 | 18.37<br>18.18 [a] | 17.33<br>17.14 [a] | 1.03 | 17.21<br>17.20 [a] | 17.97<br>17.96 [a] | -0.76 |
| **I53** | 17.32 | 17.32 | 0.00 | 18.18<br>17.99 [a] | 17.32<br>17.13 [a] | 0.87 | 17.05<br>17.04 [a] | 18.09<br>18.08 [a] | -1.04 |
| **K56** | 17.43 | 17.14 | 0.29 | 18.09<br>17.90 [a] | 17.02<br>16.83 [a] | 1.07 | 17.57<br>17.56 [a] | 18.21<br>18.20 [a] | -0.64 |
| **W36sc** | 17.57 | 17.57 | 0.00 | 18.52<br>18.33 [a] | 17.51<br>17.32 [a] | 1.01 | 17.30<br>17.29 [a] | 18.26<br>18.25 [a] | -0.95 |
| **Average** | 17.4 ± 0.1 | 17.3 ± 0.2 | 0.1 ± 0.2 | 18.3 ± 0.2<br>18.1 ± 0.2 [a] | 17.3 ± 0.2<br>17.1 ± 0.2 [a] | 1.0 ± 0.2 | 17.2 ± 0.2<br>17.2 ± 0.2 [a] | 18.0 ± 0.2<br>18.0 ± 0.2 [a] | -0.8 ± 0.2 |

Values per residue calculated with Eqs. 3.15 and 3.16. $\Delta G_{F\rightarrow TS\ddagger}$ and $\Delta G_{U\rightarrow TS\ddagger}$ are the modified standard-state activation free energies for unfolding and folding, respectively. Uncertainties represent the standard deviation from the average. [a]Viscosity-adjusted rates. The relative viscosity ($\eta_{rel}$) used in 0.35 M [ChGlu], and 0.15 M [Bmim][dca] are 1.39, and 1.02, respectively.

As I cannot know absolute free energies, I can only understand the IL effects to the folded state or the unfolded ensemble by the quantification of IL-induced changes in the barriers to folding ($\Delta\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}$) and unfolding ($\Delta\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$), compared to pure water,

$$\Delta\Delta G_u^{0'} = \Delta G_{u,cosolute}^{0'} - \Delta G_{u,water}^{0'} = \Delta\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger} - \Delta\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}. \quad [3.17]$$

A decreased folding rate in aqueous-IL results from a positive change in the activation energy of folding, indicating that the co-solute makes folding more difficult by raising the barrier, and *vice versa*. The results for the equilibrium stability ($\Delta\Delta G_u^{0'}$) in the different studied conditions (Table 3.7) agree with the earlier results from temperature dependence experiments (Table 3.4).

**Table 3.7: Activation parameters and excess changes for SH3 folding and unfolding.**

| | $\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$, kcal/ mol | $\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}$, kcal/mol | $\Delta G_u^{0'}$, kcal/mol | $\Delta\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$, kcal/mol | $\Delta\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}$, kcal/mol | $\Delta\Delta G_u^{0'}$, kcal/mol |
|---|---|---|---|---|---|---|
| **Water** | 17.4 ± 0.1 | 17.3 ± 0.2 | 0.1 ± 0.2 | - | - | - |
| **0.35 M [Ch][Glu]** | 18.3 ± 0.2 | 17.3 ± 0.2 | 1.0 ± 0.2 | 0.9 ± 0.2 | 0.0 ± 0.3 | 0.9 ± 0.3 |
| | 18.1 ± 0.2 [a] | 17.1 ± 0.2 [a] | | 0.7 ± 0.2 [a] | -0.2 ± 0.3 [a] | |
| **0.15 M [Bmim][dca]** | 17.2 ± 0.2 | 18.0 ± 0.2 | -0.8 ± 0.2 | -0.2 ± 0.2 | 0.7 ± 0.1 | -0.9 ± 0.1 |
| | 17.2 ± 0.2 [a] | 18.0 ± 0.2 [a] | | -0.3 ± 0.2 [a] | 0.7 ± 0.1 [a] | |

Average 15-residues values calculated with Eqs. 3.15-17 and excess changes using as reference water. Uncertainties represent the standard deviation from the average. [a] Viscosity-adjusted rates.

The results indicate that [Ch][Glu] IL stabilises drkN SH3, raising the energy barrier $\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$ and lowering or having no effect on $\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}$ ($\Delta\Delta G_{F\rightarrow TS\ddagger} \gg \Delta\Delta G_{U\rightarrow TS\ddagger}$, Fig. 3.30 and Table 7). The IL substantially slows the protein unfolding rate 3-fold and increases folding 1.4-fold (Table 3.5). This behaviour has been shown for macromolecular crowders (e.g., Ficoll) which is due to entropically-driven excluded volume effects, and since the viscosity increase is accounted in this analysis, an entropic pressure for protein compaction is a plausible explanation for that stabilisation[92], arguing for a entropically favourable release of water from the surface[73]. I hypothesise that [Ch][Glu] acts *via* preferential folded-state accumulation (comparing to preferential hydration as discussed above), where the folded state is promoted through slow unfolding, raising the F → TS‡ energy barrier, assuming that the change in the nature of the transition state is insignificant.

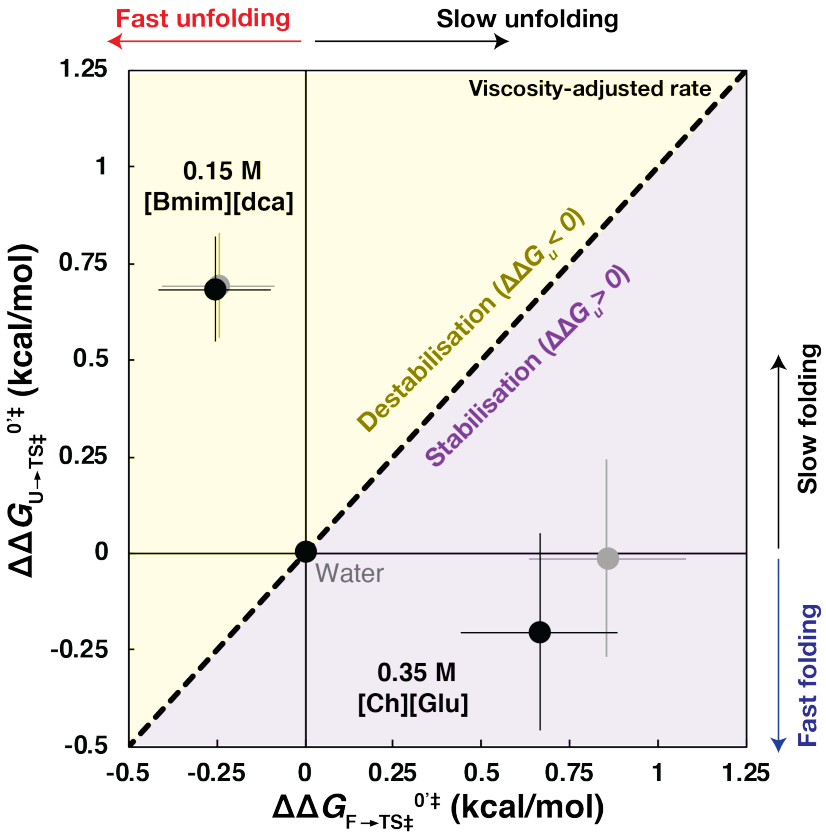



**Figure 3.30. Unfolding-folding compensation plot.**

The different segments correspond to different contributions of $\Delta\Delta G_u^{0'}$, $\Delta\Delta G_{UI\rightarrow TS\ddagger}^{0'\ddagger}$, $\Delta\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$, here calculated at 293 K, approximately $T_{m,\,water}$. The blue diagonal corresponds to a complete unfolding-folding compensation and separates the protein destabilising region ($\Delta\Delta G_u^{0'} < 0$) from the stabilising region ($\Delta\Delta G_u^{0'} > 0$). Positive (negative) values of $\Delta\Delta G_{F,UI\rightarrow TS\ddagger}^{0'\ddagger}$ imply stabilisation (destabilisation) of F and/or U state by the raising (decreasing) of the barrier to transition state formation. Black dots - Experimental data with viscosity-adjusted rate; Grey dots - Experimental data before viscosity correction. Error bars were calculated by error propagation of the primary standard deviation from the 15 residues.

[Bmim][dca] destabilises drkN SH3, lowering $\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$ and increasing $\Delta G_{UI\rightarrow TS\ddagger}^{0'\ddagger}$ ($\Delta\Delta G_{F\rightarrow TS\ddagger} \ll \Delta\Delta G_{UI\rightarrow TS\ddagger}$, Fig. 3.30 and Table 3.7). [Bmim][dca] slowed folding 3.2-fold and increased unfolding 1.6-fold (Table 3.5). Thus, it seems that the unfolded ensemble is more affected than the folded state. This is similar to the effect of urea at a high concentration (> 1.5 M) on SH3 which slowed folding 5-fold and increased unfolding 3-fold as observed by $^{19}$F experiments[20,22]. I speculate that $[Bmim]^+$ and $[dca]^-$ ions penetrate the folded state to speed unfolding, but more importantly, also interact as strong hydrophobic ion-pair with the non-native hydrophobic segment in the unfolded ensemble slowing folding (more than threefold). As discussed above, these interactions should be unfavourable with the protein since a positive $\Delta\Delta H_u$ is raised (Table 3.4 and Fig. 3.24) and not a negative $\Delta\Delta H_u$ as expected with urea or guanidinium chloride from favourable hydrogen bonding with the protein backbone[93].

Note that the increase of helical propensity in the unfolded ensemble under [Bmim][dca] conditions does not introduce a separate helical unfolded state distinct from other unfolded state conformations but simply lowers the average energy level of the unfolded state which is due to the rapid interconversion between various conformational states in the ensemble[3]. This is validated by the fact that backbone U-state NH resonances (Fig. 3.21c) correspond to unfolded conformations in fast exchange on the NMR chemical shift time scale. Thus, accordingly with ZZex data, the non-native helical propensity appears to slow protein folding (stabilisation of unfolded state) in a predictable manner. In agreement, this has been shown for diverging-turn T22 mutants of drkN SH3[15], or for β1-strand mutants of α-spectrin SH3 domain[94].

In summary, assuming that the energy level of the transition state does not change, different evidence supports the two energy level diagrams for drkN SH3 as shown in Fig. 3.31, where changes in the folding and unfolding rates are accounted for by changes in the energy levels of the unfolded and folded states.

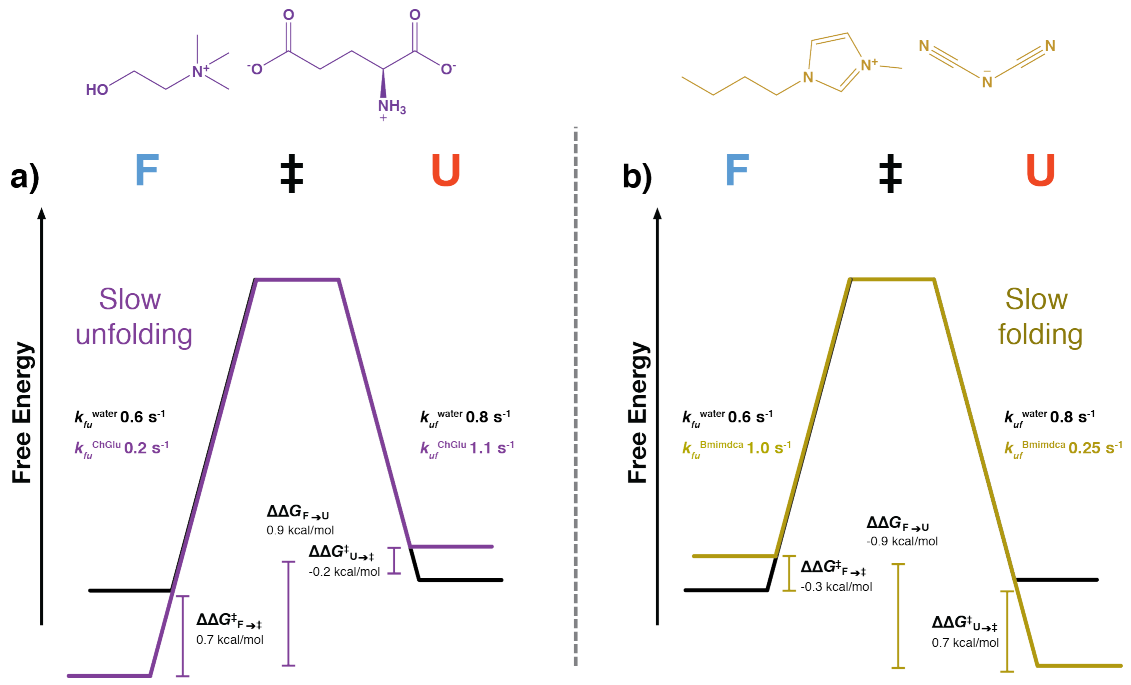



**Figure 3.31. ILs dependence of the drkN SH3 free energy landscape.**

**a)** free energy landscape model for the stabilisation of SH3 in [Ch][Glu]. **b)** free energy landscape model for the destabilisation of SH3 in [Bmim][dca]. Models **a**) and **b**) assume that the transition state energy is unaffected by ionic liquid and that the thermodynamic and kinetic behaviour can be explained by changes in the energies of both the unfolded and folded states. The terms $k_{fu}$ and $k_{uf}$ indicate the interconversion rate of drkN SH3 (F→TS‡ or U→TS‡). The [Ch][Glu] (purple) or [Bmim][dca] (dark yellow) superscripts reflect the average measured rate from 15-residue values in the presence of a) 0.35 M [Ch][Glu] or b) 0.15 M [Bmim][dca], respectively. $\Delta\Delta G_u^{0'}$, $\Delta\Delta G_{F,U\rightarrow TS\ddagger}^{0'\ddagger}$ average values and uncertainties were calculated from the equations described in the text.

Although our results point out largest differences in the unfolding or folding rate of SH3 in [Ch][Glu] or [Bmim][dca] are attributed to changes in the folded or unfolded state energies, respectively, I must consider how these effects could be explained by changes in the energy level of the folding transition state (Fig. 3.32).

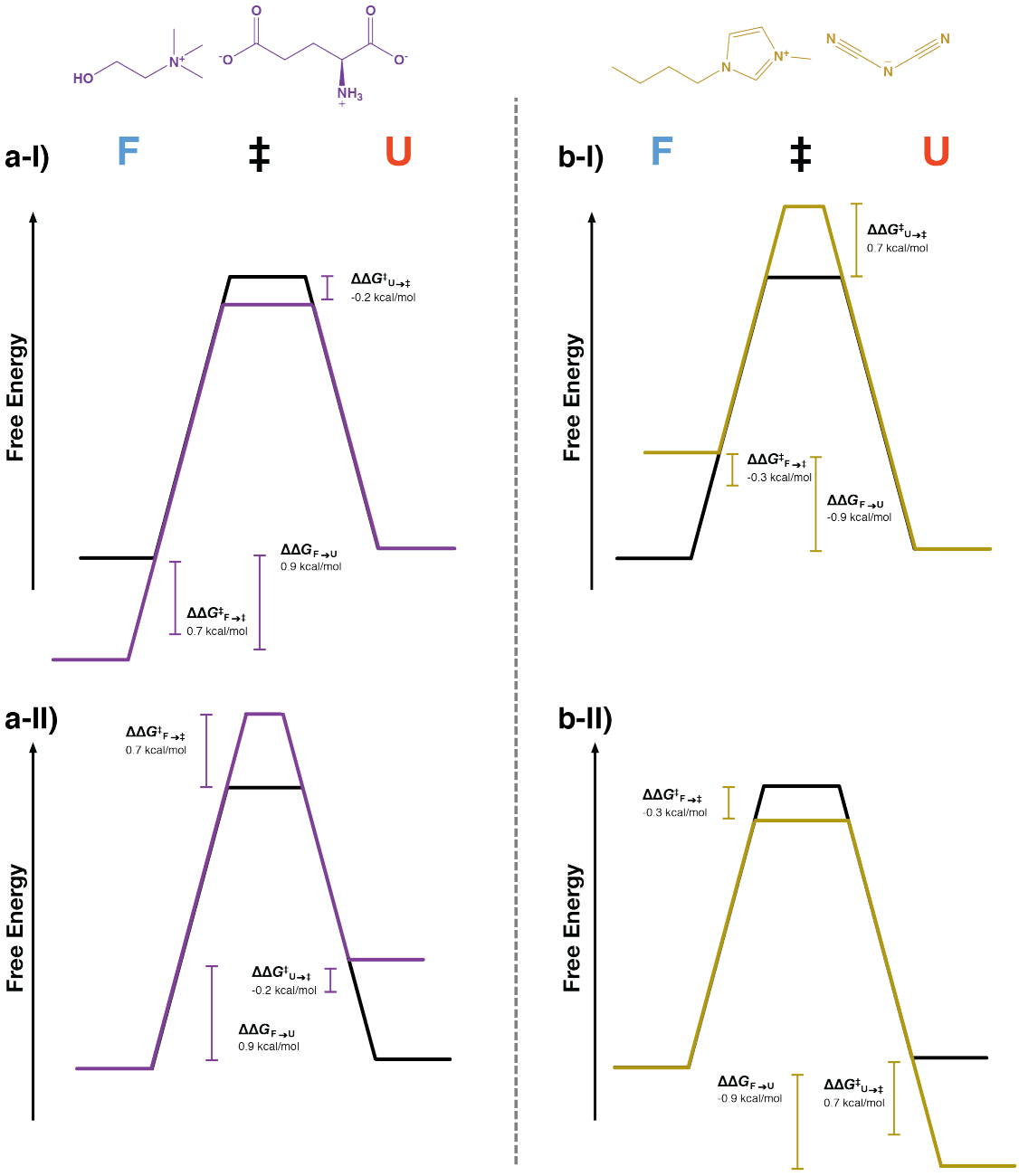



**Figure 3.32. ILs dependence of the drkN SH3 free energy landscape with changes in the transition state.**

**a)** free energy landscape model for the stabilisation of SH3 in [Ch][Glu]. **b)** free energy landscape model for the destabilisation of SH3 in [Bmim][dca]. Models **a**) assumes that the unfolded state energy is unaffected by ionic liquid and that the thermodynamic and kinetic behaviour can be explained by changes in the energies of the folding transition and folded states. Models **b**) assumes that the folded state energy is unaffected by ionic liquid and that the thermodynamic and kinetic behaviour can be explained by changes in the

energies of the folding transition and unfolded states. $\Delta\Delta G_u^{0'}$, $\Delta\Delta G_{F,U\rightarrow TS\ddagger}^{0'\ddagger}$ average values and uncertainties were calculated from the equations described in the text.

A proportionally constant $\alpha$ can be defined by relating changes in activation free energy for folding, $\Delta\Delta G_{U\rightarrow TS\ddagger}$ to changes in the folding free energy, $\Delta\Delta G_{U\rightarrow F}$, upon perturbation as,

$$\alpha = \frac{\Delta\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}}{\Delta\Delta G_{U\rightarrow F}^{0'}}, \qquad [3.18]$$

where $\alpha$ values provide a measure of the energetic sensitivity of the transition state to the perturbation relative to the folded and unfolded states and can also be indicative of the relative solvent exposure of the folding TS‡ ensemble[86]. Typically, Eq. 3.18 is used to yield $\Phi_f$ values upon single mutations[95], where a $\Phi_f$ value of 0 means that the interaction measured is as poorly formed in the transition state as it is in the unfolded state, and a value of 1 means that it is as well formed as it is in the native state. Based on the average $\Delta\Delta G_{U\rightarrow TS\ddagger}$ and $\Delta\Delta G_{U\rightarrow F}$ values extracted from the 15 residues followed by ZZex and corrected by the effect of viscosity (Table 3.6 and Table 3.7), I measured for 0.35 M [Ch][Glu] and 0.15 M [Bmim][dca], $\alpha$ values for folding of $\alpha_{[Ch][Glu]}= 0.2 \pm 0.6$ and $\alpha_{[Bmim][dca]}= 0.7 \pm 0.3$, respectively. Interestingly, the $\alpha_{[Bmim][dca]}$ obtained for 0.15 M of [Bmim][dca] is equivalent to those with ~ 1 M of urea or ~ 1.4 M of glycerol ($\alpha_{[Glycerol]} \approx \alpha_{[urea]} \approx 0.75 \pm 0.1$), indicating that TS‡ ensemble is is ~25% hydrated[86] and it is partially destabilised since the helix is partially formed in the transition state (closer to native-like structure). The low $\alpha_{[Ch][Glu]}$ value obtained for 0.35 M of [Ch][Glu] indicates that the transition state is not destabilised, but instead, partially stabilised (closer to the unfolded structure). This is assuming that the slowing of the folding rates of stabilised (or destabilised) SH3 as compared to water is due solely to the decreased (or raised) energy level of the TS‡ relative to the unfolded state (Fig. 3.32-I).

However, interpretation of the native structural propensities discussed above is somewhat ambiguous since the origin of these propensities is not well understood. The three models (including Fig. 3.32-II – folded state unaffected) should be used to relate the three states (unfolded, transition state and folded state) and highlights the importance to consider the unfolded state to understand protein stability.

## Conclusion

In this chapter, NMR was used to quantify the changes in structural and folding-unfolding kinetics of drkN SH3 in stabiliser or destabilisers aqueous-IL conditions (i.e., the presence of [Ch][Glu] or [Bmim][dca], respectively). The protein folding landscape was also characterized with changes in the energy barriers, unfolding (F→TS‡) and folding (U→TS‡), discerning the (de)stabilisation degree of folded and unfolded state by ILs. Our data shows, that for the drkN SH3 domain, the stabilisation observed with [Ch][Glu] is due to a native state stabilisation (i.e., [Ch][Glu] molecules accumulate in the surface of the folded state), while the observed denaturation is in fact caused by a stabilisation of a more compact structure of the unfolded ensemble (preferential binding of [Bmim][dca] molecules), with

evidence from local backbone perturbations and a significative increase of $\alpha$-helix propensity (between residues R20 and L28). This is rather unexpected since proteins under denaturing conditions tend to lose residual structure and approach random-coil-like-behavior[9], pointing that destabilisation by IL is different from common denaturation using agents as urea or GdmCl. The data gathered can also point to new direction for the understanding of how changes in the cellular homeostasis can influence the protein folding landscapes.

## Materials and Methods

Description of chemicals and materials and a detailed protocol for expression and purification $^{15}N$ or $^{15}N/\ ^{13}C$ isotopically labelled drkN SH3 can be found in this section. The synthesis and characterization of [Ch][Glu] IL was already described in the previous chapter 2. NMR sample preparation, data collection and analysis procedures (quantification of the F/U population; temperature scan protocol; and ZZex data including thermodynamic and kinetic examination) are described along the section, including further details in viscosity measurements and extrapolations, tables of affected residues in chemical shift perturbation, thermodynamic and kinetic parameters.

**Chemicals and Materials.** L-glutamic acid (> 98.5 % of purity) was purchased from PanReac. 1-butyl-3-methylimidazolium dicyanamide ([Bmim][dca]) and 1-butyl-3-methylimidazolium chloride ([Bmim][Cl]) ILs were sourced from IoLiTec (Denzlingen, Germany). The ILs were at least 98% pure and were dried for 24 h under vacuum at 60 ºC before solution preparation. Isotopically enriched chemical compounds, [$^{15}N$]-$NH_4Cl$, [$^{13}C$]-glucose and deuterium oxide ($D_2O$) were purchased from Cambridge Isotope Laboratories. Sodium-L-glutamate, ion exchange resin Amberlyst A-26 (OH) and SnakeSkin 3.5K MWCO dialysis tubing were purchased from Thermo Fisher Scientific. Unless otherwise described, all other chemicals were purchased from Sigma-Aldrich. Pure *milliQ* water with resistance 18.4 M$\Omega$ $cm^{-1}$ was used in all experiments. pH values are direct meter readings uncorrected for any isotope effect and were measured with Docu-pH meter (Sartorius) calibrated with standard solutions.

**Protein expression and purification.** $^{15}N$ or $^{15}N/\ ^{13}C$ isotopically enriched, N-terminal Src homology 3 (SH3) domain of *Drosophila* signal transduction protein drk (drkN SH3) containing 59 residues, was overexpressed and purified as previously described[8]. *Escherichia coli* BL21 (DE3) competent cells (NZYTech) were transformed with the pET-11d plasmid vector containing the gene encoding drkN SH3 protein by heat shock. Cells were grown with shaking at 37 ºC and 180 rpm in M9 minimal medium supplemented with 100 μM $FeSO_4$, 100 μM $CaCl_2$, 2 mM MgSO4, 10 mg/L Thiamine-HCl, 0.5% MEM Vitamins, 1.5 g/L of $^{15}NH_4Cl$ and 4 g/L of glucose or $^{13}C$-glucose, and 100 mg/L ampicillin. Protein expression was induced by adding 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG, NZYTech) at $OD_{600}$ (optical density at 600 nm) = 0.7. After 2h induction in the same conditions, cells were harvested [4,425 x *g*, 30 min, 4ºC in a JA-10 rotor (Avanti J-26S XPI, Beckman Coulter)] and frozen at -20 ºC overnight. Pellet was resuspended in 30 mL of lysis buffer per liter of culture [50 mM Tris-HCl (pH 7.5), 2 mM EDTA, 7 mM β-mercaptoethanol and protease inhibitor cocktail (cOmplete ULTRA tablets, Roche)], and then lysed by sonication on ice [10 min at 80% amplitude using 1 min on/off pulse program (UP100H ultrasonic processor, Hielscher)] and centrifuged [30,000 x *g*, 40 min, 4ºC in a JA-25.50 rotor (Avanti J-26S XPI, Beckman Coulter)]. Supernatant was dialyzed (SnakeSkin 3.5K MWCO) overnight at 4ºC against Buffer A [50 mM Tris-HCl (pH 7.5), 2 mM EDTA and 7 mM β-mercaptoethanol]. The dialyzed sample was loaded to a HiTrap Q HP anion-exchange column (GE Healthcare) on an ÄKTA start

system (GE Healthcare), and buffer B [50 mM Tris-HCl (pH 7.5), 1 M NaCl, 2 mM EDTA and 7 mM β-mercaptoethanol] was used to produce a gradient 0 - 500 mM NaCl where drkN SH3 elutes at around 150 mM NaCl since the binding with the column is weak.

Fractions containing the protein of interest were pooled and concentrated with Vivaspin turbo 15 3 K MWCO centrifugal concentrator (Sartorius). In a final step, a size exclusion chromatography step [Superdex 75 10/300 GL column (GE Healthcare) in Shimadzu prominence machine, at 4ºC] was applied to further purify and adjust drkN SH3 to a different buffer [50 mM phosphate buffer (pH 7.2) and 150 mM NaCl]. Purity was analysed in each step by SDS-PAGE (BioRad). Pure samples were extensively desalted by dialysis in *milliQ* water, flash frozen, lyophilized (Edwards Modulyo Freeze Dyer) and stored at -20 ºC until usage. Final yield of purified protein was 7 mg of drkN SH3 per liter of M9 minimal medium. Concentrations were assessed spectrophotometrically by absorption measurements at 280 nm ($\varepsilon$= 8,480 $M^{-1}$ $cm^{-1}$) using NanoDrop ND-1000 UV-Vis (Thermo Fisher Scientific).

**NMR spectroscopy**

Sample preparation, data acquisition and processing. All samples for NMR spectroscopy were prepared in *milliQ* water with 10% (v/v) $D_2O$, 0.1% (v/v) $NaN_3$ and 50 $\mu$M sodium-2,2-dimethyl-2-silapentane-5-sulfonate-$d_6$ (DSS, Eurisotop), pH was adjusted for 7.1 ± 0.1 with negligible microliter addition of HCl or NaOH solutions. Except where stated, spectra were recorded at 298.2 K on a 600 MHz Bruker Avance II+ spectrometer operating at a proton Larmor frequency of 600.13 MHz equipped with a 5-mm TCI cryoprobe. Spectrometer temperature was calibrated using a pure methanol-$d_4$ standard Bruker sample[96]. Further specific details of data acquisition (pulse programs, spectral windows, etc) and processing for each experiment are given in the following subsections. Proton chemical shifts were referenced against internal DSS while nitrogen and carbon chemical shifts were referenced indirectly to DSS using the absolute frequency ratio[97]. Processing and visualization of NMR spectra were performed with NMRPipe (version 10.1)[98] and CcpNmr Analysis 2.5[99]. Preliminary drkN SH3 assignments were taken from the BMRB databank under the accession code 25501[18].

NMR Chemical Shift Titrations. The standard Bruker $^{1}H$-$^{15}N$ HSQC (hsqcetfpf3gpsi2) pulse sequence employs a sensitivity-enhanced pulse field gradient[100]. The spectra were acquired with 2048 ($^{1}H$) and 128 ($^{15}N$) complex points for a spectral width (SW) of 12 ppm ($^{1}H$) and 32 ppm ($^{15}N$), with 8 scans.

Two-dimensional $^{1}H$-$^{15}N$ HSQC experiments were subsequent collected for different ionic liquids ([Ch][Glu], [Bmim][dca], [Bmim][Cl]) and salts ([NaGlu], [Na][dca], [ChCl], [NaCl]) titrations with 350 $\mu$M $^{15}N$-labelled SH3 where the co-solute concentrations were incrementally increased up to 1 M (e.g., 0, 0.05, 0.1, 0.25, 0.5, 1.0 M of IL/ salt concentration). Similarly, $^{1}H$-$^{15}N$ HSQC spectra were obtained for [$Na_2SO_4$] and [GdmCl] salts at 0.4 M and 2.0 M of concentration, respectively.

Combined $^{1}H$-$^{15}N$ chemical shift differences of the amide in 2D [$^{1}H$-$^{15}N$]-HSQC spectra, $\Delta\delta_{comb}$ or chemical shift perturbations (CSP), were calculated as

$$\Delta\delta_{comb} = \sqrt{(\Delta\delta(H_N)^2 + (\alpha\Delta\delta N_H)^2}, \quad [3.19]$$

where $\Delta\delta H_N$ and $\Delta\delta N_H$ are the $^1H$ and $^{15}N$ chemical shift differences of SH3 in the presence of added IL/ salt minus the same resonance in the absence of added species, normalized with the scaling factor $\alpha$= 0.14 for most residues but $\alpha$= 0.2 for glycine[31]. To decide whether a given residue belongs to the class of interacting or non-interacting residues, I have calculated a cut-off value, based on the corrected standard deviation to zero, according to the method developed by Schumann et al[32]. Since the total concentrations of ligand and protein are known during titration, $[L]_t$ and $[P]_t$, for residues above the cut-off, the $\Delta\delta_{comb}$ values were used to obtain the dissociation constant ($K_d$) from the titration experiments ([Ch][Glu] and [Bmim][dca]) by non-linear regression analysis according to[31],

$$\Delta\delta_{comb} = \Delta\delta_{max}\frac{(K_d + [L]_t + [P]_t) - \sqrt{(K_d + [L]_t + [P]_t)^2 - (4[P]_t\,[L]_t)}}{2[P]_0}, \quad [3.20]$$

where $\Delta\delta_{comb}$ is the combined chemical shift deviation from the free state defined by Eq. 3.19 and $\Delta\delta_{max}$ is the maximum chemical shift change on saturation. Data are fitted to a single site binding model, using a least square fitting search of Microsoft Excel Solver to find the values of $K_d$ and the chemical shift of the fully saturated protein.

SH3 backbone assignments. Protein backbone resonance assignment experiments (Bruker standard pulse sequences) were recorded for 0.75 mM $^{15}N^{13}C$ drkN SH3 at 0, 0.35, 0.65, 1.0 M [Bmim][dca] and 1.0 M [Ch][Glu], comprising 2D $^1H$-$^{15}N$ HSQC and 3D triple resonance experiments HNCACB, HNcoCACB, HNCO, HNcaCO spectra. All spectra were collected using Echo/Antiecho-TPPI gradient selection, which was efficient at suppressing signals from ionic liquid. Acquisition parameters for 2D $^1H$-$^{15}N$ HSQC experiments were 2048 ($^1H$) and 512 ($^{15}N$) complex points for a SW of 12 ppm ($^1H$) and 32 ppm ($^{15}N$), with 16 scans. For 3D experiments, 2048 ($^1H$), 40 ($^{15}N$) and 256 ($^{13}C$) complex points for a SW of 12 ppm ($^1H$), 32 ppm ($^{15}N$) and 75 ppm for HNCAB/ HNcoCACB or 18 ppm for HNCO/ HNcaCO ($^{13}C$), were used with 16 scans except HNCACB with 32 scans. Non-uniform sampling (NUS) was used to optimize resolution of the indirect dimensions in the available experiment time, where just 20-25% of sparse data was recorded. NUS acquired data were processed with SMILE algorithm within NMRPipe for spectra reconstruction[101]. Sequential connectivities were performed with CcpNmr AnalysisAssign 3.0[102] followed by manual verification. Backbone assignments for both folded and unfolded drkN SH3 in water were consistent with previous assignments under buffer conditions[7,18], and tryptophan indole side chain resonances were assigned by analogy to previously published data[8]. The backbone assignment for $U_{[Bmim][dca]}$ state led to unambiguous assignment in the 2D $^1H$-$^{15}N$ HSQC even for very overlapped peaks.

Secondary structure propensity (SSP). SSPs were calculated using $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts of drkN SH3 as input according to the SSP protocol by Marsh et al[53]. Positive SSP values ranging from 0 to 1 and negative values from 0 to -1 represent the propensities of $\alpha$ and $\beta$ secondary structures, respectively. Reference values for random coil, secondary structure chemical shifts and standard deviations came from RefDB[103].

ZZex experiments. SH3 interconversion rates were determined by 2D longitudinal $^{15}$N ZZex spectroscopy[80], using the pulse sequence provided by Bruker (hsqcetexf3gp). Sample conditions consisted of 1.1 mM protein in the absence or presence of 350 mM [ChGlu] or 150 mM [Bmim][dca]. A series of 13/14 2D exchange spectra were acquired at 293.2 K with variable mixing time ranging from 10 to 750 ms (e.g. 10, 20, 40, 65, 100, 140, 180, 230, 320, 450, 550, 650, 750 ms of mixing time). Each 2D spectrum was recorded as a complex data matrix of 2048 ($^{1}$H) and 256 ($^{15}$N) points for a spectral width of 12 ppm ($^{1}$H) and 32 ppm ($^{15}$N), 32 scans per FID were obtained with a recycle delay of 1.2 s. These experiments were also employed to verify some of the assignments discussed above, and the ZZex data analysis is explained further below.

Amide proton temperature coefficients. 2D $^{1}$H-$^{15}$N HSQC were performed from 278.2 to 313.2 K in steps of 5 K with 5 min equilibrium time between each measurement to ensure temperature stability, for the $^{15}$N-ZZex samples (water, 0.35 M [ChGlu] and 0.15 M [Bmim][dca]). Each 2D spectrum was recorded as a complex data matrix of 2048 ($^{1}$H) and 256 ($^{15}$N) points for a SW of 12 ppm ($^{1}$H) and 32 ppm ($^{15}$N), with 4 scans. The amide proton temperature coefficients ($\Delta\sigma_{HN}/\Delta T$) were determined from the change in the upfield $^{1}H_N$ chemical shifts with increasing temperature[51]. The $^{1}H_N$ chemical shifts were extracted from peak assignments for each $^{1}$H-$^{15}$N HSQC with CcpNmr 2.5. The values and their uncertainties were determined using a linear least-squares fit of $^{1}H_N$ chemical shifts from temperature increments. The fitting was performed with Prism 8 (GraphPad Software).

**NMR quantification of the F/U population.** The volume intensity for the two resonances from the same residue in each condition were extracted with CcpNmr 2.5. As the accuracy of the population determination is directly linked to the accuracy in peak volume determination that, in turn, is dependent on the signal-to-noise ratio. I also calculate the volume intensity error for each peak volume using a gaussian line shape fitting analysis by PINT[64,104]. The uncertainties in $\Delta G_u^{0'}$ values, as described in the main text, were propagated from peak volume error.

Obtained $\Delta G_u^{0'}$ values were then fitted by a least square analysis using the Prism 8 (GraphPad Software). The *m*-value uncertainty is the standard error from the slope using a simple linear regression.

**NMR temperature scan analysis.** NMR temperature scan data were used to determine the population of folded ($p_f$) and unfolded ($p_u$) protein as a function of temperature, and from these populations, $\Delta G_u^{0'}$ was calculated as described in the main text. The data were fitted to the integrated Gibbs-Helmholtz[55], assuming a constant heat capacity change upon unfolding ($\Delta C_{p,u}^{0'} = C_{p,U}^{0'} - C_{p,F}^{0'}$) over the temperature range studied (278 K to 313 K), as described in the main text. Data were analysed using Prism 8 (GraphPad Software) and the uncertainties are the standard deviation from the best-fit curves with nonlinear regression.

**Analysis of the ZZex spectroscopy data.** The time evolution of the resulting four peaks (auto-peaks (*ff* and *uu*) and exchange cross-peaks (*fu* and *uf*)) can be fit to the appropriate exchange model using the longitudinal Bloch-McConnell equations[83]. Peak heights were used for *I(t)* values, where the full assignment was prepared with CcpNmr 2.5 and the integration was done with PINT[64,104] using a gaussian line shape fitting analysis, optimizing peak positions, line widths and intensity. At $t$=0, the total intensity was set to the sum of the integrations of the *ff* and *uu* peaks and the cross-peak intensities were normalized to the nominal cross peak intensities. Non-linear least-squares and simultaneous fitting of $I_{ff}(t)$, $I_{uu}(t)$, $I_{fu}(t)$, $I_{uf}(t)$ curves for the measurement of chemical exchange ($k_{ex}$) and longitudinal $^{15}$N decay ($R_{1f}$ and $R_{1u}$) rates was conducted with a modified MATLAB 2017b script[82] in MATLAB (Mathworks) employing an exchange model for two-state interconversion as described by Eqs. 3.21-24 that allow the extraction of $k_{fu}$ and $k_{uf}$ interconversion rates. The equations, for a two-state exchange process, describing the dependence of auto (*ff*, *uu*)- and cross (*fu*, *uf*) -peak heights on the variable mixing period, $t$, have been described before[5,80,81] and are given below:

$$I_{ff}(t) = \frac{1}{2}\left[\left(\frac{(1 - R_{1f}^0 - R_{1u}^0 + k_{ex}(p_u - p_f))}{(\lambda_+ - \lambda_-)}\right) e^{(-\lambda_- t)} + \left(\frac{(1 + R_{1f}^0 - R_{1u}^0 + k_{ex}(p_u - p_f))}{(\lambda_+ - \lambda_-)}\right) e^{(-\lambda_+ t)}\right], \quad [3.21]$$

$$I_{uu}(t) = \frac{1}{2}\left[\left(\frac{(1 + R_{1f}^0 - R_{1u}^0 + k_{ex}(p_u - p_f))}{(\lambda_+ - \lambda_-)}\right) e^{(-\lambda_- t)} + \left(\frac{(1 - R_{1f}^0 - R_{1u}^0 + k_{ex}(p_u - p_f))}{(\lambda_+ - \lambda_-)}\right) e^{(-\lambda_+ t)}\right], \quad [3.22]$$

$$I_{fu}(t) = \frac{k_{ex} p_f}{(\lambda_+ - \lambda_-)}\left(e^{(-\lambda_- t)} - e^{(-\lambda_+ t)}\right), \quad [3.23]$$

$$I_{uf}(t) = \frac{k_{ex} p_u}{(\lambda_+ - \lambda_-)}\left(e^{(-\lambda_- t)} - e^{(-\lambda_+ t)}\right), \quad [3.24]$$

*I* refer to the time dependence of the transfer amplitudes (build-up curves) for the *ff*, *uu*, *f* to *u* (*fu*) and *u* to *f* (*uf*) interconversions. *p* refers to the fractional population of the indicated state, $k_{ex}$ is the stochastic exchange of molecules between the two states per second (s), $t$ is time in s. $R_{1f}$ and $R_{1u}$ are the longitudinal relaxation rate constants of magnetization in sites *f* and *u*, respectively. $I_{ff}(0)$ and $I_{uu}(0)$ denote the amount of longitudinal nitrogen magnetization associated with the folded and unfolded states at the start of the mixing period t. In the limit, $R_{1f} = R_{1u}$ and $k_{fu} = k_{uf}$. $\lambda_\pm$ denotes the eigenvalues of the 2x2 dynamics matrix ($a_{1,2}$) describing the loss of magnetization in the folded and unfolded states due to longitudinal relaxation and chemical exchange, given by,

$$\lambda_\pm = \frac{1}{2}\left\{R_{1f}^0 + R_{1u}^0 + k_{ex} \pm \left[\left(R_{1f}^0 + R_{1u}^0 + k_{ex}\,(p_u - p_f)\right)^2 + 4 p_f p_u {k_{ex}}^2\right]\right\}, \quad [3.25]$$

the intensities of the auto-peaks decrease with increasing $t$, while the exchange peaks increase to a maximum in intensity (at $t = \frac{ln\left(\frac{\lambda_-}{\lambda_+}\right)}{(\lambda_- - \lambda_+)}$) and subsequently decrease in intensity.

Analysis of the folding thermodynamics and kinetics. A two-site exchange process is considered according to

$$F \underset{k_{uf}}{\overset{k_{fu}}{\rightleftarrows}} U, \qquad [3.26]$$

where the chemical shift difference between sites is $\delta\omega$, the equilibrium populations of states $F$ and $U$ are $p_f$ and $p_u$ with $p_f + p_u = 1$. The exchange first order rates for the magnetization converting from site f to $u$ ($k_{fu}$) – unfolding rate constant, and $u$ to $f$ ($k_{uf}$) – folding rate constant, were calculated using the relative populations of the two states taken from integrations,

$$k_{ex} = k_{fu} + k_{uf} = \frac{k_{fu}}{p_u} = \frac{k_{uf}}{p_f}, \qquad [3.27]$$

$$k_{fu} = k_{ex} p_u, \qquad [3.28]$$

$$k_{uf} = k_{ex} p_f. \qquad [3.29]$$

The unfolding free energy (F→U, $\Delta G_u^{0'}$) per residue can be calculated using the following relation

$$\Delta G_u^{0'} = -RTln\left(\frac{k_{fu}}{k_{uf}}\right). \qquad [3.30]$$

The (un)folding rates extracted from ZZex data allows the characterization of how the barriers that define SH3 folding (U→TS‡) and unfolding (F→TS‡) are modified with ILs. Based on the transition-state theory[89], the free energies required to reach the transition state (TS‡) from the unfolded ensemble, $\Delta G_{U\rightarrow TS\ddagger}^{0'\ddagger}$, and the folded state, $\Delta G_{F\rightarrow TS\ddagger}^{0'\ddagger}$, are determined by the Eyring-Polanyi equation[90,91],

$$\Delta G_{F,U\rightarrow TS\ddagger}^{0'\ddagger} = -RTln\left(\frac{k_{F,U\rightarrow TS\ddagger} h}{k_B T}\right), \qquad [3.31]$$

where $\Delta G_{F,U\rightarrow TS\ddagger}^{0'\ddagger}$ is the modified standard-state activation free energy for folding or unfolding at absolute temperature $T$, $R$ is the gas constant, $k$ is the folding or unfolding rate at T, $h$ is Planck's constant, and $k_B$ is the Boltzmann constant. For clarity, $k_{fu} = k_{F\rightarrow TS\ddagger}$, $k_{uf} = k_{U\rightarrow TS\ddagger}$, and $\Delta G_{F\rightarrow U}^{0'} = \Delta G_u^{0'}$.

The viscosity effects were also into account, since Tollinger et al[86] show drkN SH3 folding is linearly-dependent on solvent viscosity. In this context, Kramers' model for diffusive barrier crossing[84] appears to be more appropriate than the original transition state formalism because protein folding necessarily involves diffusional events, as the expanded and hydrated unfolded state polypeptide chain collapses toward a more compact and conformationally restricted folded protein (water is expelled from the interior of the protein). In cases where the rate-limiting step involves such diffusive processes and based on the Stokes law, a $1/\eta$ dependence of $k_{fu}$ and $k_{uf}$ on solvent viscosity is expected, whereas rate constants will

be independent of solvent viscosity if the rate-limiting step involves only rearrangements that are limited by the internal friction of the protein.

$$k_{F,U\rightarrow TS\ddagger}(T,\eta) = \frac{A}{\eta} exp\left(\frac{-\Delta G^{0'\ddagger}_{F,U\rightarrow TS\ddagger}}{k_B T}\right), \quad [3.32]$$

where A is a temperature and viscosity-independent constant. Thus, to separate the viscosity effects from those result from increased stability (i.e., viscosity does not change the equilibrium of the folding reaction but affects both folding and unfolding to the same extent), I adjusted the folding/unfolding rates accordingly[22,88]

$$k_{0\ (F,U\rightarrow TS\ddagger)} = k_{F,U\rightarrow TS\ddagger}\left(\frac{\eta_c}{\eta_{water}}\right) = k_{F,U\rightarrow TS\ddagger}\eta_{rel}, \quad [3.33]$$

where $k_0$ is the viscosity-adjusted rate, $k$ is the folding/unfolding rate before viscosity correction, $\eta_c$ is the viscosity of the solution, $\eta_{water}$ is the viscosity of pure water at 293.2 K (1.002 cP)[87], and $\eta_{rel}$ is the viscosity adjusted to water at 293.2 K. Viscosity measurements and extrapolations are described below. Non- and viscosity-adjusted rates were listed for the different conditions (water, 0.35 M [Ch][Glu] and 0.15 M [Bmim][dca] at 293.2 K) in Table B4 of Appendix B.

Rearranging the equation 3.31, the free energies per residue required to reach the transition state (TS‡) are determined by

$$\left(\Delta G^{0'\ddagger}_{F,U\rightarrow TS\ddagger}\right)_{corr} = -RTln\left(\frac{k_{0\ (F,U\rightarrow TS\ddagger)}h}{k_B T}\right). \quad [3.34]$$

Together, $(\Delta G_{F\rightarrow TS\ddagger}{}^{0'\ddagger})_{corr}$ and $(\Delta G_{U\rightarrow TS\ddagger}{}^{0'\ddagger})_{corr}$ describe the correct equilibrium between the thermodynamic states and the transition state,

$$\Delta G^{0'}_u = \left(\Delta G^{0'\ddagger}_{F\rightarrow TS\ddagger}\right)_{corr} - \left(\Delta G^{0'\ddagger}_{U\rightarrow TS\ddagger}\right)_{corr}. \quad [3.35]$$

Average of residue-specific values of $k_{fu}$, $k_{uf}$ and other extracted parameters were employed in all analyses and errors were estimated from the standard deviation of the values over the residues that were included in the average.

**Viscosity measurements and extrapolations.** Viscosity of pure water at 293.2 K is 1.002 cP[87]. The relative viscosity ($\eta_{rel}$) to water at 293.2 K determined to 0.35 M [ChGlu] and 0.15 M [Bmim][dca] are 1.39 and 1.02, respectively. The details of the measurements and extrapolations are as follows.

<u>0.35 M [Ch][Glu] viscosity solution.</u> According to [Ch][Glu] viscosity measurements in chapter 2, the viscosity as a function of $T$ was fit to an exponential function, which was used to extrapolate the $\eta$ at 293.2 K. Likewise, the dependency of the solvent viscosity as a function of [Ch][Glu] concentration ([Ch][Glu]) at 293.2 K can be described by an empirical exponential equation

$$\eta\ (293.2\ K, [Ch][Glu]) = 1.046677 \exp(0.809516\ [Ch][Glu]), \quad [3.36]$$

this equation has no physical meaning but describes the dependence of viscosity on [Ch][Glu] concentration very well at 293.2 K. Viscosity of [Ch][Glu] IL 0.35 M at 293.2 K was calculated to 1.390 cP.

0.15 M [Bmim][dca] viscosity solution. Viscosity of aqueous-[Bmim][dca] IL 0.15 M at 293.2 K was determined using an equation (Eq. 3.32) based upon the Eyring's theory and a modified two-suffix-margules excess Gibbs energy model (Eyring-MTSM) to correlate the dynamic viscosities of binary mixtures of ILs with water[105]:

$$\ln(\eta_{mix}) = x_1 ln(\eta_1) + x_2 ln(\eta_2) + \alpha_{12} \frac{x_1 x_2 G_{12} G_{21}}{(x_1 G_{12} + x_1)(x_1 G_{12} + x_1)}, \quad [3.37]$$

with

$$G_{12} = \exp\left(\frac{\tau_{12}}{RT}\right), \qquad \tau_{12} = g_{12} - g_{22}, \quad [3.38]$$

$$G_{21} = \exp\left(\frac{\tau_{21}}{RT}\right), \qquad \tau_{21} = g_{21} - g_{11}, \quad [3.39]$$

and

$$\alpha_{12} = \alpha_{12}^{(0)} + \frac{\alpha_{12}^{(1)}}{T}, \quad [3.40]$$

where $x_1$, $x_2$, $\eta_1$, $\eta_2$ are the mole fraction and viscosities of component 1 and 2, respectively, $g_{ij}$ is the potential energy between components i and j, R is the gas constant, $\alpha 12$, $\tau 12$ and $\tau 21$ are adjustable parameters, and $\alpha_{12}(0)$ and $\alpha_{12}(1)$ are adjustable temperature dependent considerations. The interaction parameters of the model for [Bmim][dca] (component 1) and water (component 2) were taken from López-Córtez et al[106], $\alpha_{12}(0)$, $\alpha_{12}(1)$, $\tau 1_2$ and $\tau_{21}$ are 0.9838, 428.2, 407.3, -467.9, respectively. [Bmim][dca] pure viscosity ($\eta_1$) is 39.14 cP at 293.2 K[107]. The viscosity of aqueous-[Bmim][dca] IL 0.15 M at 293.2 K was calculated to 1.018 cP.

# References


1. Dyson, H. J. & Wright, P. E. Elucidation of the Protein Folding Landscape by NMR. in *Methods in Enzymology* **394**, 299–321 (2005).
2. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).
3. Marsh, J. A. *et al.* Improved Structural Characterizations of the drkN SH3 Domain Unfolded State Suggest a Compact Ensemble with Native-like and Non-native Structure. *J. Mol. Biol.* **367**, 1494–1510 (2007).
4. Uversky, V. N. Intrinsically Disordered Proteins and Their "Mysterious" (Meta)Physics. *Front. Phys.* **7**, 8–23 (2019).
5. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D. & Kay, L. E. Slow Dynamics in Folded and Unfolded States of an SH3 Domain. *J. Am. Chem. Soc.* **123**, 11341–11352 (2001).
6. Bezsonova, I., Singer, A., Choy, W.-Y., Tollinger, M. & Forman-Kay, J. D. Structural Comparison of the Unstable drkN SH3 Domain and a Stable Mutant † , ‡. *Biochemistry* **44**, 15550–15560 (2005).
7. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845–858 (1994).
8. Zhang, O. & Forman-Kay, J. D. Structural Characterization of Folded and Unfolded States of an SH3 Domain in Equilibrium in Aqueous Buffer. *Biochemistry* **34**, 6784–6794 (1995).
9. Zhang, O. & Forman-Kay, J. D. NMR Studies of Unfolded States of an SH3 Domain in Aqueous Solution and Denaturing Conditions †. *Biochemistry* **36**, 3959–3970 (1997).
10. Crowhurst, K. A., Tollinger, M. & Forman-Kay, J. D. Cooperative Interactions and a Non-native Buried Trp in the Unfolded State of an SH3 Domain. *J. Mol. Biol.* **322**, 163–178 (2002).
11. Choy, W.-Y. *et al.* Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J. Mol. Biol.* **316**, 101–112 (2002).
12. Marsh, J. A. & Forman-Kay, J. D. Structure and Disorder in an Unfolded State under Nondenaturing Conditions from Ensemble Models Consistent with a Large Number of Experimental Restraints. *J. Mol. Biol.* **391**, 359–374 (2009).
13. Mazouchi, A. *et al.* Conformations of a Metastable SH3 Domain Characterized by smFRET and an Excluded-Volume Polymer Model. *Biophys. J.* **110**, 1510–1522 (2016).
14. Lincoff, J. *et al.* Extended experimental inferential structure determination method in determining the structural ensembles of disordered protein states. *Commun. Chem.* **3**, 74 (2020).
15. Mok, Y.-K., Elisseeva, E. L., Davidson, A. R. & Forman-Kay, J. D. Dramatic stabilization of an SH3 domain by a single substitution: roles of the folded and unfolded states. *J. Mol. Biol.* **307**, 913–928 (2001).
16. Mok, Y.-K., Kay, C. M., Kay, L. E. & Forman-Kay, J. NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**, 619–638 (1999).
17. Crowhurst, K. A. & Forman-Kay, J. D. Aromatic and Methyl NOEs Highlight Hydrophobic Clustering in the Unfolded State of an SH3 Domain †. *Biochemistry* **42**, 8687–8695 (2003).
18. Lee, J. H. *et al.* Heterogeneous binding of the SH3 client protein to the DnaK molecular chaperone. *Proc. Natl. Acad. Sci.* **112**, E4206–E4215 (2015).
19. Sekhar, A. *et al.* Conserved conformational selection mechanism of Hsp70 chaperone-substrate interactions. *Elife* **7**, 1–29 (2018).
20. Smith, A. E., Zhou, L. Z., Gorensek, A. H., Senske, M. & Pielak, G. J. In-cell thermodynamics and a new role for protein surfaces. *Proc. Natl. Acad. Sci.* **113**, 1725–1730 (2016).
21. Stadmiller, S. S., Gorensek-Benitez, A. H., Guseman, A. J. & Pielak, G. J. Osmotic Shock Induced Protein Destabilization in Living Cells and Its Reversal by Glycine Betaine. *J. Mol. Biol.* **429**, 1155–1161 (2017).
22. Gorensek-Benitez, A. H., Smith, A. E., Stadmiller, S. S., Perez Goncalves, G. M. & Pielak, G. J. Cosolutes, Crowding, and Protein Folding Kinetics. *J. Phys. Chem. B* **121**, 6527–6537 (2017).
23. Rydeen, A. E., Brustad, E. M. & Pielak, G. J. Osmolytes and Protein–Protein Interactions. *J. Am. Chem. Soc.* **140**, 7441–7444 (2018).
24. Cheng, K., Wu, Q., Jiang, L., Liu, M. & Li, C. Protein stability analysis in ionic liquids by 19F NMR. *Anal. Bioanal. Chem.* **411**, 4929–4935 (2019).
25. Tanford, C. Protein Denaturation. in *Advances in protein chemistry* **23**, 121–282 (1968).
26. Greene, R. F. & Pace, C. N. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J. Biol. Chem.* **249**, 5388–5393 (1974).
27. Auton, M. & Bolen, D. W. Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc. Natl. Acad. Sci.* **102**, 15065–15068 (2005).
28. Saum, S. H. & Müller, V. Regulation of osmoadaptation in the moderate halophile Halobacillus halophilus: chloride, glutamate and switching osmolyte strategies. *Saline Systems* **4**, 4 (2008).
29. Figueiredo, A. M., Sardinha, J., Moore, G. R. & Cabrita, E. J. Protein destabilisation in ionic liquids: the role of preferential interactions in denaturation. *Phys. Chem. Chem. Phys.* **15**, 19632 (2013).

30. Silva, M., Figueiredo, A. M. & Cabrita, E. J. Epitope mapping of imidazolium cations in ionic liquid–protein interactions unveils the balance between hydrophobicity and electrostatics towards protein destabilisation. *Phys. Chem. Chem. Phys.* **16**, 23394–23403 (2014).

31. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).

32. Schumann, F. H. *et al.* Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. *J. Biomol. NMR* **39**, 275–289 (2007).

33. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. Characterization of the Backbone Dynamics of Folded and Denatured States of an SH3 Domain †. *Biochemistry* **36**, 2390–2402 (1997).

34. Hayes, R., Warr, G. G. & Atkin, R. Structure and Nanostructure in Ionic Liquids. *Chem. Rev.* **115**, 6357–6426 (2015).

35. Lim, W. A., Richards, F. M. & Fox, R. O. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375–379 (1994).

36. Constatinescu, D., Herrmann, C. & Weingärtner, H. Patterns of protein unfolding and protein aggregation in ionic liquids. *Phys. Chem. Chem. Phys.* **12**, 1756 (2010).

37. Gontrani, L. Choline-amino acid ionic liquids: past and recent achievements about the structure and properties of these really "green" chemicals. *Biophys. Rev.* **10**, 873–880 (2018).

38. Schröder, C. Proteins in Ionic Liquids: Current Status of Experiments and Simulations. *Top. Curr. Chem.* **375**, 25 (2017).

39. Bisht, M. & Venkatesu, P. Influence of cholinium-based ionic liquids on the structural stability and activity of $\alpha$-chymotrypsin. *New J. Chem.* **41**, 13902–13911 (2017).

40. Kushwah, N., Jain, V. & Yadav, D. Osmolytes: A Possible Therapeutic Molecule for Ameliorating the Neurodegeneration Caused by Protein Misfolding and Aggregation. *Biomolecules* **10**, 132 (2020).

41. Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* **5**, 593–599 (2009).

42. Nordwald, E. M. & Kaar, J. L. Mediating electrostatic binding of 1-butyl-3-methylimidazolium chloride to enzyme surfaces improves conformational stability. *J. Phys. Chem. B* **117**, 8977–86 (2013).

43. Garajová, K., Sedláková, D., Berta, M., Gazova, Z. & Sedlák, E. Destabilization effect of imidazolium cation-Hofmeister anion salts on cytochrome c. *Int. J. Biol. Macromol.* **164**, 3808–3813 (2020).

44. Toral, A. R. *et al.* Cross-linked Candida antarctica lipase B is active in denaturing ionic liquids. *Enzyme Microb. Technol.* **40**, 1095–1099 (2007).

45. Zhao, H. Protein stabilization and enzyme activation in ionic liquids: specific ion effects. *J. Chem. Technol. Biotechnol.* **91**, 25–50 (2016).

46. Muñoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Mol. Biol.* **1**, 399–409 (1994).

47. Muñoz, V. & Serrano, L. Elucidating the Folding Problem of Helical Peptides using Empirical Parameters. II†. Helix Macrodipole Effects and Rational Modification of the Helical Content of Natural Peptides. *J. Mol. Biol.* **245**, 275–296 (1995).

48. Lacroix, E., Viguera, A. R. & Serrano, L. Elucidating the folding problem of $\alpha$-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.* **284**, 173–191 (1998).

49. Choy, W. Y. & Forman-Kay, J. D. Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J. Mol. Biol.* **308**, 1011–1032 (2001).

50. Baxter, N. J. & Williamson, M. P. Temperature dependence of 1H chemical shifts in proteins. *J. Biomol. NMR* **9**, 359–369 (1997).

51. Cierpicki, T. & Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **21**, 249–261 (2001).

52. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. Comparison of the Backbone Dynamics of a Folded and an Unfolded SH3 Domain Existing in Equilibrium in Aqueous Buffer. *Biochemistry* **34**, 868–878 (1995).

53. Marsh, J. A., Singh, V. K., Jia, Z. & Forman-Kay, J. D. Sensitivity of secondary structure propensities to sequence differences between $\alpha$- and $\gamma$-synuclein: Implications for fibrillation. *Protein Sci.* **15**, 2795–2804 (2006).

54. Lumry, R. & Rajender, S. Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous properly of water. *Biopolymers* **9**, 1125–1227 (1970).

55. Becktel, W. J. & Schellman, J. A. Protein stability curves. *Biopolymers* **26**, 1859–1877 (1987).

56. Senske, M. *et al.* The temperature dependence of the Hofmeister series: thermodynamic fingerprints of cosolute–protein interactions. *Phys. Chem. Chem. Phys.* **18**, 29698–29708 (2016).

57. Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. (W. H. Freeman and Company, New York, 1999).

58. Privalov, P. L. & Makhatadze, G. I. Heat capacity of proteins. *J. Mol. Biol.* **213**, 385–391 (1990).

59. Senske, M., Smith, A. E. & Pielak, G. J. Protein Stability in Reverse Micelles. *Angew. Chemie Int. Ed.* **55**, 3586–3589 (2016).

60. Stadmiller, S. S. & Pielak, G. J. Enthalpic stabilization of an SH3 domain by D2O. *Protein Sci.* **27**, 1710–1716 (2018).

61. Privalov, P. L. & Khechinashvili, N. N. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol.* **86**, 665–684 (1974).

62. Myers, J. K., Nick Pace, C. & Martin Scholtz, J. Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148 (1995).

63. Oliveberg, M., Tan, Y. J. & Fersht, A. R. Negative activation enthalpies in the kinetics of protein folding. *Proc. Natl. Acad. Sci.* **92**, 8926–8929 (1995).

64. Ahlner, A., Carlsson, M., Jonsson, B.-H. H. & Lundström, P. PINT: A software for integration of peak volumes and extraction of relaxation rates. *J. Biomol. NMR* **56**, 191–202 (2013).

65. Monteith, W. B., Cohen, R. D., Smith, A. E., Guzman-Cisneros, E. & Pielak, G. J. Quinary structure modulates protein stability in cells. *Proc. Natl. Acad. Sci.* **112**, 1739–1742 (2015).

66. Danielsson, J. *et al.* Thermodynamics of protein destabilization in live cells. *Proc. Natl. Acad. Sci.* **112**, 12402–12407 (2015).

67. Senske, M. *et al.* Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc.* **136**, 9036–9041 (2014).

68. Sukenik, S., Sapir, L., Gilman-Politi, R. & Harries, D. Diversity in the mechanisms of cosolute action on biomolecular processes. *Faraday Discuss.* **160**, 225–237 (2013).

69. Oprzeska-Zingrebe, E. A. & Smiatek, J. Aqueous ionic liquids in comparison with standard co-solutes. *Biophys. Rev.* **10**, 809–824 (2018).

70. Zhou, H.-X., Rivas, G. & Minton, A. P. Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences. *Annu. Rev. Biophys.* **37**, 375–397 (2008).

71. Sapir, L. & Harries, D. Macromolecular Stabilization by Excluded Cosolutes: Mean Field Theory of Crowded Solutions. *J. Chem. Theory Comput.* **11**, 3478–3490 (2015).

72. Timasheff, S. N. The Control of Protein Stability and Association by Weak Interactions with Water: How Do Solvents Affect These Processes? *Annu. Rev. Biophys. Biomol. Struct.* **22**, 67–97 (1993).

73. Xie, G. & Timasheff, S. N. Temperature dependence of the preferential interactions of ribonuclease a in aqueous co-solvent systems: Thermodynamic analysis. *Protein Sci.* **6**, 222–232 (1997).

74. Street, T. O., Bolen, D. W. & Rose, G. D. A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci.* **103**, 13997–14002 (2006).

75. Auton, M., Rösgen, J., Sinev, M., Holthauzen, L. M. F. & Bolen, D. W. Osmolyte effects on protein stability and solubility: A balancing act between backbone and side-chains. *Biophys. Chem.* **159**, 90–99 (2011).

76. Politi, R. & Harries, D. Enthalpically driven peptide stabilization by protective osmolytes. *Chem. Commun.* **46**, 6449 (2010).

77. Benton, L. A., Smith, A. E., Young, G. B. & Pielak, G. J. Unexpected Effects of Macromolecular Crowding on Protein Stability. *Biochemistry* **51**, 9773–9775 (2012).

78. Tischer, A., Lilie, H., Auton, M. & Lange, C. Oxidative refolding of rPA in l-ArgHCl and in ionic liquids: A correlation between hydrophobicity, salt effects, and refolding yield. *Biopolymers* **101**, 1129–40 (2014).

79. Ben-Amotz, D. Water-Mediated Hydrophobic Interactions. *Annu. Rev. Phys. Chem.* **67**, 617–638 (2016).

80. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734 (1994).

81. Palmer, A. G., Kroenke, C. D. & Patrick Loria, J. Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules. in *Methods in Enzymology* **339**, 204–238 (Elsevier Masson SAS, 2001).

82. Gustafson, C. L. *et al.* A Slow Conformational Switch in the BMAL1 Transactivation Domain Modulates Circadian Rhythms. *Mol. Cell* **66**, 447-457.e7 (2017).

83. McConnell, H. M. Reaction Rates by Nuclear Magnetic Resonance. *J. Chem. Phys.* **28**, 430–431 (1958).

84. Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304 (1940).

85. Jacob, M., Schindler, T., Balbach, J. & Schmid, F. X. Diffusion control in an elementary protein folding reaction. *Proc. Natl. Acad. Sci.* **94**, 5622–5627 (1997).

86. Tollinger, M., Neale, C., Kay, L. E. & Forman-Kay, J. D. Characterization of the Hydrodynamic Properties of the Folding Transition State of an SH3 Domain by Magnetization Transfer NMR Spectroscopy †. *Biochemistry* **45**, 6434–6445 (2006).

87. Kestin, J., Sokolov, M. & Wakeham, W. A. Viscosity of liquid water in the range −8 °C to 150 °C. *J. Phys. Chem. Ref. Data* **7**, 941–948 (1978).

88. Perl, D. *et al.* Thermodynamics of a diffusional protein folding reaction. *Biophys. Chem.* **96**, 173–190 (2002).

89. Laidler, K. J. & King, M. C. The development of transition-state theory. *J. Phys. Chem.* **87**, 2657–2664 (1983).

90. Eyring, H. The Activated Complex and the Absolute Rate of Chemical Reactions. *Chem. Rev.* **17**, 65–77 (1935).

91. Evans, M. G. & Polanyi, M. Some applications of the transition state method to the calculation of reaction velocities, especially in solution. *Trans. Faraday Soc.* **31**, 875 (1935).

92. Hong, J. & Gierasch, L. M. Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc.* **132**, 10445–10452 (2010).

93. Makhatadze, G. I. & Privalov, P. L. Protein interactions with urea and guanidinium chloride. *J. Mol. Biol.* **226**, 491–505 (1992).

94. Prieto, J., Wilmans, M., Jiménez, M. A., Rico, M. & Serrano, L. Non-native local interactions in protein folding and stability: introducing a helical tendency in the all β-sheet α-spectrin SH3 domain. *J. Mol. Biol.* **268**, 760–778 (1997).

95. Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122–126 (1989).

96. Findeisen, M., Brand, T. & Berger, S. A 1H-NMR thermometer suitable for cryoprobes. *Magn. Reson. Chem.* **45**, 175–178 (2007).

97. Wishart, D. S. *et al*. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135–140 (1995).

98. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 687–692 (1995).

99. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Bioinforma.* **59**, 687–696 (2005).

100. Kay, L. E., Keifer, P. & Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665 (1992).

101. Ying, J., Delaglio, F., Torchia, D. A. & Bax, A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **68**, 101–118 (2017).

102. Skinner, S. P. *et al*. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* **66**, 111–124 (2016).

103. Zhang, H., Neal, S. & Wishart, D. S. RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR* **25**, 173–195 (2003).

104. Niklasson, M. *et al*. Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* **69**, 93–99 (2017).

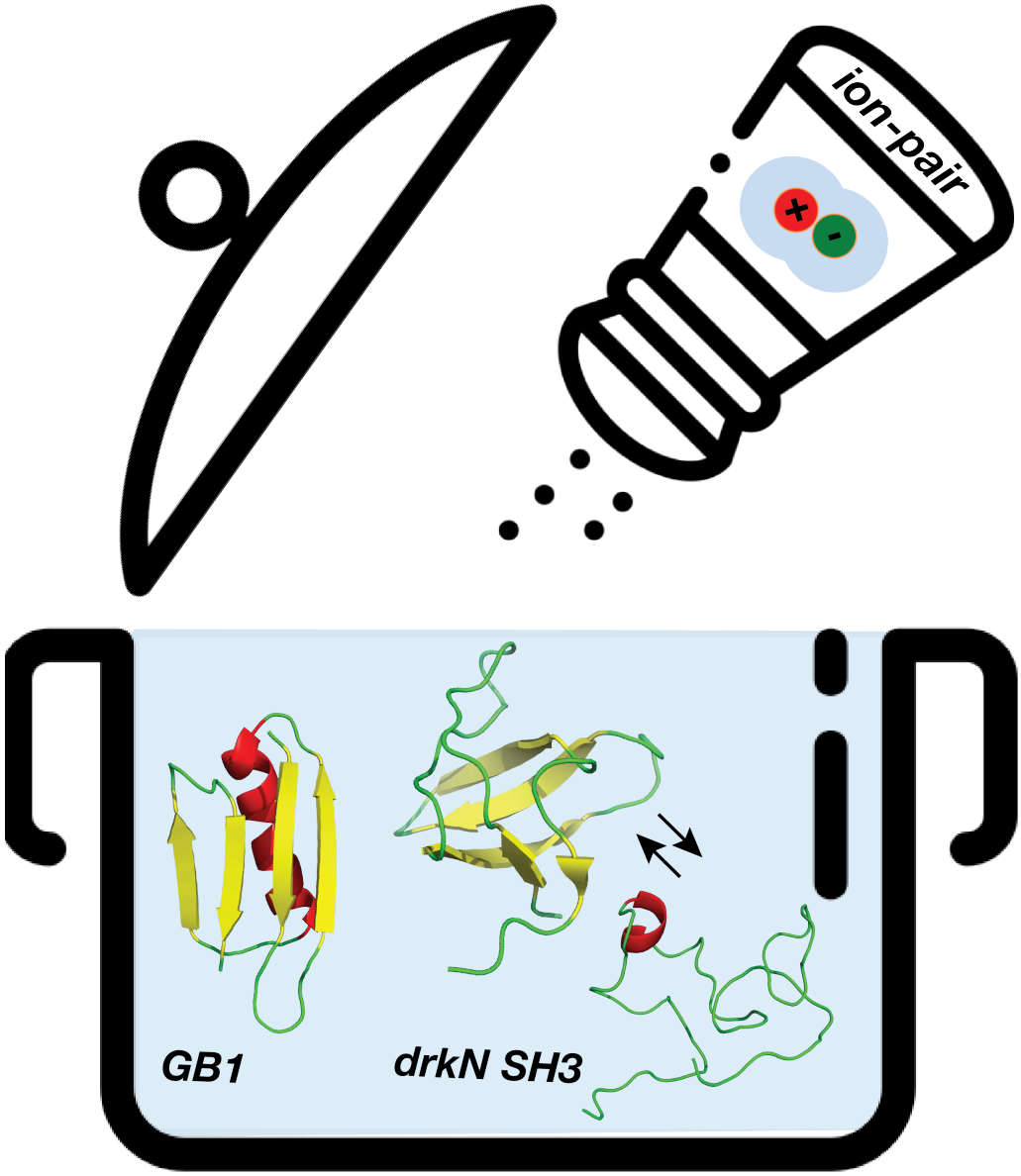
105. Atashrouz, S., Zarghampour, M., Abdolrahimi, S., Pazuki, G. & Nasernejad, B. Estimation of the viscosity of ionic liquids containing binary mixtures based on the Eyring's theory and a modified Gibbs energy model. *J. Chem. Eng. Data* **59**, 3691–3704 (2014).

106. López-Cortés, I. Y., Iglesias-Silva, G. A., Ramos-Estrada, M. & Rivera-Rojas, J. L. A correlation for the viscosity of binary mixtures of ionic liquids with organic solvents and water. *Fluid Phase Equilib.* **514**, 112543 (2020).

107. Carvalho, P. J., Regueira, T., Santos, L. M. N. B. F., Fernandez, J. & Coutinho, J. A. P. Effect of water on the viscosities and densities of 1-butyl-3- methylimidazolium dicyanamide and 1-butyl-3-methylimidazolium tricyanomethane at atmospheric pressure. *J. Chem. Eng. Data* **55**, 645–652 (2010).

4

# Chapter 4: Overall conclusions and Outlook



The work developed and presented in this thesis provided key insights into the interactions between ionic liquids and proteins. Here, I provide a summary of the main results from chapter 2 and 3 and discuss the possibility for a general mechanism (or universal hypothesis) of their effects on proteins. Taking in account the limited data in terms of the number of studied ILs and proteins, I hypothesize that the effect of ILs on protein stability is explained by indirect (and stabiliser) or preferential binding (and denaturing) mechanisms.

## Overview of protein (de)stabilisation in ILs

Considering the stabilisation of the negatively charged GB1 and drkN SH3 proteins, in terms of $T_m$ or $\Delta G_u^{0'}$, Figure 4.1 illustrates 1) the $\Delta T_m$ of GB1 in the presence of 0.75 M or 2.0 M of salts or ILs, which corresponds to the minimum (red squares) and maximum (blue circles) of $T_m$ measured for [Ch][Glu]; and 2) the *m*-value that corresponds to the $\Delta\Delta G_u^{0'}$ of SH3 at 1 M of IL/salt concentration. In general, the tested cosolutes can be ranked with the stability order [Ch][Glu] ≈ Na[Glu] > NaCl > [Ch]Cl > [Bmim]Cl > Na[dca] ≫ [Bmim]dca, where the effects from $[Glu]^-$ or $[dca]^-$ anions seem to prevail over cationic effects ($[Ch]^+$ or $[Bmim]^+$). [Ch][Glu] effect seems to result mainly from the interaction between protein and anion. On the other hand, both the $[Bmim]^+$ cation and the $[dca]^-$ anion appear to be equally destabilisers, with their combination and concomitant ion-pair resulting in a very strong denaturant.

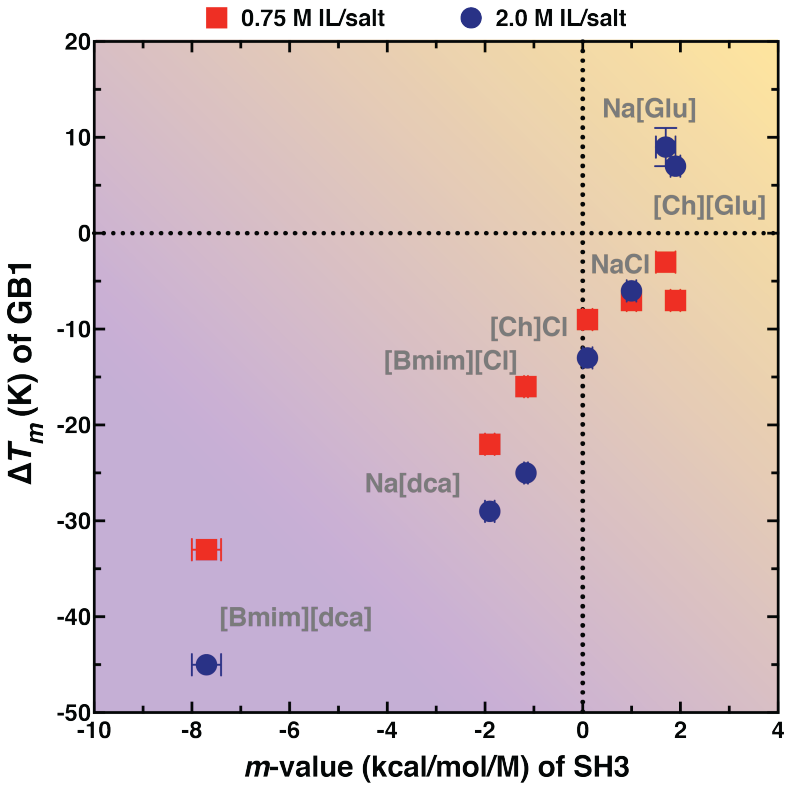



**Figure 4.1. Protein (de)stabilisation with ILs and salts.**

Y-axis, $\Delta T_m$ of GB1 at 0.75 (red squares) and 2.0 M (blue squares) of cosolute concentration using $T_m$ of water as reference, the errors were propagated from the fitting error for each condition. X-axis, *m*-values of drkN

SH3 as extracted from the slope of each titration at 298.2 K and the errors were derived from the linear fitting. The gradient of the background from purple to yellow colour scales with increasing stability for the two proteins.

Because drkN SH3 domain is thermodynamically unstable, it is much more sensible to the solvent environment than GB1. Low concentrations of cosolute are enough to produce large changes in stability and the ion-specific effects are too strong to trace a minima (folding-unfolding equilibrium is linearly dependent on IL or salt concentration). Likewise, GB1 is also monotonously destabilised by addition of [Bmim][dca], [Bmim][Cl] ILs and Na[dca], and [Ch][Cl] salts. However, the scenario is different for [Ch][Glu] and Na[Glu], where it is speculated that these ions and the protein establish electrostatic interactions (e.g., coulombic attractions) at 0.75 M of IL/salt concentration (destabilisation), while at high concentrations (2.0 M), ion-specific effects become dominant (stabilisation). Why [Ch][Glu] stabilises GB1 only at high IL concentration?

## Overview of ILs-protein interactions: a unified molecular mechanism?

A large amount of research data were acquired to discriminate, at molecular level, the [Ch][Glu] and [Bmim][dca] effects observed in the (de)stabilisation of GB1 and drkN SH3 proteins.

### [Ch][Glu] effects: an indirect mechanism?

The preferential hydration of the protein does not occur uniquely in the presence of a preferential exclusion behaviour at the protein surface (due to the increase of surface tension of water), as often observed for stabilising osmolytes (e.g., sucrose, sorbitol, trehalose)[1–4]. There is considerable evidence (as discussed along chapter 2 and 3), particularly for high concentrations, that [Ch][Glu] cosolute molecules are preferentially attracted or accumulated to specific hydrophilic surfaces on the native protein which is accompanied by a preferential hydration resultant from the strong kosmotropic and hygroscopic properties of [Ch][Glu] (mainly because of the strongly hydrated anion). While the stabilisation of the folded state is entropically driven, it does not exclude that the IL is more strongly excluded from the unfolded state than the folded state. This agrees with the observations that protectants, such as TMAO or hydroxyectoine, carry the hydration water in front of the solute surface when being attracted to a globular polymer, such as PNIPAM (poly(N-isopropylacrylamide))[5].

In addition, the fine balance between direct cosolute-protein and indirect water-mediated interactions may change with cosolute concentration (and/or temperature) and that could lead to preferential accumulation or preferential exclusion of the cosolute[6–11]. At concentrations greater than 1.0 M, GB1 is not only additionally stabilised but it is also, for example, more affected in rotation than in translation or more protected in terms of water exchange. These observations suggest a predominant transition, on the folded state, from preferential exclusion (hydration) to accumulation (Fig. 4.2a), with more protein adhering to the native state relatively to the unfolded ensemble.

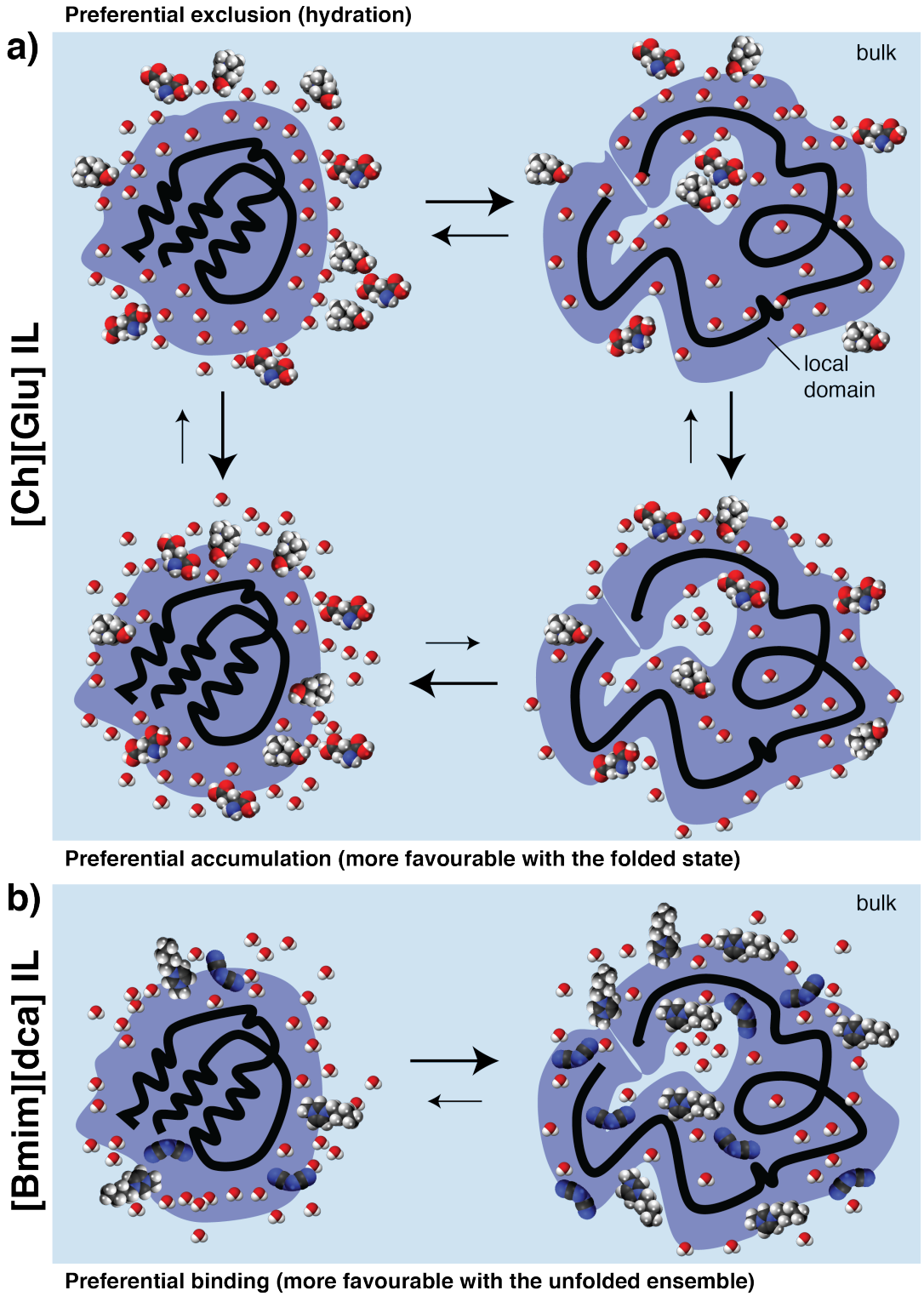



**Figure 4.2. Dominant mechanisms of ILs on protein stabilisation.**

Schematic representation of the effects of **a**) preferential exclusion-to-preferential accumulation of [Ch][Glu] as cosolute (more favourable with the folded state), and **b**) preferential binding of [Bmim][dca] as cosolute (more favourable with the unfolded ensemble), at the surface of a protein (local domain) and on its folding equilibrium. The quantity and the relative sizes of the molecules do not correspond to reality.

This switch could also be sensible to a change in temperature. It should be noted that effects of the protein-on-solvent cannot be disregarded[4], as the driving force for preferential hydration is the perturbation by the protein on the chemical potential of the cosolute. Noteworthy, this macroscopic mechanism is also concomitant with ion-specific-effects[12,13]. In particular, the unfavourable interaction observed for the larger $Glu^-$ anion[14] with the known melting hotspot of GB1 backbone amide (around Ala23 residue)[15] may change from an indirect and solvent-mediated (anion-water-protein interaction) for a more direct one, which does not mean a local dehydration but a reinforcement of that interaction. This change is expected to be determinant for GB1 stability. For drkN SH3, [Ch][Glu] is preferential accumulated in the folded state at low and high concentration but without significant preferential interactions (Figure 4.2a - bottom). This leads to an entropic stabilisation of the native state (excluded-volume effects) by significantly slowing the rate of protein unfolding.

Since it is generally accepted that the bulk water structure is not greatly affected by ions, the kosmotropicity concept renaissances on the effect of ion impact on protein hydration, and direct ion-protein interactions[16–18]. In this sense, kosmotropicity of the ions of a given IL could partially explain the increase of protein stability[12,19]. Accordingly to Nikawa et al[20], if an IL is prepared with a small amount of water, by the molar ratio of IL: $H_2O$ = 1:3 (i.e., seven water molecules per ion pair[21]), there is no free water and all the water molecules are in the form of hydration water, which interact strongly with the ions. The formation of the hydrogen bonding network between the [Ch][Glu] ions and water molecules was confirmed by the downfield shift of the chemical shift of $H_2O$ relative to bulk water (Figure 4.3). In fact, the trend of this shift was already observed for titrations up to 1.5 M of [Ch][Glu] (chapter 2) and it is similar to those cholinium based-ILs with typical kosmotropic anions, such as phosphate or citrate[20]. This confirms the anion's ability to facilitate the structuring of nearby water molecules and support the idea that preferential accumulation (via water-mediated interactions) can be considered as the fundamental mechanism for the onset of preferential hydration (here, the cause of exclusion is the solvophobic effect) which, indeed, stabilises the protein.

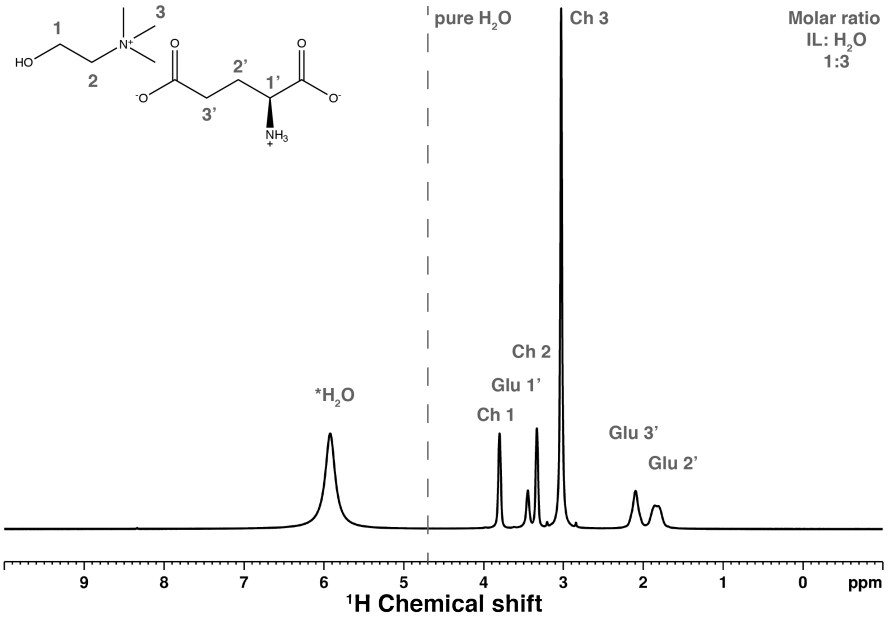



**Figure 4.3. Downfield shift of the water signal in [Ch][Glu] IL.**

$^1$H NMR spectra of ≈ 3.8 M of hydrated [Ch][Glu] with the molar ratio of IL: $H_2O$ = 1:3 (at 318.2 K and 400.15 MHz). Under these conditions, there is no "free" water molecules, and the ions are strongly hydrated. The asterisk (*) indicates that the signal of $H_2O$ is merged with that of exchangeable protons in the IL. The dotted vertical line represents the signal of pure $H_2O$ (4.7 ppm).

**[Bmim][dca] effects: preferential binding mechanism?**

In general, destabiliser cosolutes are known to favourably interact with the protein surface and to have a preferential binding with the unfolded ensemble. For example, urea accumulates in the vicinity of the protein due to favourable direct and chemical interactions (e.g., hydrogen bonding) with the protein backbone and most of the side chains groups. Because the unfolded ensemble has a larger preferential interaction with urea than the folded ensemble, an enthalpic driving force for unfolding is provided[22]. In a different thermodynamic mechanism, hydrophobic ILs that impose strong denaturation (e.g., imidazolium-based ILs) establish unfavourable and hydrophobic interactions with non-polar residues of the protein by their ions and/or ion-pair[23–28]. Here, the origin of such denaturing effects are direct anion-protein interactions which can be partially overcompensated by combination with a suitable cation (through cation–anion pair and cation–protein interactions). The ability of weakly hydrated anions, such as $[dca]^-$ with high H-bond basicity[29], to bind (and to dehydrate) positively charged or polar residues at the protein surface could ultimately lead to protein unfolding[25]. Yet their mechanism of action is not fully understood; is it the folded state that is destabilised, or the unfolded state that is stabilised?

In agreement with the aforementioned studies, I found that [Bmim][dca] denatures GB1 and drkN SH3 by preferential hydrophobic interactions not only with the folded state but also with the unfolded one (Figure 4.2b). Surprisingly, in the case of SH3, this leads to a stabilisation of a non-native $\alpha$-helical structure (previously observed[30–33]) in the unfolded ensemble in opposition to the random coil-like found in urea or GdmCl salt, thus shifting the

equilibrium to the unfolded state. A cooperative effect of the direct binding of $[dca]^-$ anion with Lys26 and hydrophobic contacts of $[Bmim]^+$ and/or ion-pair seems to stabilise a particular hydrophobic patch (with concomitant increase of helical propensity) within the unfolded ensemble. The preferential binding or accumulation of [Bmim][dca] ions in the vicinity of the backbone of a particular unfolded conformation significantly slow the rate of folding. Although less evident for GB1, I also found a residual $\alpha$-helical structure in the denatured state of GB1 under IL which it is indicative of a stabilisation of a non-native hydrophobic staple motif within the unfolded ensemble[34,35].

However, the [Bmim][dca]-protein contacts are unfavourable with the protein backbone (even if the interaction with the side chains groups is favourable[36]) as suggested by the thermodynamic fingerprint of the SH3 unfolding transition (entropic destabilisation offset by enthalpic stabilisation contribution). The hydration of the hydrophobic groups of $[Bmim]^+$ reduces water entropy which diminishes the loss of entropy due to the solvation of exposed hydrophobic residues of the protein[37]. The IL denaturing effect is not only explained by a direct mechanism (as in the case of urea and GdmCl) , but also by water-mediated hydrophobic interactions[38]. This is in line with the contacts observed with the hydrophobic core of the well-folded GB1.

**Final conclusions**

In this dissertation, the identification of a stabiliser [Ch][Glu] IL based on charged metabolites, as well as the study of the effects of a denaturant [Bmim][dca] IL on two charged proteins with distinct stability, led to a thorough understanding of the IL-protein interactions and their effect on the equilibrium thermodynamics. Cholinium-based ILs with similar anions to glutamate anion (e.g., acetate, aspartate, citrate, dihydrogen phosphate) should improve protein stability, independently of the protein, as the major effects are due to the combination of preferential hydration with excluded volume effects (entropic-driven stabilisation mechanism). Imidazolium-based ILs with an anion as dicyanamide (or thiocyanate) is a powerful denaturant and have the potential to be used instead of urea or GdmCl in common protein purification methods in the laboratory. The extension of the cation alkyl chain, which it increases hydrophobic contacts with the protein, should lead to higher denaturation.

Additionally, the NMR methodologies used represent a robust and effective way to study the elusive unfolded and disordered states of proteins.

Together, the key results of this thesis highlight the mechanism of action of charged metabolites that are able to form ILs on the modulation of protein stability *in vitro*, helping to rationalise the Hofmeister series and contributing to a universal interaction mechanism of salt ions, osmolytes, or ionic liquids. This have a directly impact in the rational design of solvents in industry which could improve biotechnological processes increasing not only catalytic activity but also the protein thermostability in such media.

The relevance *in vivo* remains to be studied.

**Future perspectives**

As the next step, the existence of non-native and residual structure in the unfolded ensemble state, as suggested in the chapter 2 and 3 for GB1 and drkN SH3 under denaturant-IL conditions, could be validated using sequential amide-amide NOEs data. The fact that residual structure exists in the unfolded state will reduce the conformational search in the early stages of protein folding, a crucial process to understand misfolding of proteins as a harbinger of disease. Moreover, as the kinetics of protein folding occurs over an extensive range of timescales (chapter 1) and the CMPG relaxation dispersion as well as the CEST method are sensible to intermediate-to-fast timescale exchange processes, the characterization of conformational ensembles with a very small populations of substates can be performed[39,40]. It could be interesting to carry out such experiments, especially for the drkN SH3, at selected salt and IL conditions.

To complement IL-protein structural interactions studies, molecular mechanics calculations and molecular dynamics (MD) simulations can also be very useful[41,42]. In particular, on the analysis of the distribution of anions and cations around GB1 or drkN SH3 surface, which may provide further insights not only in the ion solvation but also in their repercussions on the water of hydration.

Since all the biophysical studies performed with GB1 and drkN SH3 are just concerning the monomeric form, the IL effects on protein aggregation, for example, on αSynuclein (αSyn), remains to be investigated. Therefore, the monitorization of protein amyloid fibrillation in the presence of ILs is central and it only depends on the optimization of experimental conditions in fluorescence assays. Can [Ch][Glu] (or [Bmim][dca]) ionic liquid suppress or promote a toxic off-pathway of aggregation? Can it dissolve mature fibrils back to native structures?

# References


1. Timasheff, S. N. The Control of Protein Stability and Association by Weak Interactions with Water: How Do Solvents Affect These Processes? *Annu. Rev. Biophys. Biomol. Struct.* **22**, 67–97 (1993).
2. Xie, G. & Timasheff, S. N. The thermodynamic mechanism of protein stabilization by trehalose. *Biophys. Chem.* **64**, 25–43 (1997).
3. Xie, G. & Timasheff, S. N. Mechanism of the stabilization of ribonuclease a by sorbitol: Preferential hydration is greater for the denatured than for the native protein. *Protein Sci.* **6**, 211–221 (1997).
4. Timasheff, S. N. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci.* **99**, 9721–9726 (2002).
5. Schroer, M. A., Michalowsky, J., Fischer, B., Smiatek, J. & Grübel, G. Stabilizing effect of TMAO on globular PNIPAM states: preferential attraction induces preferential hydration. *Phys. Chem. Chem. Phys.* **18**, 31459–31470 (2016).
6. Xie, G. & Timasheff, S. N. Temperature dependence of the preferential interactions of ribonuclease a in aqueous co-solvent systems: Thermodynamic analysis. *Protein Sci.* **6**, 222–232 (1997).
7. Zeeb, M., Jacob, M. H., Schindler, T. & Balbach, J. 15N relaxation study of the cold shock protein CspB at various solvent viscosities. *J. Biomol. NMR* **27**, 221–234 (2003).
8. Abriata, L. A., Spiga, E. & Peraro, M. D. Molecular Effects of Concentrated Solutes on Protein Hydration, Dynamics, and Electrostatics. *Biophys. J.* **111**, 743–755 (2016).
9. Barnett, G. V. *et al.* Osmolyte Effects on Monoclonal Antibody Stability and Concentration-Dependent Protein Interactions with Water and Common Osmolytes. *J. Phys. Chem. B* **120**, 3318–3330 (2016).
10. Gao, M. *et al.* Crowders and Cosolvents-Major Contributors to the Cellular Milieu and Efficient Means to Counteract Environmental Stresses. *ChemPhysChem* **18**, 2951–2972 (2017).
11. Gorensek-Benitez, A. H., Smith, A. E., Stadmiller, S. S., Perez Goncalves, G. M. & Pielak, G. J. Cosolutes, Crowding, and Protein Folding Kinetics. *J. Phys. Chem. B* **121**, 6527–6537 (2017).
12. Zhao, H. Protein stabilization and enzyme activation in ionic liquids: specific ion effects. *J. Chem. Technol. Biotechnol.* **91**, 25–50 (2016).
13. Okur, H. I. *et al.* Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **121**, 1997–2014 (2017).
14. Cheng, X. *et al.* Basis of Protein Stabilization by K Glutamate: Unfavorable Interactions with Carbon, Oxygen Groups. *Biophys. J.* **111**, 1854–1865 (2016).
15. Ding, K., Louis, J. M. & Gronenborn, A. M. Insights into Conformation and Dynamics of Protein GB1 During Folding and Unfolding by NMR. *J. Mol. Biol.* **335**, 1299–1307 (2004).
16. Zhang, Y. & Cremer, P. S. Interactions between macromolecules and ions: the Hofmeister series. *Curr. Opin. Chem. Biol.* **10**, 658–663 (2006).
17. Jungwirth, P. & Cremer, P. S. Beyond Hofmeister. *Nat. Chem.* **6**, 261–263 (2014).
18. Ball, P. & Hallsworth, J. E. Water structure and chaotropicity: their uses, abuses and biological implications. *Phys. Chem. Chem. Phys.* **17**, 8297–8305 (2015).
19. Fujita, K. *et al.* Solubility and Stability of Cytochrome c in Hydrated Ionic Liquids: Effect of Oxo Acid Residues and Kosmotropicity. *Biomacromolecules* **8**, 2080–2086 (2007).
20. Nikawa, Y., Fujita, K. & Ohno, H. Quantitative assessment of kosmotropicity of hydrated ionic liquids by nuclear magnetic resonance. *Phys. Chem. Chem. Phys.* **19**, 8148–8151 (2017).
21. Ohno, H., Fujita, K. & Kohno, Y. Is seven the minimum number of water molecules per ion pair for assured biological activity in ionic liquid–water mixtures? *Phys. Chem. Chem. Phys.* **17**, 14454–14460 (2015).
22. Canchi, D. R. & García, A. E. Cosolvent effects on protein stability. *Annu. Rev. Phys. Chem.* **64**, 273–293 (2013).
23. Constantinescu, D., Weingärtner, H. & Herrmann, C. Protein Denaturation by Ionic Liquids and the Hofmeister Series: A Case Study of Aqueous Solutions of Ribonuclease A. *Angew. Chemie Int. Ed.* **46**, 8887–8889 (2007).
24. Nordwald, E. M. & Kaar, J. L. Mediating electrostatic binding of 1-butyl-3-methylimidazolium chloride to enzyme surfaces improves conformational stability. *J. Phys. Chem. B* **117**, 8977–86 (2013).
25. Figueiredo, A. M., Sardinha, J., Moore, G. R. & Cabrita, E. J. Protein destabilisation in ionic liquids: the role of preferential interactions in denaturation. *Phys. Chem. Chem. Phys.* **15**, 19632 (2013).
26. Silva, M., Figueiredo, A. M. & Cabrita, E. J. Epitope mapping of imidazolium cations in ionic liquid–protein interactions unveils the balance between hydrophobicity and electrostatics towards protein destabilisation. *Phys. Chem. Chem. Phys.* **16**, 23394–23403 (2014).
27. Warner, L. *et al.* The effects of high concentrations of ionic liquid on GB1 protein structure and dynamics probed by high-resolution magic-angle-spinning NMR spectroscopy. *Biochem. Biophys. Reports* **8**, 75–80 (2016).
28. Cheng, K., Wu, Q., Jiang, L., Liu, M. & Li, C. Protein stability analysis in ionic liquids by 19F NMR. *Anal. Bioanal. Chem.* **411**, 4929–4935 (2019).
29. MacFarlane, D. R., Pringle, J. M., Johansson, K. M., Forsyth, S. A. & Forsyth, M. Lewis base ionic liquids. *Chem. Commun.* 1905 (2006). doi:10.1039/b516961p

30. Zhang, O. & Forman-Kay, J. D. NMR Studies of Unfolded States of an SH3 Domain in Aqueous Solution and Denaturing Conditions †. *Biochemistry* **36**, 3959–3970 (1997).

31. Mok, Y.-K., Kay, C. M., Kay, L. E. & Forman-Kay, J. NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**, 619–638 (1999).

32. Mok, Y.-K., Elisseeva, E. L., Davidson, A. R. & Forman-Kay, J. D. Dramatic stabilization of an SH3 domain by a single substitution: roles of the folded and unfolded states. *J. Mol. Biol.* **307**, 913–928 (2001).

33. Marsh, J. A. *et al.* Improved Structural Characterizations of the drkN SH3 Domain Unfolded State Suggest a Compact Ensemble with Native-like and Non-native Structure. *J. Mol. Biol.* **367**, 1494–1510 (2007).

34. Blanco, F. J., Ortiz, A. R. & Serrano, L. Role of a nonnative interaction in the folding of the protein G B1 domain as inferred from the conformational analysis of the α-helix fragment. *Fold. Des.* **2**, 123–133 (1997).

35. Cohen, R. D. & Pielak, G. J. Quinary interactions with an unfolded state ensemble. *Protein Sci.* **26**, 1698–1703 (2017).

36. Tischer, A., Lilie, H., Auton, M. & Lange, C. Oxidative refolding of rPA in l-ArgHCl and in ionic liquids: A correlation between hydrophobicity, salt effects, and refolding yield. *Biopolymers* **101**, 1129–40 (2014).

37. Senske, M. *et al.* The temperature dependence of the Hofmeister series: thermodynamic fingerprints of cosolute–protein interactions. *Phys. Chem. Chem. Phys.* **18**, 29698–29708 (2016).

38. Ben-Amotz, D. Water-Mediated Hydrophobic Interactions. *Annu. Rev. Phys. Chem.* **67**, 617–638 (2016).

39. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D. & Kay, L. E. Slow Dynamics in Folded and Unfolded States of an SH3 Domain. *J. Am. Chem. Soc.* **123**, 11341–11352 (2001).

40. Sekhar, A. *et al.* Conserved conformational selection mechanism of Hsp70 chaperone-substrate interactions. *Elife* **7**, 1–29 (2018).

41. Micaêlo, N. M. & Soares, C. M. Protein Structure and Dynamics in Ionic Liquids. Insights from Molecular Dynamics Simulation Studies. *J. Phys. Chem. B* **112**, 2566–2572 (2008).

42. Schröder, C. Proteins in Ionic Liquids: Current Status of Experiments and Simulations. *Top. Curr. Chem.* **375**, 25 (2017).

# Publications and conference communications during this dissertation

The core of my PhD research, divided in chapter 2 and 3, have been prepared in two manuscripts, respectively, and they are about to be submitted. It is also worth to mention the collaborations with André Faustino and Ivo Martins from Nuno Santo's Lab (IMM), as well as with Francisco Leisico and Teresa Santos-Silva from the XTAL – Macromolecular Crystallography Lab (FCT - NOVA). These resulted in two publications. The results from a third collaboration, with Filipe Sousa and Manuela Pereira from Biological Energy Transduction Group (FCUL - UL), have been prepared in a manuscript.

### Papers in scientific peer reviewed journals

5. **[2021] Silva, M. S.,** Viegas, A., Figueiredo, A. M., Cabrita, E. J. Molecular basis of protein (de)stabilisation in Ionic Liquids. *In preparation*.
4. **[2021]** Sousa, F. M., Refojo, P. N., Barreto, A., Pires, P., **Silva, M. S.,** Pereira, P., Carapeto, A. P., Robalo, T. T., Rodrigues, M. S., Pinho, M. G., Cabrita, E. J., Pereira, M. M. Unveiling the membrane bound dihydroorotate: quinone oxidoreductase (DHOQO) from Staphylococcus aureus. *In preparation*.
3. **[2021] Silva, M. S.,** Viegas, A., O'Toole P., Felix, S., Figueiredo, A. M., Cabrita, E. J. Shifting the equilibrium – the role of the unfolded state for protein stability in ionic liquids. *In preparation*.
2. **[2019]** Faustino, A.F., Barbosa, G.M., **Silva, M.**, Castanho, M.A.R.B., Da Poian, A. T., Cabrita, E.J., Santos, N.C., Almeida, F.C.L., Martins, I.C. Silva, M., Fast NMR method to probe solvent accessibility and disordered regions in proteins. *Scientific Reports*, **9,** 1647 (2019).
1. **[2018]** Leisico, F., Vieira, D., Figueiredo, T., **Silva, M.**, Cabrita, E., Sobral, R., Ludovice, A., Trincão, J., Romao, M., de Lencastre, H., Santos-Silva, T. First insights of peptidoglycan amidation in Gram-positive bacteria - the high-resolution crystal structure of *Staphylococcus aureus* glutamine amidotransferase GatD. *Scientific Reports*, **8,** 5313 (2018).


### Communications in scientific conferences

11. **Silva, M. S.**, Viegas, A., Figueiredo, A. M., Cabrita, E. J. Protein stability and ionic liquids: indirect versus preferential binding mechanism. Poster presented at X Biennial Meeting of the NMR Group of the Spanish Royal Society of Chemistry / VII Iberian NMR Meeting / IX Ibero-American NMR Meeting (online), April 2021.
10. **Silva, M. S.** Understanding the effects of ionic liquids on protein stability. **Oral presentation** at the 11th ITQB NOVA PhD Student's (online) Meeting, Oeiras, February 2021.
9. **Silva, M.** Molecular understanding of protein stability in ionic liquids. **Oral presentation** at Bioseminars organized by UCIBIO, FCT NOVA, July 2020.

8. **Silva, M.**, Viegas, A., O'Toole, P., Figueiredo, A. M., Cabrita, E. J. Choline glutamate modulates protein folding and stability: an NMR view. Poster and **oral communication** at 2020 CCPN Conference online, July 2020.

7. **Silva, M. S.**, O'Toole, P., Viegas, A., Figueiredo, A. M., Cabrita, E. J. Modulation of protein (de)stabilisation by ionic liquids: an NMR view. Poster presented at the 21th ISMAR - 15th EUROMAR joint conference, Berlin, August 2019.

6. **Silva, M. S.**, O'Toole, P., Viegas, A., Figueiredo, A. M., Cabrita, E. J. Protein (de)stabilisation in Ionic Liquids: a biophysical approach. **Oral communication** presented at the I NOVA Biophysica international conference, Lisbon, September 2019.

5. **Silva, M. S.**, Viegas, A., Figueiredo, A. M., Cabrita, E. J. Atomistic picture of charged metabolite – protein interaction with repercussions on protein stabilization. Poster presented at the 9th ITQB NOVA PhD Student's Meeting, Oeiras, November 2018.

4. **Silva, M. S.**, O'Toole, P., Viegas, A., Figueiredo, A. M., Cabrita, E. J. Tuning protein stability with ionic liquids: investigation of destabilizing vs stabilizing effects. Poster presented by Eurico J. Cabrita at the 28th ICMRBS - International Conference on Magnetic Resonance in Biological Systems, Dublin, August 2018.

3. **Silva, M. S.**, Viegas, A., Figueiredo, A. M., Cabrita, E. J. Discerning the effect of charged metabolite interactions on proteins. Poster presented at the VIII Ibero-American NMR Meeting, Lisbon, June 2018.

2. **Silva, M. S.**, Viegas, A., Figueiredo, A. M., Cabrita, E. J. Discriminating the effect of charged metabolite interactions on protein structure and stability. Poster presented at the 6th International Iberian Biophysics Congress and X Ibero-American Congress of Biophysics, Castellón, June 2018.

1. **Silva, M. S.**, Viegas, A., Figueiredo, A. M., Cabrita, E. J. Ionic Liquids as tunable modulators for protein stability. **Oral communication** presented at the 6th Portuguese Young Chemists Meeting, Setúbal, May 2018.

## Attendance to Scientific Meetings, Courses and Training visits

### Conferences

12. EUROMAR online meeting, December 2020

12. 2020 CCPN Conference online, July 2020

11. 10th ITQB NOVA PhD Student's Meeting. Oeiras, January 2020

10. 1st Meeting COST Action 18132 "Functional Glyconanomaterials for the development of diagnostics and targeted therapeutic probes", Caparica, November 2019

9. 21th ISMAR - 15th EUROMAR joint conference, Berlin, August 2019

8. I NOVA Biophysica international conference, Lisbon, September 2019

7. 9th ITQB NOVA PhD Student's Meeting, Oeiras, November 2018

6. VIII Ibero-American NMR Meeting, Lisbon, June 2018

5. 6th International Iberian Biophysics Congress and X Ibero-American Congress of Biophysics, Castellón, June 2018 (awarded with ISMAR travel stipend)

4. 6th Portuguese Young Chemists Meeting, Setúbal, May 2018

3. COST Action CM1306 WG3 workshop "Emerging Magnetic Resonance Techniques to study dynamic molecular machines", Setúbal, April 2018

2. 1st Biophysics party, Oeiras, July 2017 (organized by me and colleagues in the establishment of the Portuguese Young Biophysicists)

1. 8th ITQB NOVA PhD Student's Meeting, Oeiras, November 2017

**Courses and Training visits**

3. EMBO Practical Course – Structure, dynamics, and function of biomacromolecules, with main lectures of Dr. Stephan Grzesiek, Dr. Michael Nilges and Dr. Michael Sattler, Garching/ Munich, August 2019 (awarded with EMBO travel stipend)

2. "Data Processing Automation (Python)" workshop from NOVA Doctoral School. Lisbon, May 2018

1. "Research Skills Development" workshop from NOVA Doctoral School. Setúbal, May 2017

## Teaching and assistance: Co-supervision

During this dissertation, I co-supervised students in their master thesis and/ or their undergraduate projects:

Fredi Brigham, Structural studies of $\alpha$-Synuclein and N-acetylated $\alpha$-Synuclein under the influence of common metabolites, Last Year Project in Biochemistry B. Sc. Degree, 2020.

Ana L. Dias, *Estudos estruturais da α-Sinucleina sob a influencia de metabolitos comuns*, Last Year Project in Biochemistry B. Sc. Degree, 2019.

Philip O'Toole, Characterization of crowding effects on protein structure and functionality, Last Year Project in Biochemistry B. Sc. Degree, 2018.

Cláudia Afonso, Development of in-cell NMR methodologies, Bioorganic Chemistry M. Sc. Degree, 2017.

# Appendix A

**Table A.1. Melting temperatures (Tm) of GB1 in water and in the presence of buffer and/ or cosolute.**

| | 0.05 M Sodium Phosphate | | Water | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Concentration (M)** | **[Ch][Glu]** | **[Bmim][dca]** | **[Ch][Glu]** | **[Bmim][dca]** | **[ChCl]** | **Na[Glu]** | **NaCl** | **[Bmim]Cl** | **Na[dca]** |
| **0.00** | 70.8 ± 0.6* | | | | | 80 ± 1* | | | |
| **0.05** | 68.7 ± 0.1 | 68.8 ± 0.3 | 77.5 ± 0.3 | 75.8 ± 0.3 | 78.7 ± 0.5 | 76.4 ± 0.3 | 74.6 ± 0.2 | - | 78.3 ± 0.5 |
| **0.10** | 72.0 ± 0.1 | 66.7 ± 0.2 | 77.6 ± 0.2 | 73.7 ± 0.2 | 79.5 ± 0.4 | 77.6 ± 0.3 | 78.3 ± 0.4 | 77.8 ± 0.7 | 61.0 ± 0.1 |
| **0.25** | 70.1 ± 0.1 | 61.1 ± 0.1 | 75.5 ± 0.2 | 65.0 ± 0.2 | 75.3 ± 0.2 | 76.1 ± 0.2 | 72.2 ± 0.2 | 72.3 ± 0.6 | 70.0 ± 0.3 |
| **0.50** | 68.6 ± 0.1 | 52.0 ± 0.1 | 73.7 ± 0.2 | 55.6 ± 0.2 | 72.9 ± 0.2 | 75.4 ± 0.2 | 70.9 ± 0.1 | 67.9 ± 0.3 | 64.9 ± 0.5 |
| **0.75** | 71.9 ± 0.1 | 46.2 ± 0.1 | 73.5 ± 0.2 | 46.7 ± 0.2 | 70.7 ± 0.1 | 77.0 ± 0.3 | 73.1 ± 0.2 | 64.6 ± 0.2 | 58.5 ± 0.3 |
| **1.00** | 73.4 ± 0.2 | 39.2 ± 0.2 | 74.6 ± 0.2 | 40.5 ± 0.2 | 69.4 ± 0.1 | 78.6 ± 0.3 | 73.3 ± 0.2 | 60.1 ± 0.2 | 54.85 ± 0.09 |
| **1.25** | 73.9 ± 0.2 | 36.5 ± 0.1 | 76.1 ± 0.2 | 37.3 ± 0.1 | 68.66 ± 0.09 | 79.7 ± 0.4 | 74.2 ± 0.2 | 57.8 ± 0.3 | 53.5 ± 0.3 |
| **1.50** | 77.3 ± 0.2 | 33.1 ± 0.1 | 79.1 ± 0.4 | 33.7 ± 0.1 | 68.1 ± 0.1 | 77.6 ± 0.1 | 74.4 ± 0.2 | 55.5 ± 0.2 | 51.1 ± 0.2 |
| **2.00** | 81.6 ± 0.6 | 29.12 ± 0.09 | 87.4 ± 0.8 | 35.5 ± 0.2 | 66.62 ± 0.09 | 89 ± 2 | 74.6 ± 0.2 | 54.8 ± 0.1 | 50.77 ± 0.04 |
| **2.50** | - | 25.6 ± 0.1 | - | 26.7 ± 0.1 | - | - | - | 51.6 ± 0.1 | 44.2 ± 0.1 |

A series of cosolute-solutions with varying IL ([Ch][Glu], [Bmim][dca], [Bmim]Cl) or salt ([Ch]Cl, Na[Glu], Na[dca], and NaCl) concentration from 0 to 2.0 M or 2.5 M were prepared from each solution stock in *milliQ* water to titrate GB1 at 52 μM of concentration. Additionally, [Ch][Glu] and [Bmim][dca] solutions were also prepared in buffer 0.05 M sodium phosphate pH 7.25. $T_m$ values were determined by fitting the intensity decay to a sigmoidal function using Prism 8 and the standard error was obtained from the fitting. *Values and uncertainties estimated from a triplicate measurement.

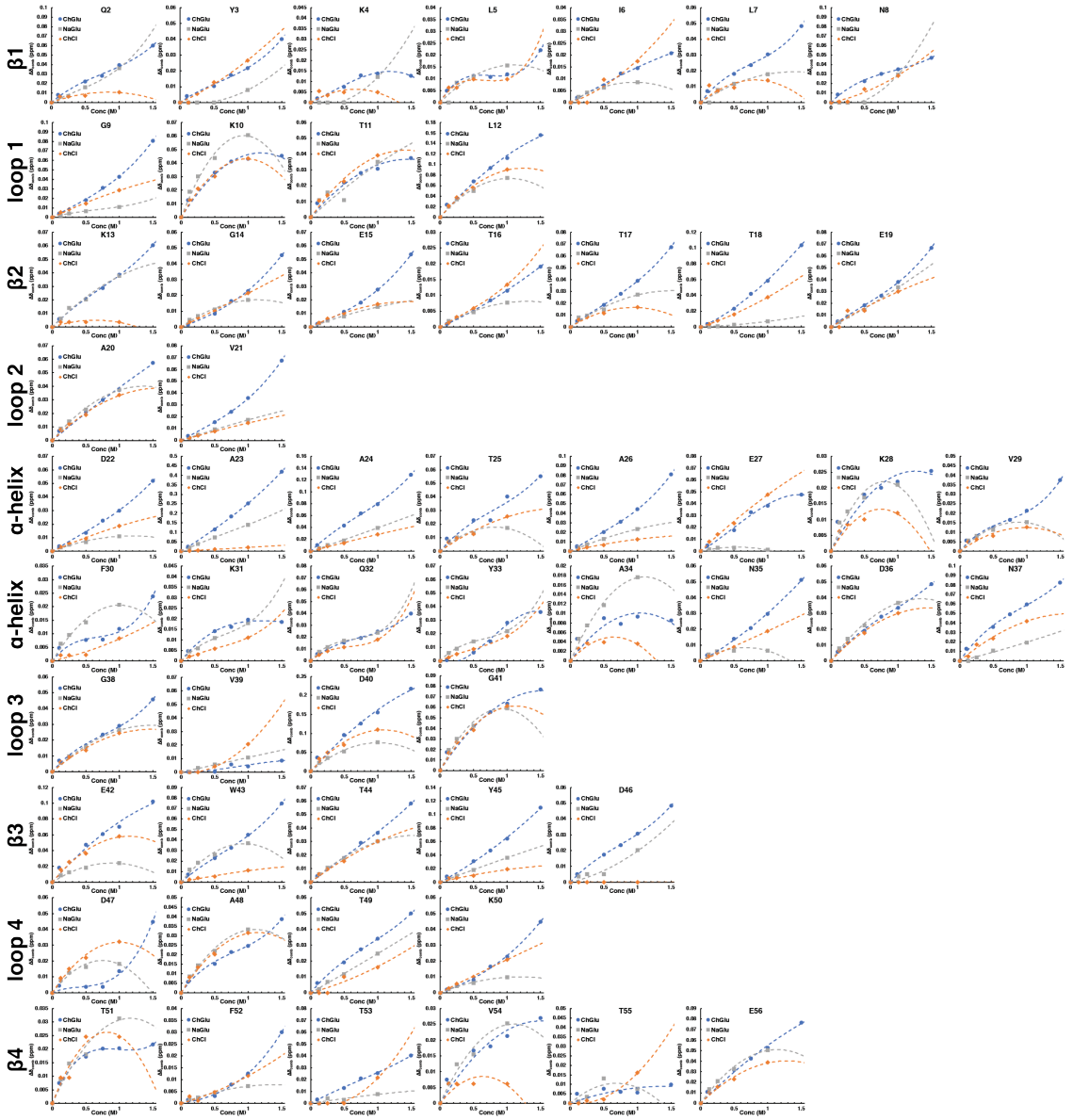



**Figure A.1. Effect of concentration of [Ch][Glu] and related salts on backbone amides of GB1.**

Plot of amide $\Delta\delta_{comb}$ as function of [Ch][Glu] (blue), Na[Glu] (gray) and [Ch]Cl (orange) concentration for all residues of GB1. The dashed lines are indicative only (polynomial with two variables). The combined chemical shifts were calculated against the chemical shifts in water.

**Table A.2. Backbone assignment of GB1 for the folded and unfolded states in ILs.**

| | Folded state, 1.5 M [Ch][Glu] IL | | | | | Unfolded state, 3.5 M [Bmim][dca] IL | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | δCα (± 0.03 ppm) | δCβ (± 0.03 ppm) | δCO (± 0.01 ppm) | $\delta H_N$ (± 0.002 ppm) | $\delta N_H$ (± 0.03 ppm) | δCα (± 0.02 ppm) | δCβ (± 0.02 ppm) | δCO (± 0.01 ppm) | $\delta H_N$ (± 0.003 ppm) | $\delta N_H$ (± 0.02 ppm) |
| **M1** | 54.72 | 32.74 | 171.11 | - | - | - | - | - | - | - |
| **Q2** | 56.04 | 30.62 | 174.83 | 8.422 | 123.87 | 56.33 | 29.96 | 175.80 | - | - |
| **Y3** | 57.24 | 43.49 | 174.89 | 9.081 | 124.51 | 58.62 | 38.69 | 175.62 | 8.204 | 119.54 |
| **K4** | 55.22 | 35.94 | 172.87 | 9.107 | 122.45 | 57.04 | 33.47 | 176.24 | 8.058 | 120.01 |
| **L5** | 52.69 | - | 174.60 | 8.638 | 126.68 | 55.84 | 42.46 | 177.18 | 8.028 | 120.50 |
| **I6** | 60.33 | 38.12 | 174.80 | 9.109 | 126.27 | 61.75 | 38.80 | 176.31 | 7.842 | 118.92 |
| **L7** | 54.50 | 42.09 | 174.82 | 8.765 | 125.81 | 55.54 | 42.18 | 176.80 | 8.119 | 122.89 |
| **N8** | 51.08 | 38.08 | 175.71 | 8.915 | 126.42 | 53.55 | 39.33 | 175.64 | 8.150 | 117.39 |
| **G9** | 44.80 | - | 173.44 | 7.906 | 109.37 | 45.98 | - | 173.99 | 8.297 | 107.90 |
| **K10** | 59.09 | 32.64 | 178.84 | 9.517 | 121.10 | 56.69 | 33.35 | 176.57 | 8.117 | 119.59 |
| **T11** | 62.03 | 69.71 | 173.82 | 8.824 | 108.41 | 62.19 | 70.07 | 174.80 | 8.055 | 112.90 |
| **L12** | 55.23 | 43.59 | 173.31 | 7.292 | 124.85 | 55.68 | 42.33 | 177.10 | 8.227 | 122.44 |
| **K13** | 53.86 | 35.04 | 176.31 | 8.171 | 124.14 | 56.71 | 33.25 | 176.74 | 8.130 | 119.16 |
| **G14** | 45.26 | - | 171.40 | 8.433 | 109.22 | 45.54 | - | 173.87 | 8.266 | 108.34 |
| **E15** | 54.73 | 34.02 | 175.20 | 8.474 | 118.98 | 56.65 | 30.84 | 176.58 | 8.315 | 119.81 |
| **T16** | 60.76 | 69.34 | 172.04 | 8.768 | 115.81 | 61.99 | 69.98 | 174.73 | 8.252 | 113.31 |
| **T17** | 60.04 | 73.43 | 174.04 | 8.132 | 112.23 | 62.13 | 69.84 | 174.77 | 8.176 | 114.66 |
| **T18** | 62.43 | 70.14 | 170.78 | 8.940 | 114.11 | 62.47 | 69.64 | 174.61 | 8.102 | 114.25 |
| **E19** | 54.58 | 30.44 | 175.84 | 8.019 | 126.26 | 57.15 | 30.60 | 176.27 | 8.408 | 121.79 |
| **A20** | 50.87 | 23.65 | 177.56 | 9.331 | 127.25 | 52.94 | 19.62 | 177.72 | 8.227 | 122.97 |
| **V21** | 63.52 | 32.08 | 174.90 | 8.570 | 115.84 | 62.78 | 32.94 | 175.78 | 7.966 | 117.25 |
| **D22** | 52.68 | 42.34 | 175.06 | 7.321 | 114.97 | 54.52 | 41.84 | 176.36 | 8.274 | 122.50 |
| **A23** | 54.60 | 17.62 | 179.44 | 8.725 | 121.97 | 54.08 | 19.24 | 178.40 | 8.351 | 123.56 |
| **A24** | 54.84 | 17.94 | 180.71 | 8.184 | 120.91 | 53.91 | 19.00 | 178.98 | 8.331 | 120.78 |
| **T25** | 66.90 | 67.82 | 176.15 | 8.320 | 116.75 | 64.33 | 69.23 | 175.31 | 8.033 | 113.46 |
| **A26** | 54.94 | 17.53 | 177.16 | 7.239 | 123.71 | 54.46 | 18.84 | 178.47 | 8.269 | 123.65 |
| **E27** | 59.73 | 29.35 | 177.33 | 8.291 | 116.69 | 58.55 | 29.98 | 177.40 | 8.228 | 117.67 |
| **K28** | 59.80 | 32.22 | 179.92 | 6.959 | 116.85 | 58.17 | 32.88 | 177.79 | 7.905 | 119.08 |
| **V29** | 66.06 | 31.82 | 179.80 | 7.339 | 120.56 | 64.73 | 32.46 | 177.01 | 8.043 | 118.28 |
| **F30** | 56.62 | 37.40 | 178.34 | 8.529 | 120.66 | 59.64 | 39.13 | 176.66 | 8.196 | 119.51 |
| **K31** | 60.01 | 31.67 | 179.58 | 9.108 | 123.12 | 57.92 | 32.82 | 177.18 | 8.102 | 119.31 |
| **Q32** | 58.83 | 28.20 | 177.13 | 7.470 | 119.66 | 57.23 | 29.23 | 176.43 | 8.139 | 118.40 |
| **Y33** | 61.93 | 38.65 | 179.01 | 8.234 | 121.04 | 58.91 | 38.81 | 175.92 | 7.994 | 118.23 |
| **A34** | 56.38 | 18.01 | 179.37 | 9.212 | 122.67 | 53.02 | 19.42 | 177.23 | 8.105 | 122.21 |
| **N35** | 56.97 | 38.89 | 179.37 | 8.313 | 117.85 | 53.50 | 39.40 | 174.93 | 8.137 | 116.48 |
| **D36** | 57.06 | 40.21 | 177.08 | 8.923 | 121.37 | 54.58 | 41.67 | 175.95 | 8.211 | 120.16 |
| **N37** | 53.86 | 40.15 | 174.05 | 7.374 | 115.05 | 53.39 | 39.30 | 175.60 | 8.350 | 118.30 |
| **G38** | 46.96 | - | 173.93 | 7.834 | 108.00 | 45.68 | - | 174.16 | 8.444 | 107.98 |
| **V39** | 61.92 | 33.09 | 174.21 | 8.103 | 121.07 | 62.40 | 32.73 | 175.65 | 8.039 | 117.95 |
| **D40** | 52.76 | 42.77 | 174.53 | 8.818 | 128.56 | 54.59 | 41.80 | 176.54 | 8.427 | 122.65 |
| **G41** | 45.29 | - | 171.88 | 7.686 | 107.13 | 45.52 | - | 173.82 | 8.220 | 108.11 |
| **E42** | 55.52 | 31.77 | 177.01 | 8.264 | 120.47 | 56.64 | 30.70 | 176.15 | 8.327 | 119.67 |
| **W43** | 58.02 | 30.46 | 176.93 | 9.420 | 128.85 | 57.16 | 29.76 | 175.71 | 8.343 | 121.17 |
| **T44** | 60.58 | 72.26 | 172.76 | 9.267 | 114.62 | 61.32 | 70.55 | 174.12 | 8.020 | 113.46 |
| **Y45** | 57.00 | 41.52 | 173.05 | 8.665 | 121.24 | 58.68 | 38.80 | 175.38 | 8.235 | 121.03 |
| **D46** | 51.86 | 43.22 | 174.57 | 7.674 | 128.47 | 54.52 | 42.05 | 175.85 | 8.243 | 120.97 |
| **D47** | 56.43 | 42.20 | 178.07 | 8.642 | 125.17 | 54.92 | 41.75 | 176.74 | 8.233 | 120.60 |
| **A48** | 55.04 | 18.48 | 179.68 | 8.332 | 119.75 | 53.74 | 19.26 | 178.26 | 8.411 | 123.17 |
| **T49** | 60.52 | 70.16 | 175.18 | 7.004 | 103.05 | 63.02 | 69.58 | 175.04 | 8.112 | 109.61 |
| **K50** | 56.40 | 29.29 | 175.07 | 7.852 | 122.97 | 56.58 | 32.69 | 176.16 | 7.911 | 121.06 |
| **T51** | 62.38 | 72.06 | 174.83 | 7.399 | 111.18 | 62.19 | 70.09 | 174.11 | 7.813 | 112.30 |
| **F52** | 57.16 | 42.69 | 174.61 | 10.380 | 130.84 | 57.98 | 39.78 | 175.20 | 8.209 | 120.27 |
| **T53** | 61.68 | 71.02 | 172.60 | 9.091 | 117.20 | 61.47 | 70.39 | 174.06 | 8.016 | 113.54 |
| **V54** | 58.06 | 32.37 | 173.20 | 8.192 | 123.51 | 62.37 | 32.75 | 175.76 | 8.143 | 120.03 |
| **T55** | 61.43 | 70.88 | 174.08 | 8.372 | 123.85 | 61.55 | 70.08 | 173.41 | 8.020 | 114.95 |
| **E56** | 58.23 | 32.69 | 180.30 | 7.824 | 133.64 | 58.24 | 31.51 | 180.34 | 7.883 | 126.69 |
| **W43sc** | - | - | | 10.517 | 130.86 | - | - | | 10.244 | 129.66 |

The backbone and aliphatic $H^N$, $N^H$, C$\alpha$, C$\beta$ and CO chemical shifts are shown for GB1 folded and unfolded states in 1.5 M [Ch][Glu] ($F_{[Ch][Glu]}$) or 3.5 M [Bmim][dca] ($U_{[Bmim][dca]}$), respectively. Spectra were acquired at 298.2 K, 600.13 MHz. For the assignment, 2D $^{1}H$-$^{15}N$ HSQC and 3D triple resonance experiments (HNCACB, HNcoCACB, HNCO, HNcaCO spectra) were used. The errors are the average of the standard deviation per residue.

**Table A.3. Backbone amide proton exchange ($k_{ex}^{obs}$, $s^{-1}$) for GB1 in water and [Ch][Glu].**

| | Too slow / Too fast | | | | Slow enough to quantify / Fast enough to quantify | | | | Superposition of peaks / Ambiguous from fitting | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | water | | | | 0.1 M [Ch][Glu] | | | | 0.75 M [Ch][Glu] | | | | 1.0 M [Ch][Glu] | | | | 1.5 M [Ch][Glu] | | | |
| | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | |
| | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE |
| M1 | | | | | | | | | | | | | | | | | | | | |
| Q2 | | | | | | | | | | | | | | | | | | | | |
| Y3 | | | 4.2E-05 | 1.1E-06 | | | 1.7E-04 | 1.2E-06 | 3.2E-05 | 1.5E-06 | 9.0E-04 | 1.3E-05 | 3.0E-05 | 2.3E-06 | 1.0E-03 | 2.2E-05 | 2.5E-05 | 3.2E-06 | 3.6E-04 | 1.1E-05 |
| K4 | | | 3.4E-05 | 1.2E-06 | | | 9.8E-05 | 1.6E-06 | | | 4.8E-04 | 5.1E-06 | | | 5.6E-04 | 1.0E-05 | | | 1.8E-04 | 5.9E-06 |
| L5 | | | 2.0E-05 | 1.9E-06 | | | 4.9E-05 | 5.4E-07 | 5.5E-06 | 3.0E-06 | 2.7E-04 | 2.3E-06 | 6.7E-06 | 4.9E-06 | 2.9E-04 | 4.0E-06 | | | 9.2E-05 | 3.7E-06 |
| I6 | | | 1.7E-05 | 3.8E-06 | | | 2.8E-05 | 1.5E-06 | | | 1.1E-04 | 2.0E-06 | | | 1.3E-04 | 2.9E-06 | | | 4.1E-05 | 6.0E-06 |
| L7 | | | 5.9E-05 | 9.4E-07 | 4.9E-06 | 1.2E-06 | 1.4E-04 | 8.3E-07 | 8.9E-05 | 9.9E-07 | 6.4E-04 | 4.4E-06 | 1.3E-04 | 1.5E-06 | 8.0E-04 | 1.0E-05 | 1.5E-04 | 2.7E-06 | 4.8E-04 | 9.5E-06 |
| N8 | 4.4E-04 | 1.4E-06 | 4.3E-03 | 1.2E-04 | 1.1E-03 | 6.4E-06 | 9.8E-03 | 1.1E-03 | 8.1E-03 | 2.7E-04 | 2.5E-02 | 2.8E-03 | 1.1E-02 | 5.8E-04 | | | 1.3E-02 | 4.0E-03 | | |
| G9 | 1.7E-03 | 9.0E-06 | 1.6E-02 | 8.9E-04 | 4.9E-03 | 3.1E-05 | | | 1.1E-02 | 1.7E-03 | | | | | | | | | | |
| K10 | 4.1E-02 | - | | | | | | | | | | | | | | | | | | |
| T11 | | | | | | | | | | | | | | | | | | | | |
| L12 | 2.1E-02 | 3.3E-04 | | | | | | | | | | | | | | | | | | |
| K13 | 3.0E-03 | 1.7E-05 | | | 1.2E-02 | 2.5E-04 | | | | | | | | | | | | | | |
| G14 | 1.3E-03 | 5.5E-06 | 1.6E-02 | 6.2E-04 | 4.0E-03 | 2.0E-05 | 6.0E-03 | - | 4.1E-02 | 2.8E-02 | | | | | | | | | | |
| E15 | | | | | | | | | | | | | | | | | | | | |
| T16 | 1.0E-04 | 1.2E-06 | 1.2E-03 | 2.9E-06 | 2.8E-04 | 8.2E-07 | 3.1E-03 | 1.7E-05 | 2.4E-03 | 2.8E-05 | 1.1E-02 | 6.9E-04 | 3.5E-03 | 1.3E-04 | 1.1E-02 | 1.5E-04 | 4.4E-03 | 2.0E-04 | 1.4E-02 | 3.5E-03 |
| T17 | 6.1E-03 | 3.8E-05 | | | | | | | | | | | | | | | | | | |
| T18 | 7.2E-06 | 2.3E-06 | 1.2E-04 | 8.0E-07 | 2.4E-05 | 7.8E-07 | 3.7E-04 | 3.5E-06 | 3.7E-04 | 2.8E-06 | 3.1E-03 | 4.6E-05 | 6.5E-04 | 9.6E-06 | 4.4E-03 | 1.3E-04 | 7.6E-04 | 1.8E-05 | 2.6E-03 | 3.1E-04 |
| E19 | 1.9E-02 | - | | | | | | | | | | | | | | | | | | |
| A20 | 5.8E-05 | 9.3E-07 | 2.5E-03 | 1.3E-05 | 4.4E-04 | 1.3E-06 | 6.1E-03 | 5.0E-05 | 3.6E-03 | 8.6E-05 | 1.7E-02 | 8.9E-04 | 3.9E-03 | 2.6E-04 | 2.9E-02 | - | 6.3E-03 | 7.7E-04 | 1.4E-02 | 3.4E-03 |
| V21 | | | | | | | | | | | | | | | | | | | | |
| D22 | 1.1E-03 | 3.2E-06 | | | 8.2E-03 | 1.1E-04 | | | 6.6E-03 | 2.1E-03 | | | 6.6E-03 | 1.9E-03 | | | 2.6E-03 | 1.5E-03 | | |
| A23 | 1.7E-02 | 7.1E-04 | | | | | | | | | | | | | | | | | | |
| A24 | 1.1E-03 | 5.3E-06 | | | 1.0E-02 | 1.7E-04 | | | | | | | | | | | 6.0E-04 | 6.5E-05 | | |
| T25 | 1.1E-04 | 1.8E-06 | | | 9.7E-04 | 4.7E-06 | | | | | | | | | | | 9.7E-03 | 6.1E-04 | 2.4E-02 | 6.7E-03 |
| A26 | | | 2.7E-05 | 2.1E-06 | | | 6.4E-05 | 6.3E-07 | 1.5E-05 | 1.5E-06 | 4.2E-04 | 3.0E-06 | 2.3E-05 | 1.8E-06 | 5.3E-04 | 5.9E-06 | 2.1E-05 | 2.9E-06 | 2.1E-04 | 4.0E-06 |
| E27 | 1.2E-04 | 2.8E-06 | 5.3E-05 | 6.0E-06 | | | 9.4E-05 | 3.8E-06 | | | 2.4E-04 | 2.2E-06 | 3.0E-05 | 6.0E-06 | 3.0E-04 | 3.6E-06 | | | 9.9E-05 | 3.0E-06 |
| K28 | 2.0E-05 | 3.1E-06 | 2.2E-04 | 6.9E-07 | 4.0E-05 | 6.0E-07 | 4.7E-04 | 3.9E-06 | 2.4E-04 | 1.5E-06 | 1.3E-03 | 1.1E-05 | 3.4E-04 | 3.4E-06 | 1.6E-03 | 2.5E-05 | 4.1E-04 | 6.5E-06 | 9.9E-04 | 3.6E-05 |
| V29 | | | 9.4E-05 | 8.7E-07 | 1.0E-05 | 8.0E-07 | 2.0E-04 | 1.3E-06 | 1.1E-04 | 1.1E-06 | 5.5E-04 | 4.4E-06 | 1.5E-04 | 1.8E-06 | 7.1E-04 | 7.9E-06 | 1.8E-04 | 2.7E-06 | 4.4E-04 | 9.9E-06 |
| F30 | | | 5.0E-05 | 2.4E-06 | | | 3.4E-05 | 7.6E-07 | | | 2.4E-04 | 1.8E-06 | | | 2.8E-04 | 3.5E-06 | | | 7.2E-05 | 2.9E-06 |
| K31 | | | 1.8E-05 | 3.6E-06 | | | 7.0E-05 | 1.0E-06 | 2.5E-05 | 2.1E-06 | 4.1E-04 | 6.5E-06 | 4.2E-05 | 3.6E-06 | 5.0E-04 | 7.1E-06 | 3.4E-05 | 6.6E-06 | 1.7E-04 | 4.4E-06 |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Q32** | 1.8E-05 | 1.7E-06 | 6.4E-04 | 1.2E-06 | 1.5E-04 | 7.5E-07 | 1.7E-03 | 1.7E-05 | 1.2E-03 | 1.2E-05 | 5.3E-03 | 1.2E-04 | 2.0E-03 | 3.0E-05 | 7.7E-03 | 3.0E-04 | 2.3E-03 | 7.0E-05 | 4.7E-03 | 2.4E-04 |
| **Y33** | 1.9E-05 | 3.1E-06 | 1.3E-04 | 6.6E-07 | 2.3E-05 | 6.1E-07 | 3.0E-04 | 3.7E-06 | 1.8E-04 | 1.6E-06 | 2.6E-03 | 4.1E-05 | 2.6E-04 | 4.4E-06 | 3.4E-03 | 7.2E-05 | 3.7E-04 | 7.1E-06 | 3.1E-03 | 1.6E-04 |
| **A34** | 3.0E-05 | 3.3E-06 | 6.1E-05 | 8.4E-07 | 9.2E-06 | 1.1E-06 | 1.4E-04 | 9.1E-07 | 6.0E-05 | 1.0E-06 | 7.6E-04 | 9.5E-06 | 9.2E-05 | 1.8E-06 | 9.3E-04 | 1.3E-05 | 1.2E-04 | 2.8E-06 | 4.6E-04 | 1.5E-05 |
| **N35** | 4.8E-05 | 1.2E-06 | 3.0E-04 | 2.3E-06 | 5.8E-05 | 6.4E-07 | 7.8E-04 | 1.0E-05 | 5.2E-04 | 3.4E-06 | 3.7E-03 | 5.5E-05 | 9.5E-04 | 1.2E-05 | 5.0E-03 | 1.2E-04 | 1.4E-03 | 2.9E-05 | 4.1E-03 | 2.2E-04 |
| **D36** | 3.2E-04 | 1.6E-06 | 1.1E-03 | 2.7E-06 | 2.5E-04 | 9.0E-07 | 2.2E-03 | 1.4E-05 | 1.7E-03 | 2.8E-05 | 9.1E-03 | 3.0E-04 | 3.2E-03 | 9.2E-05 | 1.5E-02 | 1.3E-03 | 2.8E-03 | 3.2E-04 | 1.3E-02 | 1.7E-03 |
| **N37** | 1.3E-04 | 1.2E-06 | 6.6E-04 | 2.0E-06 | 1.4E-04 | 1.0E-06 | 1.4E-03 | 2.2E-05 | 9.9E-04 | 1.3E-05 | 9.2E-03 | 2.0E-03 | 1.8E-03 | 3.9E-05 | 8.1E-03 | 6.6E-04 | 2.7E-03 | 8.2E-05 | 7.6E-03 | 7.7E-04 |
| **G38** | 4.4E-04 | 1.9E-06 | 8.4E-03 | 5.1E-05 | 2.7E-03 | 1.8E-05 | 2.1E-02 | 2.1E-03 | 4.4E-02 | - | | | | | | | | | | |
| **V39** | 1.0E-05 | 1.9E-06 | 1.3E-04 | 1.1E-06 | 2.3E-05 | 8.1E-07 | 2.5E-04 | 1.4E-06 | 2.0E-04 | 1.3E-06 | 1.2E-03 | 1.1E-05 | 3.3E-04 | 3.2E-06 | 1.6E-03 | 2.3E-05 | 4.5E-04 | 5.8E-06 | 1.2E-03 | 3.4E-05 |
| **D40** | 3.3E-02 | 3.6E-03 | | | 1.2E-02 | 3.1E-04 | | | 1.3E-02 | 1.1E-03 | | | | | | | | | | |
| **G41** | 9.2E-03 | 6.3E-05 | | | 1.8E-02 | 7.2E-04 | | | | | | | | | | | | | | |
| **E42** | 3.9E-05 | 1.2E-06 | 3.6E-04 | 1.8E-06 | 8.3E-05 | 8.2E-07 | 9.2E-04 | 9.4E-06 | 6.5E-04 | 4.0E-06 | 8.5E-04 | 3.1E-05 | 9.0E-04 | 7.7E-06 | 9.0E-04 | 7.4E-05 | 1.1E-03 | 1.7E-05 | 9.2E-04 | 3.3E-05 |
| **W43** | 3.0E-03 | 3.9E-05 | | | 1.3E-02 | - | | | | | | | | | | | | | | |
| **T44** | | | 1.5E-05 | 2.5E-06 | | | 3.0E-05 | 7.1E-07 | 1.0E-05 | 2.8E-06 | 2.8E-04 | 2.6E-06 | 1.2E-05 | 4.3E-06 | 3.5E-04 | 6.4E-06 | 1.3E-05 | 7.0E-06 | 1.4E-04 | 3.7E-06 |
| **Y45** | 2.5E-02 | 9.5E-04 | | | | | | | | | | | | | | | | | | |
| **D46** | | | 3.0E-05 | 2.8E-06 | | | 5.2E-05 | 5.7E-07 | 2.2E-05 | 2.0E-06 | 3.8E-04 | 4.7E-06 | 3.5E-05 | 2.6E-06 | 4.8E-04 | 9.6E-06 | 4.1E-05 | 3.4E-06 | 2.4E-04 | 7.6E-06 |
| **D47** | | | | | | | | | | | | | | | | | | | | |
| **A48** | 7.9E-03 | 5.5E-05 | | | | | | | | | | | | | | | | | | |
| **T49** | 5.6E-04 | 2.3E-06 | 2.2E-02 | 3.5E-03 | 4.6E-03 | 2.9E-05 | | | 4.1E-02 | 4.2E-02 | | | | | | | | | | |
| **K50** | 1.2E-05 | 2.1E-06 | 5.8E-04 | 1.4E-06 | 1.1E-04 | 7.3E-07 | 1.5E-03 | 1.7E-05 | 8.0E-04 | 1.3E-05 | 4.2E-03 | 1.3E-04 | 1.1E-03 | 2.5E-05 | 5.1E-03 | 1.3E-04 | 1.2E-03 | 3.9E-05 | 3.0E-03 | 2.4E-04 |
| **T51** | | | 5.6E-05 | 3.7E-06 | | | 6.9E-05 | 7.0E-07 | 1.9E-05 | 1.8E-06 | 5.6E-04 | 5.1E-06 | 1.4E-05 | 2.8E-06 | 7.0E-04 | 1.0E-05 | 8.6E-06 | 3.7E-06 | 2.4E-04 | 5.1E-06 |
| **F52** | | | 3.3E-05 | 2.1E-06 | | | 6.3E-05 | 7.7E-07 | 2.9E-05 | 3.0E-06 | 5.2E-04 | 8.7E-06 | 2.6E-05 | 5.0E-06 | 6.8E-04 | 1.8E-05 | 6.3E-06 | 8.9E-06 | 2.4E-04 | 1.0E-05 |
| **T53** | | | 2.5E-05 | 1.8E-06 | | | 7.2E-05 | 7.5E-07 | 2.4E-05 | 2.0E-06 | 6.2E-04 | 7.1E-06 | 3.3E-05 | 2.9E-06 | 7.8E-04 | 1.7E-05 | | | 2.9E-04 | 7.1E-06 |
| **V54** | | | 2.0E-05 | 2.8E-06 | | | 3.1E-05 | 6.8E-07 | 7.2E-06 | 2.8E-06 | 2.7E-04 | 2.0E-06 | 1.4E-05 | 3.5E-06 | 3.4E-04 | 3.9E-06 | 1.1E-05 | 6.6E-06 | 1.3E-04 | 3.3E-06 |
| **T55** | | | 6.8E-05 | 8.5E-07 | 8.9E-06 | 9.7E-07 | 1.7E-04 | 9.0E-07 | 1.1E-04 | 9.4E-07 | 9.2E-04 | 7.6E-06 | 1.7E-04 | 2.2E-06 | 1.3E-03 | 2.2E-05 | 2.2E-04 | 4.7E-06 | 8.2E-04 | 2.8E-05 |
| **E56** | 4.8E-04 | 1.6E-06 | 1.2E-03 | 3.0E-06 | 3.1E-04 | 9.8E-07 | 1.8E-03 | 1.4E-05 | 1.1E-03 | 9.5E-06 | 4.6E-03 | 9.1E-05 | 1.6E-03 | 2.7E-05 | 6.1E-03 | 3.9E-04 | 2.0E-03 | 5.4E-05 | 4.8E-03 | 3.3E-04 |
| **W43 sc** | 4.2E-03 | 2.9E-05 | | | 1.5E-02 | 6.7E-04 | | | | | | | | | | | | | | |

Rates of exchange, $k_{ex}$, were obtained for each backbone amide through plots of peak intensity against exchange time fitted to a single exponential function ($I(t)=I_0$ exp (-$k_{ex}$ $t$) + Plateau). Green indicates residues for which exchange is slow enough to quantify, red indicates those residues fast enough to quantify, pink indicates residues that decay too rapidly to quantify, and blue indicates those residues that decay too slowly to quantify. The obtained values are separated by condition and then by temperature. Moreover, the residues in the sequence (left) are coloured by their secondary structure, and the values of the indole side chain of W43 were added to the bottom. SE, standard error was obtained directly from the fitting. An ambiguous fitting was found for a few residues with fast exchange, as indicated by yellow. Thr25 and Glu27 peaks are overlapped in the $^{1}$H-$^{15}$N HSQC spectra recorded for GB1 in 0.75 M [Ch][Glu] at 298.2 K (indicated by orange). All the data was recorded for 1 mM $^{15}$N GB1 at 298.2 K or 310.2 K on a 600.13 MHz spectrometer. The $pD_{corr}$ for the (per-)deuterated samples, $D_2O$, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu], used for 298.2 K and 310.2 K measurements were $pD_{corr} = 7.1 \pm 0.2$ (6.9, 7.1, 7.2, 7.2, 7.4, respectively) and $pD_{corr} = 7.3 \pm 0.1$ (7.2, 7.5, 7.2, 7.2, 7.4, respectively).

**Table A.4. $\Delta G_{HX}$ values (kcal/mol) for GB1 in water and [Ch][Glu].**

| | water | | | | 0.1 M [Ch][Glu] | | | | 0.75 M [Ch][Glu] | | | | 1.0 M [Ch][Glu] | | | | 1.5 M [Ch][Glu] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | |
| | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE | $\Delta G_{HX}$ (kcal/mol) | SE |
| **M1** | | | | | | | | | | | | | | | | | | | | |
| **Q2** | | | | | | | | | | | | | | | | | | | | |
| **Y3** | | | **7.69** | 0.02 | | | **7.212** | 0.005 | 6.84 | 0.03 | **5.719** | 0.009 | 7.09 | 0.04 | **5.65** | 0.01 | 7.43 | 0.08 | **6.58** | 0.02 |
| **K4** | | | **7.93** | 0.02 | | | **7.65** | 0.01 | | | **6.221** | 0.007 | | | **6.12** | 0.01 | | | **7.14** | 0.02 |
| **L5** | | | **7.61** | 0.06 | | | **7.413** | 0.007 | 7.3 | 0.3 | **5.921** | 0.005 | 7.4 | 0.4 | **5.863** | 0.008 | | | **6.87** | 0.02 |
| **I6** | | | **7.0** | 0.1 | | | **7.08** | 0.03 | | | **5.79** | 0.01 | | | **5.70** | 0.01 | | | **6.68** | 0.09 |
| **L7** | | | **6.44** | 0.01 | 6.8 | 0.1 | **6.280** | 0.004 | 5.222 | 0.007 | **4.879** | 0.004 | 5.211 | 0.007 | **4.742** | 0.008 | 5.34 | 0.01 | **5.36** | 0.01 |
| **N8** | 5.467 | 0.002 | **5.34** | 0.02 | 5.089 | 0.003 | **5.20** | 0.07 | 4.04 | 0.02 | **4.17** | 0.07 | 4.05 | 0.03 | | | 4.2 | 0.2 | | |
| **G9** | 5.082 | 0.003 | **4.96** | 0.03 | 4.648 | 0.004 | | | 4.30 | 0.09 | | | | | | | | | | |
| **K10** | 2.59 | - | | | | | | | | | | | | | | | | | | |
| **T11** | | | | | | | | | | | | | | | | | | | | |
| **L12** | 2.275 | 0.009 | | | | | | | | | | | | | | | | | | |
| **K13** | 3.611 | 0.003 | | | 2.97 | 0.01 | | | | | | | | | | | | | | |
| **G14** | 4.996 | 0.003 | **4.70** | 0.02 | 4.498 | 0.003 | **5.66** | - | 3.2 | 0.4 | | | | | | | | | | |
| **E15** | | | | | | | | | | | | | | | | | | | | |
| **T16** | 5.657 | 0.007 | **5.404** | 0.001 | 5.237 | 0.002 | **5.199** | 0.003 | 4.081 | 0.007 | **3.98** | 0.04 | 4.06 | 0.02 | **3.972** | 0.008 | 4.16 | 0.03 | **4.1** | 0.2 |
| **T17** | 3.710 | 0.004 | | | | | | | | | | | | | | | | | | |
| **T18** | 7.7 | 0.2 | **7.339** | 0.004 | 7.173 | 0.019 | **7.011** | 0.006 | 5.656 | 0.004 | **5.251** | 0.009 | 5.545 | 0.009 | **5.03** | 0.02 | 5.68 | 0.01 | **5.65** | 0.07 |
| **E19** | 2.45 | - | | | | | | | | | | | | | | | | | | |
| **A20** | 6.089 | 0.009 | **5.061** | 0.003 | 5.071 | 0.002 | **4.882** | 0.005 | 3.93 | 0.01 | **3.78** | 0.03 | 4.10 | 0.04 | **3.46** | - | 4.05 | 0.07 | **4.2** | 0.1 |
| **V21** | | | | | | | | | | | | | | | | | | | | |
| **D22** | 3.973 | 0.002 | | | 2.941 | 0.008 | | | 3.2 | 0.2 | | | 3.4 | 0.2 | | | 4.2 | 0.3 | | |
| **A23** | 2.69 | 0.02 | | | | | | | | | | | | | | | | | | |
| **A24** | 4.528 | 0.003 | | | 3.390 | 0.009 | | | | | | | | | | | 5.65 | 0.07 | | |
| **T25** | 5.791 | 0.009 | | | 4.703 | 0.003 | | | | | | | | | | | 3.90 | 0.04 | **4.0** | 0.2 |
| **A26** | | | **8.34** | 0.05 | | | **8.187** | 0.006 | 7.63 | 0.06 | **6.570** | 0.004 | 7.62 | 0.05 | **6.434** | 0.007 | 7.91 | 0.08 | **7.30** | 0.01 |
| **E27** | 5.16 | 0.01 | **6.93** | 0.07 | | | **6.94** | 0.02 | | | **5.900** | 0.006 | 6.5 | 0.1 | **5.775** | 0.007 | | | **6.75** | 0.02 |
| **K28** | 6.67 | 0.09 | **6.494** | 0.002 | 6.435 | 0.009 | **6.400** | 0.005 | 5.473 | 0.004 | **5.334** | 0.005 | 5.492 | 0.006 | **5.21** | 0.01 | 5.610 | 0.009 | **5.79** | 0.02 |
| **V29** | | | **6.471** | 0.006 | 6.72 | 0.05 | **6.382** | 0.004 | 5.432 | 0.006 | **5.296** | 0.005 | 5.445 | 0.007 | **5.146** | 0.007 | 5.551 | 0.009 | **5.74** | 0.01 |
| **F30** | | | **7.15** | 0.03 | | | **7.76** | 0.01 | | | **6.100** | 0.005 | | | **6.005** | 0.008 | | | **7.13** | 0.03 |
| **K31** | | | **8.4** | 0.1 | | | **7.876** | 0.009 | 7.11 | 0.05 | **6.33** | 0.01 | 7.02 | 0.05 | **6.215** | 0.009 | 7.4 | 0.1 | **7.19** | 0.02 |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Q32** | 7.24 | 0.06 | **6.369** | 0.001 | 6.157 | 0.003 | **6.130** | 0.006 | 5.021 | 0.006 | **4.98** | 0.01 | 4.957 | 0.009 | **4.76** | 0.02 | 5.10 | 0.02 | **5.35** | 0.03 |
| **Y33** | 6.86 | 0.10 | **7.008** | 0.003 | 6.92 | 0.02 | **6.845** | 0.008 | 5.816 | 0.005 | **5.08** | 0.01 | 5.82 | 0.01 | **4.89** | 0.01 | 5.83 | 0.01 | **5.27** | 0.03 |
| **A34** | 6.74 | 0.06 | **7.631** | 0.008 | 7.63 | 0.07 | **7.479** | 0.004 | 6.63 | 0.01 | **5.997** | 0.008 | 6.59 | 0.01 | **5.871** | 0.009 | 6.66 | 0.01 | **6.60** | 0.02 |
| **N35** | 7.07 | 0.02 | **7.275** | 0.005 | 7.139 | 0.007 | **7.057** | 0.008 | 5.946 | 0.004 | **5.637** | 0.009 | 5.810 | 0.007 | **5.46** | 0.01 | 5.83 | 0.01 | **5.88** | 0.03 |
| **D36** | 5.315 | 0.003 | **5.804** | 0.001 | 5.623 | 0.002 | **5.740** | 0.004 | 4.62 | 0.01 | **4.43** | 0.02 | 4.45 | 0.02 | **4.14** | 0.05 | 4.77 | 0.07 | **4.51** | 0.08 |
| **N37** | 6.240 | 0.005 | **6.534** | 0.002 | 6.358 | 0.004 | **6.440** | 0.010 | 5.321 | 0.008 | **4.8** | 0.1 | 5.19 | 0.01 | **4.91** | 0.05 | 5.18 | 0.02 | **5.24** | 0.06 |
| **G38** | 5.901 | 0.003 | **5.368** | 0.004 | 5.004 | 0.004 | **5.17** | 0.06 | 3.46 | - | | | | | | | | | | |
| **V39** | 6.6 | 0.1 | **6.353** | 0.005 | 6.30 | 0.02 | **6.318** | 0.004 | 5.134 | 0.004 | **4.903** | 0.006 | 5.040 | 0.006 | **4.701** | 0.009 | 5.088 | 0.008 | **5.18** | 0.02 |
| **D40** | 1.93 | 0.06 | | | 2.71 | 0.02 | | | 2.78 | 0.05 | | | | | | | | | | |
| **G41** | 3.416 | 0.004 | | | 3.20 | 0.02 | | | | | | | | | | | | | | |
| **E42** | 6.07 | 0.02 | **5.984** | 0.003 | 5.796 | 0.006 | **5.784** | 0.006 | 4.685 | 0.004 | **5.371** | 0.022 | 4.708 | 0.005 | **5.34** | 0.05 | 4.826 | 0.009 | **5.63** | 0.02 |
| **W43** | 3.191 | 0.008 | | | 2.51 | - | | | | | | | | | | | | | | |
| **T44** | | | **8.2** | 0.1 | | | **8.12** | 0.01 | 7.4 | 0.2 | **6.282** | 0.006 | 7.5 | 0.2 | **6.14** | 0.01 | 7.7 | 0.3 | **7.01** | 0.02 |
| **Y45** | 2.59 | 0.02 | | | | | | | | | | | | | | | | | | |
| **D46** | | | **7.65** | 0.06 | | | **7.684** | 0.007 | 6.82 | 0.05 | **6.005** | 0.008 | 6.75 | 0.04 | **5.85** | 0.01 | 6.90 | 0.05 | **6.59** | 0.02 |
| **D47** | | | | | | | | | | | | | | | | | | | | |
| **A48** | 3.136 | 0.004 | | | | | | | | | | | | | | | | | | |
| **T49** | 4.856 | 0.002 | **3.8** | 0.1 | 3.783 | 0.004 | | | 2.6 | 0.6 | | | | | | | | | | |
| **K50** | 7.5 | 0.1 | **6.401** | 0.001 | 6.309 | 0.004 | **6.188** | 0.007 | 5.244 | 0.009 | **5.09** | 0.02 | 5.28 | 0.01 | **4.98** | 0.02 | 5.47 | 0.02 | **5.60** | 0.05 |
| **T51** | | | **7.68** | 0.04 | | | **7.932** | 0.006 | 7.32 | 0.06 | **6.180** | 0.006 | 7.7 | 0.1 | **6.042** | 0.009 | 8.2 | 0.3 | **7.01** | 0.01 |
| **F52** | | | **7.88** | 0.04 | | | **7.852** | 0.008 | 6.93 | 0.06 | **6.11** | 0.01 | 7.2 | 0.1 | **5.94** | 0.02 | 8.3 | 0.8 | **6.88** | 0.03 |
| **T53** | | | **8.11** | 0.04 | | | **7.820** | 0.006 | 7.09 | 0.05 | **6.037** | 0.007 | 7.11 | 0.05 | **5.90** | 0.01 | | | **6.81** | 0.02 |
| **V54** | | | **7.54** | 0.09 | | | **7.64** | 0.01 | 7.1 | 0.2 | **5.849** | 0.004 | 7.0 | 0.1 | **5.708** | 0.007 | 7.3 | 0.4 | **6.62** | 0.02 |
| **T55** | | | **7.198** | 0.008 | 7.29 | 0.06 | **7.008** | 0.003 | 5.923 | 0.005 | **5.505** | 0.005 | 5.873 | 0.008 | **5.30** | 0.01 | 5.96 | 0.01 | **5.88** | 0.02 |
| **E56** | 2.184 | 0.002 | **2.733** | 0.002 | 2.620 | 0.002 | **2.871** | 0.005 | 1.993 | 0.005 | **1.84** | 0.01 | 1.97 | 0.01 | **1.66** | 0.04 | 2.06 | 0.02 | **2.10** | 0.04 |

$\Delta G_{HX}$ values were obtained for each backbone amide accordingly to the equation discussed in the main text. The standard error (SE) was obtained from propagation of the error of $k_{ex}$. A few $\Delta G_{HX}$ values do not have error associated because they are obtained with an ambiguous fitting, indicated with a "-" in SE parameter. The obtained values are separated by condition and then by temperature. For clarity, the blank spaces are maintained to indicate those residues that exchange too fast or too slow to be accurately measured. The values for the condition at 310.2 K are shown in bold.

**Table A.5. Rates of exchange ($k_{ex}$, $s^{-1}$) extracted with CLEANEX-PM for GB1 in water and [Ch][Glu].**

| | water | | | | 0.1 M [Ch][Glu] | | | | 0.75M [Ch][Glu] | | | | 1.0 M [Ch][Glu] | | | | 1.5M [Ch][Glu] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | |
| | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE | $k_{ex}$ ($s^{-1}$) | SE |
| **Q2** | 2.55 | 0.05 | 24.0 | 0.9 | 4.76 | 0.06 | 33 | 1 | 5.5 | 0.2 | 39 | 2 | 2.54 | 0.06 | 17 | 1 | 3.2 | 0.2 | 22 | 1 |
| **K10** | | | 0.98 | 0.06 | | | 5.6 | 0.4 | | | 5.5 | 0.3 | | | 2.5 | 0.2 | | | 4.2 | 0.3 |
| **T11** | 2.34 | 0.08 | 11.9 | 0.3 | 6.9 | 0.2 | 64 | 1 | 6.4 | 0.1 | 42 | 2 | 3.5 | 0.2 | 20 | 1 | 4.7 | 0.2 | 37 | 2 |
| **L12** | 0.35 | 0.05 | 2.5 | 0.1 | 0.8 | 0.1 | 12.9 | 0.5 | | | 8.8 | 0.7 | | | 4.5 | 0.2 | | | 7.5 | 0.4 |
| **E15** | 1.17 | 0.04 | 6.1 | 0.1 | 4.0 | 0.1 | 33.7 | 0.7 | 3.6 | 0.1 | 23.7 | 0.9 | 1.79 | 0.09 | 10.2 | 0.2 | 3.1 | 0.1 | 17.4 | 0.9 |
| **T17** | 0.28 | 0.04 | 3.02 | 0.08 | 1.5 | 0.1 | 17.0 | 0.7 | | | 13.4 | 0.5 | | | 5.9 | 0.4 | | | 8.8 | 0.5 |
| **E19** | 1.15 | 0.03 | 8.3 | 0.1 | 2.9 | 0.1 | 37.5 | 0.8 | 2.8 | 0.2 | 25.3 | 0.9 | | | 11.3 | 0.5 | | | 23 | 1 |
| **V21** | 3.42 | 0.09 | 19.5 | 0.6 | 7.6 | 0.2 | 90 | 2 | 5.3 | 0.3 | 58.2 | 3.4 | 2.5 | 0.1 | 20.0 | 0.8 | | | 25 | 2 |
| **A23** | 0.59 | 0.02 | 1.77 | 0.04 | 3.0 | 0.2 | 18.4 | 0.6 | 2.2 | 0.1 | 30.0 | 0.5 | | | 11.9 | 0.4 | | | 21 | 1 |
| **Y45** | 0.41 | 0.02 | 2.49 | 0.05 | 3.0 | 0.2 | 18.4 | 0.4 | 4.0 | 0.3 | 30 | 1 | | | 14.8 | 0.4 | | | 24 | 1 |
| **D47** | 3.92 | 0.08 | 20.3 | 0.6 | 11.7 | 0.4 | 106 | 3 | 7.7 | 0.3 | 66 | 3 | 5.7 | 0.3 | 25 | 1 | 7.6 | 0.4 | 37 | 2 |
| **A48** | 0.45 | 0.05 | 2.69 | 0.09 | 1.13 | 0.09 | 13.8 | 0.5 | | | 7.4 | 0.2 | | | 4.2 | 0.3 | | | 5.4 | 0.3 |

Values of $k_{ex}$ were obtained for backbone amide from an initial slope analysis. For each residue, the ratio of peak volume intensity ($V/V_0$, where $V_0$ is the intensity of the corresponding peak in a fully relaxed spectrum) as a function of mixing time ($\tau_{mix}$) was fitted to a simple linear regression. At short mixing time, $k_{ex}$ reflects the initial slope ($V/V_0$ vs $\tau_{mix}$) and can be approximated by $V/V_0 = fk_{ex}\tau_{mix}$, where $f = 1/$fraction of unsaturated water and it was used to correct $k_{ex}$, assuming that the saturation of water leads to the decrease in intensity for pure exchange peaks in the same proportion. The $f$ for the different samples, water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu], obtained for 298.2 K and 310.2 K measurements were $f = 0.74, 0.51, 0.62, 0.66, 0.79$, respectively, and $f = 0.61, 0.40, 0.47, 0.52, 0.67$, respectively. The number of scans (NS) also differ accordingly with the condition, for water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu], at 298.2 K or 310.2 K measurements were 8, 8, 16, 32, 64, respectively, or NS = 8, 8, 16, 16, 32, respectively. SE, standard error was obtained directly from the fitting. For clarity, the blank spaces are maintained to indicate those residues that cannot be measured, as they are in the limit of the experiment. All the data was recorded for 0.6 mM $^{15}N$ $^{13}C$ GB1 at 298.2 K or 310.2 K on a 500.34 MHz spectrometer. The pH for the different samples: water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M [Ch][Glu], were pH = 7.2 ± 0.1 (7.1, 7.0, 7.2, 7.2, 7.4, respectively).

**Table A.6. $\Delta G_{CLEANEX}$ values (kcal/mol) for GB1 in water and [Ch][Glu].**

| | water | | | | 0.1 M [Ch][Glu] | | | | 0.75M [Ch][Glu] | | | | 1.0 M [Ch][Glu] | | | | 1.5M [Ch][Glu] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | | 298.2 K | | 310.2 K | |
| | $\Delta G_{CL}$ *EANEX* (kcal/mol) | SE | $\Delta G_{CL}$ ***EANEX*** **(kcal/mol)** | SE | $\Delta G_{CL}$ *EANEX* (kcal/mol) | SE | $\Delta G_{CL}$ ***EANEX*** **(kcal/mol)** | SE | $\Delta G_{CL}$ *EANEX* (kcal/mol) | SE | $\Delta G_{CL}$ ***EANEX*** **(kcal/mol)** | SE | $\Delta G_{CL}$ *EANEX* (kcal/mol) | SE | $\Delta G_{CL}$ ***EANEX*** **(kcal/mol)** | SE | $\Delta G_{CL}$ *EANEX* (kcal/mol) | SE | $\Delta G_{CL}$ ***EANEX*** **(kcal/mol)** | SE |
| **Q2** | 2.40 | 0.01 | **1.78** | 0.02 | 1.92 | 0.01 | **1.47** | 0.02 | 1.96 | 0.02 | **1.51** | 0.04 | 2.49 | 0.01 | **2.07** | 0.04 | 2.44 | 0.03 | **2.02** | 0.03 |
| **K10** | | | **1.65** | 0.04 | | | **0.41** | 0.05 | | | **0.63** | 0.03 | | | **1.24** | 0.04 | | | **1.09** | 0.04 |
| **T11** | 0.32 | 0.02 | **0.00** | 0.02 | -0.48 | 0.01 | **-1.21** | 0.01 | -0.23 | 0.01 | **-0.74** | 0.03 | 0.23 | 0.03 | **-0.17** | 0.03 | 0.23 | 0.02 | **-0.36** | 0.04 |
| **L12** | 0.86 | 0.09 | **0.36** | 0.02 | 0.19 | 0.10 | **-0.83** | 0.02 | | | **-0.38** | 0.05 | | | **0.14** | 0.03 | | | **0.01** | 0.03 |
| **E15** | 0.20 | 0.02 | **-0.13** | 0.01 | -0.68 | 0.02 | **-1.36** | 0.01 | -0.41 | 0.02 | **-0.93** | 0.02 | 0.10 | 0.03 | **-0.30** | 0.01 | -0.06 | 0.02 | **-0.44** | 0.03 |
| **T17** | 1.69 | 0.08 | **0.96** | 0.02 | 0.53 | 0.06 | **-0.28** | 0.02 | | | **0.08** | 0.02 | | | **0.70** | 0.04 | | | **0.64** | 0.03 |
| **E19** | 0.26 | 0.02 | **-0.28** | 0.01 | -0.45 | 0.02 | **-1.39** | 0.01 | -0.23 | 0.04 | **-0.93** | 0.02 | | | **-0.32** | 0.03 | | | **-0.57** | 0.04 |
| **V21** | -0.93 | 0.02 | **-1.37** | 0.02 | -1.57 | 0.01 | **-2.48** | 0.02 | -1.15 | 0.03 | **-2.00** | 0.04 | -0.59 | 0.03 | **-1.23** | 0.03 | | | **-1.19** | 0.04 |
| **A23** | 0.83 | 0.02 | **0.85** | 0.01 | -0.29 | 0.04 | **-0.76** | 0.02 | 0.08 | 0.03 | **-0.85** | 0.01 | | | **-0.17** | 0.02 | | | **-0.32** | 0.04 |
| **Y45** | 1.19 | 0.03 | **0.79** | 0.01 | -0.14 | 0.03 | **-0.61** | 0.01 | -0.11 | 0.04 | **-0.70** | 0.02 | | | **-0.15** | 0.02 | | | **-0.26** | 0.03 |
| **D47** | -0.70 | 0.01 | **-1.07** | 0.02 | -1.51 | 0.02 | **-2.26** | 0.02 | -1.06 | 0.02 | **-1.76** | 0.02 | -0.77 | 0.03 | **-1.05** | 0.03 | -0.77 | 0.03 | **-1.11** | 0.03 |
| **A48** | 0.99 | 0.06 | **0.59** | 0.02 | 0.28 | 0.05 | **-0.59** | 0.02 | | | **0.01** | 0.02 | | | **0.48** | 0.04 | | | **0.50** | 0.04 |

$\Delta G_{CLEANEX}$ values were obtained for each backbone amide in a similar way to H/D exchange data. The standard error (SE) was obtained from propagation of the error of $k_{ex}$. The obtained values are separated by condition and then by temperature. For clarity, the blank spaces are maintained to indicate those residues that cannot be measured, as they are in the limit of the experiment. The values for the condition at 310.2 K are shown in bold.

**Table A.7. Relaxation data for GB1 in water and in the presence of [Ch][Glu], at 298.2 K and 600 MHz.**

| | Water | | | | | | | | 0.1 M [Ch][Glu] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **$R_1$ ($s^{-1}$)** | SE | **$R_2$ ($s^{-1}$)** | SE | **NOE** | SE | **$R_2/R_1$** | SE | **$R_1$ ($s^{-1}$)** | SE | **$R_2$ ($s^{-1}$)** | SE | **NOE** | SE | **$R_2/R_1$** | SE |
| **Q2** | **2.81** | 0.14 | **4.79** | 0.13 | **0.74** | 0.03 | **1.70** | 0.10 | **3.26** | 0.13 | **5.29** | 0.29 | **0.69** | 0.04 | **1.62** | 0.11 |
| **Y3** | **2.84** | 0.09 | **4.28** | 0.05 | **0.73** | 0.03 | **1.50** | 0.05 | **2.92** | 0.10 | **4.58** | 0.08 | **0.75** | 0.03 | **1.57** | 0.06 |
| **K4** | **2.95** | 0.09 | **4.54** | 0.04 | **0.73** | 0.03 | **1.54** | 0.05 | **3.01** | 0.12 | **4.87** | 0.05 | **0.76** | 0.04 | **1.62** | 0.07 |
| **L5** | **3.01** | 0.08 | **4.55** | 0.01 | **0.67** | 0.02 | **1.51** | 0.04 | **3.10** | 0.11 | **4.90** | 0.04 | **0.67** | 0.03 | **1.58** | 0.06 |
| **I6** | **2.96** | 0.08 | **4.73** | 0.02 | **0.73** | 0.03 | **1.60** | 0.05 | **3.03** | 0.12 | **5.52** | 0.07 | **0.75** | 0.04 | **1.82** | 0.08 |
| **L7** | **2.93** | 0.08 | **4.60** | 0.02 | **0.72** | 0.03 | **1.57** | 0.04 | **3.00** | 0.10 | **4.89** | 0.02 | **0.69** | 0.03 | **1.63** | 0.06 |
| **N8** | **2.80** | 0.08 | **4.78** | 0.03 | **0.63** | 0.03 | **1.71** | 0.05 | **2.86** | 0.09 | **5.36** | 0.04 | **0.65** | 0.03 | **1.87** | 0.06 |
| **G9** | **2.87** | 0.08 | **4.34** | 0.01 | **0.63** | 0.02 | **1.51** | 0.04 | **2.90** | 0.10 | **4.63** | 0.04 | **0.65** | 0.03 | **1.59** | 0.06 |
| **K10** | **2.74** | 0.08 | **4.21** | 0.02 | **0.56** | 0.03 | **1.54** | 0.05 | **2.85** | 0.10 | **4.58** | 0.07 | **0.64** | 0.03 | **1.61** | 0.06 |
| **T11** | **2.57** | 0.11 | **4.43** | 0.08 | **0.63** | 0.02 | **1.72** | 0.08 | **2.92** | 0.11 | **5.03** | 0.26 | **0.69** | 0.04 | **1.73** | 0.11 |
| **L12** | **2.23** | 0.08 | **3.39** | 0.04 | **0.67** | 0.02 | **1.52** | 0.06 | **2.39** | 0.08 | **3.83** | 0.14 | **0.61** | 0.03 | **1.60** | 0.08 |
| **K13** | **2.65** | 0.07 | **3.99** | 0.02 | **0.66** | 0.02 | **1.51** | 0.04 | **2.69** | 0.08 | **4.27** | 0.03 | **0.73** | 0.03 | **1.59** | 0.05 |
| **G14** | **2.53** | 0.07 | **3.96** | 0.04 | **0.69** | 0.02 | **1.56** | 0.05 | **2.56** | 0.09 | **4.11** | 0.01 | **0.69** | 0.03 | **1.60** | 0.06 |
| **E15** | **2.60** | 0.09 | **4.13** | 0.06 | **0.75** | 0.02 | **1.59** | 0.06 | **2.88** | 0.10 | **4.66** | 0.19 | **0.72** | 0.03 | **1.62** | 0.09 |
| **T16** | **2.83** | 0.08 | **4.36** | 0.01 | **0.73** | 0.03 | **1.54** | 0.04 | **2.92** | 0.10 | **4.71** | 0.04 | **0.73** | 0.03 | **1.61** | 0.06 |
| **T17** | **2.72** | 0.11 | **4.18** | 0.07 | **0.76** | 0.02 | **1.53** | 0.07 | **2.98** | 0.12 | **4.65** | 0.18 | **0.77** | 0.03 | **1.56** | 0.09 |
| **T18** | **2.91** | 0.09 | **4.44** | 0.03 | **0.68** | 0.03 | **1.53** | 0.05 | **3.01** | 0.11 | **4.84** | 0.06 | **0.71** | 0.03 | **1.61** | 0.06 |
| **E19** | **2.67** | 0.11 | **4.39** | 0.09 | **0.71** | 0.02 | **1.65** | 0.08 | **2.97** | 0.11 | **5.10** | 0.26 | **0.68** | 0.03 | **1.72** | 0.11 |
| **A20** | **2.80** | 0.10 | **5.03** | 0.04 | **0.67** | 0.03 | **1.80** | 0.06 | **2.97** | 0.11 | **5.61** | 0.14 | **0.66** | 0.03 | **1.89** | 0.08 |
| **V21** | **2.88** | 0.13 | **4.72** | 0.11 | **0.74** | 0.03 | **1.64** | 0.08 | **3.27** | 0.12 | **5.28** | 0.26 | **0.76** | 0.04 | **1.62** | 0.10 |
| **D22** | **2.86** | 0.10 | **4.65** | 0.06 | **0.72** | 0.02 | **1.63** | 0.06 | **3.00** | 0.12 | **5.06** | 0.13 | **0.68** | 0.03 | **1.68** | 0.08 |
| **A23** | **2.95** | 0.11 | **4.85** | 0.07 | **0.68** | 0.03 | **1.64** | 0.07 | **3.24** | 0.12 | **5.39** | 0.19 | **0.70** | 0.04 | **1.67** | 0.08 |
| **A24** | **2.81** | 0.08 | **4.86** | 0.03 | **0.78** | 0.03 | **1.73** | 0.05 | **2.87** | 0.11 | **5.23** | 0.06 | **0.75** | 0.03 | **1.82** | 0.07 |
| **T25** | **2.89** | 0.12 | **4.94** | 0.07 | **0.71** | 0.03 | **1.71** | 0.07 | **3.03** | 0.13 | **5.40** | 0.16 | **0.73** | 0.03 | **1.78** | 0.09 |
| **A26** | **3.03** | 0.09 | **4.83** | 0.02 | **0.77** | 0.03 | **1.59** | 0.05 | **3.09** | 0.11 | **5.21** | 0.07 | **0.79** | 0.04 | **1.69** | 0.06 |
| **E27** | **2.99** | 0.09 | **5.03** | 0.01 | **0.73** | 0.02 | **1.68** | 0.05 | **3.00** | 0.11 | **5.43** | 0.02 | **0.72** | 0.03 | **1.81** | 0.07 |
| **K28** | **2.97** | 0.09 | **5.12** | 0.02 | **0.68** | 0.03 | **1.72** | 0.05 | **2.94** | 0.10 | **5.49** | 0.01 | **0.71** | 0.03 | **1.87** | 0.07 |
| **V29** | **2.96** | 0.08 | **4.92** | 0.01 | **0.70** | 0.03 | **1.66** | 0.05 | **2.96** | 0.13 | **5.29** | 0.02 | **0.73** | 0.03 | **1.78** | 0.08 |
| **F30** | **3.02** | 0.08 | **5.05** | 0.02 | **0.74** | 0.02 | **1.67** | 0.05 | **3.02** | 0.13 | **5.41** | 0.02 | **0.77** | 0.03 | **1.79** | 0.08 |
| **K31** | **3.05** | 0.09 | **5.14** | 0.01 | **0.78** | 0.03 | **1.68** | 0.05 | **3.01** | 0.12 | **5.47** | 0.01 | **0.77** | 0.04 | **1.82** | 0.07 |
| **Q32** | **2.96** | 0.08 | **5.14** | 0.01 | **0.79** | 0.03 | **1.74** | 0.05 | **2.94** | 0.12 | **5.53** | 0.02 | **0.79** | 0.03 | **1.88** | 0.08 |
| **Y33** | **3.00** | 0.09 | **5.04** | 0.02 | **0.69** | 0.03 | **1.68** | 0.05 | **3.02** | 0.11 | **5.38** | 0.01 | **0.73** | 0.03 | **1.78** | 0.06 |
| **A34** | **3.08** | 0.09 | **5.06** | 0.03 | **0.70** | 0.03 | **1.64** | 0.05 | **3.07** | 0.12 | **5.42** | 0.02 | **0.73** | 0.04 | **1.77** | 0.07 |
| **N35** | **2.94** | 0.08 | **5.10** | 0.01 | **0.67** | 0.02 | **1.73** | 0.05 | **2.87** | 0.10 | **5.43** | 0.01 | **0.70** | 0.03 | **1.89** | 0.07 |
| **D36** | **2.98** | 0.09 | **5.04** | 0.02 | **0.72** | 0.03 | **1.69** | 0.05 | **2.99** | 0.11 | **5.37** | 0.01 | **0.71** | 0.03 | **1.80** | 0.07 |
| **N37** | **2.82** | 0.08 | **4.49** | 0.02 | **0.73** | 0.03 | **1.59** | 0.05 | **2.83** | 0.12 | **4.83** | 0.03 | **0.75** | 0.03 | **1.71** | 0.07 |
| **G38** | **2.91** | 0.08 | **4.93** | 0.03 | **0.75** | 0.03 | **1.69** | 0.05 | **2.91** | 0.12 | **5.61** | 0.01 | **0.79** | 0.03 | **1.93** | 0.08 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **V39** | **2.86** | 0.08 | **4.81** | 0.01 | **0.74** | 0.02 | **1.69** | 0.05 | **2.84** | 0.11 | **5.05** | 0.03 | **0.79** | 0.03 | **1.78** | 0.07 |
| **D40** | **2.49** | 0.07 | **4.66** | 0.05 | **0.54** | 0.02 | **1.87** | 0.06 | **2.50** | 0.10 | **5.60** | 0.02 | **0.61** | 0.03 | **2.24** | 0.09 |
| **G41** | **1.95** | 0.05 | **3.62** | 0.06 | **0.53** | 0.02 | **1.85** | 0.05 | **1.94** | 0.05 | **4.26** | 0.03 | **0.46** | 0.03 | **2.19** | 0.06 |
| **E42** | **2.82** | 0.08 | **4.47** | 0.02 | **0.71** | 0.02 | **1.59** | 0.05 | **2.85** | 0.09 | **4.84** | 0.01 | **0.74** | 0.03 | **1.70** | 0.05 |
| **W43** | **2.85** | 0.08 | **4.85** | 0.08 | **0.68** | 0.03 | **1.70** | 0.06 | **2.91** | 0.11 | **5.83** | 0.03 | **0.68** | 0.03 | **2.00** | 0.08 |
| **T44** | **2.89** | 0.09 | **4.45** | 0.03 | **0.68** | 0.02 | **1.54** | 0.05 | **3.05** | 0.10 | **4.85** | 0.07 | **0.67** | 0.03 | **1.59** | 0.06 |
| **Y45** | **2.76** | 0.11 | **4.37** | 0.05 | **0.67** | 0.02 | **1.59** | 0.07 | **3.08** | 0.12 | **4.97** | 0.19 | **0.70** | 0.03 | **1.61** | 0.09 |
| **D46** | **2.88** | 0.10 | **4.91** | 0.04 | **0.68** | 0.02 | **1.71** | 0.06 | **3.04** | 0.12 | **6.10** | 0.16 | **0.72** | 0.03 | **2.01** | 0.09 |
| **D47** | **2.83** | 0.14 | **4.35** | 0.12 | **0.66** | 0.03 | **1.54** | 0.09 | **3.27** | 0.14 | **5.04** | 0.33 | **0.73** | 0.04 | **1.54** | 0.12 |
| **A48** | **2.62** | 0.10 | **4.32** | 0.05 | **0.74** | 0.03 | **1.65** | 0.06 | **2.81** | 0.10 | **4.80** | 0.16 | **0.68** | 0.03 | **1.71** | 0.08 |
| **T49** | **2.77** | 0.11 | **4.29** | 0.08 | **0.74** | 0.03 | **1.55** | 0.07 | **2.91** | 0.12 | **4.78** | 0.19 | **0.73** | 0.03 | **1.64** | 0.09 |
| **K50** | **2.99** | 0.11 | **4.54** | 0.04 | **0.72** | 0.03 | **1.52** | 0.06 | **3.14** | 0.12 | **4.94** | 0.11 | **0.73** | 0.04 | **1.57** | 0.07 |
| **T51** | **2.88** | 0.11 | **4.33** | 0.06 | **0.76** | 0.02 | **1.50** | 0.06 | **3.04** | 0.11 | **4.76** | 0.15 | **0.74** | 0.03 | **1.56** | 0.08 |
| **F52** | **2.99** | 0.10 | **4.63** | 0.12 | **0.72** | 0.03 | **1.55** | 0.06 | **3.13** | 0.13 | **5.15** | 0.04 | **0.74** | 0.04 | **1.64** | 0.07 |
| **T53** | **2.87** | 0.09 | **4.49** | 0.03 | **0.74** | 0.03 | **1.56** | 0.05 | **3.02** | 0.13 | **4.93** | 0.09 | **0.75** | 0.03 | **1.63** | 0.08 |
| **V54** | **3.03** | 0.09 | **4.52** | 0.03 | **0.70** | 0.02 | **1.49** | 0.04 | **3.13** | 0.11 | **4.87** | 0.05 | **0.75** | 0.03 | **1.56** | 0.06 |
| **T55** | **2.90** | 0.09 | **4.38** | 0.09 | **0.72** | 0.02 | **1.51** | 0.06 | **2.99** | 0.10 | **4.79** | 0.13 | **0.70** | 0.03 | **1.60** | 0.07 |
| **E56** | **2.90** | 0.08 | **4.42** | 0.10 | **0.74** | 0.03 | **1.53** | 0.06 | **2.96** | 0.10 | **4.83** | 0.06 | **0.74** | 0.03 | **1.63** | 0.06 |
| **W43sc** | **2.47** | 0.08 | **4.14** | 0.02 | **0.65** | 0.03 | **1.68** | 0.05 | **2.50** | 0.12 | **4.87** | 0.05 | **0.67** | 0.04 | **1.95** | 0.09 |
| | **0.75 M [Ch][Glu]** | | | | | | | | **1.0 M [Ch][Glu]** | | | | | | | |
| | **$R_1$ ($s^{-1}$)** | SE | **$R_2$ ($s^{-1}$)** | SE | **NOE** | SE | **$R_2/R_1$** | SE | **$R_1$ ($s^{-1}$)** | SE | **$R_2$ ($s^{-1}$)** | SE | **NOE** | SE | **$R_2/R_1$** | SE |
| **Q2** | **2.74** | 0.07 | **7.17** | 0.22 | **0.75** | 0.03 | **2.61** | 0.10 | 2.42 | 0.08 | 8.03 | 0.18 | 0.74 | 0.05 | **3.31** | 0.13 |
| **Y3** | **2.59** | 0.09 | **6.54** | 0.06 | **0.77** | 0.03 | **2.53** | 0.09 | 2.35 | 0.09 | 7.30 | 0.12 | 0.78 | 0.04 | **3.11** | 0.13 |
| **K4** | **2.62** | 0.08 | **7.16** | 0.19 | **0.77** | 0.03 | **2.74** | 0.11 | 2.34 | 0.10 | 8.09 | 0.23 | 0.82 | 0.05 | **3.46** | 0.18 |
| **L5** | **2.67** | 0.07 | **6.88** | 0.06 | **0.82** | 0.03 | **2.58** | 0.07 | 2.44 | 0.09 | 7.89 | 0.05 | 0.81 | 0.04 | **3.23** | 0.12 |
| **I6** | **2.62** | 0.08 | **6.88** | 0.10 | **0.80** | 0.03 | **2.63** | 0.09 | 2.41 | 0.10 | 7.90 | 0.11 | 0.78 | 0.05 | **3.28** | 0.15 |
| **L7** | **2.56** | 0.07 | **6.72** | 0.04 | **0.77** | 0.03 | **2.62** | 0.07 | 2.35 | 0.08 | 7.51 | 0.05 | 0.80 | 0.04 | **3.20** | 0.11 |
| **N8** | **2.43** | 0.06 | **6.52** | 0.09 | **0.77** | 0.02 | **2.68** | 0.08 | 2.21 | 0.08 | 7.58 | 0.08 | 0.76 | 0.04 | **3.44** | 0.12 |
| **G9** | **2.53** | 0.06 | **6.69** | 0.10 | **0.78** | 0.02 | **2.65** | 0.08 | 2.26 | 0.07 | 7.49 | 0.13 | 0.78 | 0.04 | **3.31** | 0.12 |
| **K10** | **2.43** | 0.07 | **6.66** | 0.06 | **0.68** | 0.03 | **2.74** | 0.09 | 2.23 | 0.06 | 7.42 | 0.07 | 0.67 | 0.04 | **3.33** | 0.10 |
| **T11** | **2.54** | 0.06 | **6.68** | 0.10 | **0.64** | 0.02 | **2.63** | 0.08 | 2.28 | 0.06 | 7.28 | 0.13 | 0.65 | 0.04 | **3.19** | 0.10 |
| **L12** | **2.09** | 0.05 | **5.16** | 0.05 | **0.66** | 0.02 | **2.47** | 0.07 | 1.96 | 0.06 | 5.81 | 0.04 | 0.68 | 0.03 | **2.97** | 0.10 |
| **K13** | **2.32** | 0.05 | **6.14** | 0.07 | **0.68** | 0.03 | **2.65** | 0.07 | 2.08 | 0.05 | 6.94 | 0.06 | 0.68 | 0.03 | **3.34** | 0.09 |
| **G14** | **2.19** | 0.06 | **6.03** | 0.05 | **0.68** | 0.02 | **2.76** | 0.07 | 1.97 | 0.05 | 6.75 | 0.09 | 0.66 | 0.03 | **3.44** | 0.09 |
| **E15** | **2.46** | 0.06 | **6.46** | 0.06 | **0.71** | 0.02 | **2.63** | 0.07 | 2.23 | 0.08 | 7.41 | 0.04 | 0.67 | 0.03 | **3.33** | 0.11 |
| **T16** | **2.55** | 0.06 | **6.92** | 0.06 | **0.74** | 0.02 | **2.71** | 0.07 | 2.24 | 0.09 | 7.64 | 0.05 | 0.74 | 0.03 | **3.42** | 0.13 |
| **T17** | **2.58** | 0.09 | **6.52** | 0.06 | **0.75** | 0.02 | **2.53** | 0.09 | 2.29 | 0.10 | 7.26 | 0.13 | 0.77 | 0.03 | **3.17** | 0.15 |
| **T18** | **2.67** | 0.07 | **6.92** | 0.06 | **0.82** | 0.03 | **2.60** | 0.07 | 2.33 | 0.10 | 7.84 | 0.11 | 0.80 | 0.04 | **3.37** | 0.15 |
| **E19** | **2.52** | 0.08 | **6.50** | 0.10 | **0.82** | 0.02 | **2.58** | 0.09 | 2.25 | 0.07 | 7.37 | 0.06 | 0.81 | 0.04 | **3.28** | 0.11 |
| **A20** | **2.56** | 0.07 | **7.18** | 0.11 | **0.80** | 0.03 | **2.80** | 0.09 | 2.29 | 0.10 | 8.12 | 0.10 | 0.82 | 0.05 | **3.55** | 0.16 |
| **V21** | **2.79** | 0.07 | **7.22** | 0.07 | **0.77** | 0.03 | **2.59** | 0.07 | 2.46 | 0.07 | 8.01 | 0.11 | 0.81 | 0.05 | **3.26** | 0.10 |
| **D22** | **2.56** | 0.07 | **7.22** | 0.06 | **0.79** | 0.02 | **2.82** | 0.08 | 2.28 | 0.09 | 8.23 | 0.05 | 0.81 | 0.03 | **3.61** | 0.14 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **A23** | **2.75** | 0.08 | **7.68** | 0.12 | **0.79** | 0.03 | **2.79** | 0.09 | 2.44 | 0.10 | 8.60 | 0.12 | 0.90 | 0.05 | **3.53** | 0.15 |
| **A24** | **2.35** | 0.06 | **8.46** | 0.82 | **0.78** | 0.02 | **3.60** | 0.36 | 2.04 | 0.06 | 8.85 | 0.10 | 0.81 | 0.04 | **4.34** | 0.14 |
| **T25** | **2.54** | 0.09 | **7.67** | 0.66 | **0.79** | 0.02 | **3.01** | 0.28 | 2.27 | 0.07 | 9.07 | 0.06 | 0.76 | 0.01 | **4.00** | 0.12 |
| **A26** | **2.66** | 0.08 | **7.56** | 0.12 | **0.79** | 0.03 | **2.84** | 0.09 | 2.34 | 0.09 | 8.85 | 0.06 | 0.82 | 0.04 | **3.78** | 0.15 |
| **E27** | **2.52** | 0.07 | **7.99** | 0.38 | **0.86** | 0.03 | **3.17** | 0.18 | 2.60 | 0.17 | 9.18 | 0.11 | 0.71 | 0.01 | **3.54** | 0.24 |
| **K28** | **2.42** | 0.07 | **8.17** | 0.04 | **0.82** | 0.03 | **3.37** | 0.10 | 2.15 | 0.07 | 9.38 | 0.11 | 0.86 | 0.04 | **4.37** | 0.16 |
| **V29** | **2.48** | 0.06 | **7.87** | 0.07 | **0.81** | 0.03 | **3.17** | 0.08 | 2.18 | 0.06 | 8.86 | 0.08 | 0.83 | 0.04 | **4.06** | 0.12 |
| **F30** | **2.54** | 0.07 | **8.01** | 0.12 | **0.80** | 0.02 | **3.15** | 0.10 | 2.28 | 0.08 | 9.36 | 0.12 | 0.82 | 0.04 | **4.10** | 0.15 |
| **K31** | **2.51** | 0.07 | **8.17** | 0.08 | **0.79** | 0.03 | **3.26** | 0.10 | 2.24 | 0.08 | 9.24 | 0.19 | 0.81 | 0.05 | **4.12** | 0.17 |
| **Q32** | **2.36** | 0.05 | **8.35** | 0.05 | **0.79** | 0.02 | **3.54** | 0.08 | 2.09 | 0.06 | 9.30 | 0.08 | 0.79 | 0.04 | **4.45** | 0.13 |
| **Y33** | **2.49** | 0.07 | **7.91** | 0.17 | **0.82** | 0.02 | **3.18** | 0.11 | 2.23 | 0.08 | 9.11 | 0.25 | 0.86 | 0.05 | **4.08** | 0.19 |
| **A34** | **2.57** | 0.07 | **7.96** | 0.08 | **0.80** | 0.03 | **3.09** | 0.09 | 2.25 | 0.07 | 9.19 | 0.08 | 0.81 | 0.05 | **4.08** | 0.13 |
| **N35** | **2.37** | 0.06 | **8.08** | 0.06 | **0.83** | 0.02 | **3.41** | 0.10 | 2.13 | 0.06 | 9.28 | 0.08 | 0.80 | 0.03 | **4.35** | 0.12 |
| **D36** | **2.47** | 0.07 | **7.99** | 0.06 | **0.83** | 0.02 | **3.24** | 0.10 | 2.18 | 0.08 | 9.07 | 0.10 | 0.83 | 0.04 | **4.15** | 0.15 |
| **N37** | **2.44** | 0.06 | **7.04** | 0.07 | **0.79** | 0.02 | **2.89** | 0.08 | 2.11 | 0.07 | 8.08 | 0.08 | 0.82 | 0.03 | **3.82** | 0.14 |
| **G38** | **2.53** | 0.07 | **7.81** | 0.08 | **0.77** | 0.02 | **3.09** | 0.09 | 2.27 | 0.07 | 8.77 | 0.07 | 0.78 | 0.03 | **3.86** | 0.12 |
| **V39** | **2.33** | 0.06 | **7.19** | 0.30 | **0.75** | 0.02 | **3.09** | 0.15 | 2.10 | 0.06 | 8.45 | 0.08 | 0.76 | 0.03 | **4.03** | 0.12 |
| **D40** | **2.05** | 0.04 | **7.05** | 0.08 | **0.66** | 0.02 | **3.44** | 0.08 | 1.84 | 0.04 | 8.33 | 0.05 | 0.63 | 0.03 | **4.54** | 0.10 |
| **G41** | **1.72** | 0.02 | **5.55** | 0.15 | **0.47** | 0.02 | **3.22** | 0.10 | 1.52 | 0.02 | 6.33 | 0.06 | 0.46 | 0.03 | **4.17** | 0.07 |
| **E42** | **2.43** | 0.05 | **7.24** | 0.14 | **0.78** | 0.02 | **2.97** | 0.09 | 2.13 | 0.06 | 8.00 | 0.15 | 0.77 | 0.03 | **3.75** | 0.12 |
| **W43** | **2.59** | 0.06 | **7.56** | 0.12 | **0.79** | 0.03 | **2.92** | 0.08 | 2.31 | 0.09 | 8.53 | 0.13 | 0.73 | 0.04 | **3.69** | 0.15 |
| **T44** | **2.63** | 0.08 | **6.93** | 0.06 | **0.81** | 0.03 | **2.63** | 0.08 | 2.38 | 0.11 | 7.82 | 0.15 | 0.86 | 0.05 | **3.29** | 0.16 |
| **Y45** | **2.76** | 0.07 | **6.94** | 0.07 | **0.82** | 0.03 | **2.52** | 0.07 | 2.51 | 0.11 | 7.80 | 0.13 | 0.81 | 0.04 | **3.11** | 0.14 |
| **D46** | **2.74** | 0.10 | **7.65** | 0.07 | **0.80** | 0.02 | **2.79** | 0.11 | 2.47 | 0.10 | 8.76 | 0.15 | 0.80 | 0.04 | **3.54** | 0.15 |
| **D47** | **2.88** | 0.08 | **6.70** | 0.10 | **0.80** | 0.03 | **2.33** | 0.07 | 2.60 | 0.11 | 7.46 | 0.08 | 0.84 | 0.05 | **2.87** | 0.12 |
| **A48** | **2.35** | 0.07 | **6.87** | 0.06 | **0.80** | 0.02 | **2.92** | 0.09 | 2.11 | 0.08 | 7.68 | 0.07 | 0.78 | 0.04 | **3.64** | 0.14 |
| **T49** | **2.51** | 0.08 | **6.62** | 0.13 | **0.81** | 0.02 | **2.64** | 0.10 | 2.23 | 0.10 | 7.59 | 0.06 | 0.80 | 0.03 | **3.41** | 0.16 |
| **K50** | **2.82** | 0.10 | **7.18** | 0.08 | **0.80** | 0.03 | **2.55** | 0.10 | 2.54 | 0.11 | 8.14 | 0.13 | 0.75 | 0.04 | **3.21** | 0.14 |
| **T51** | **2.69** | 0.10 | **6.70** | 0.04 | **0.77** | 0.02 | **2.49** | 0.10 | 2.36 | 0.14 | 7.57 | 0.09 | 0.83 | 0.03 | **3.22** | 0.19 |
| **F52** | **2.71** | 0.10 | **7.15** | 0.19 | **0.81** | 0.03 | **2.64** | 0.12 | 2.58 | 0.11 | 8.14 | 0.14 | 0.74 | 0.05 | **3.16** | 0.15 |
| **T53** | **2.63** | 0.09 | **7.11** | 0.03 | **0.77** | 0.02 | **2.70** | 0.09 | 2.44 | 0.12 | 8.06 | 0.10 | 0.83 | 0.04 | **3.31** | 0.16 |
| **V54** | **2.74** | 0.08 | **7.14** | 0.07 | **0.78** | 0.03 | **2.61** | 0.08 | 2.45 | 0.11 | 8.08 | 0.15 | 0.80 | 0.04 | **3.30** | 0.16 |
| **T55** | **2.65** | 0.08 | **6.79** | 0.07 | **0.79** | 0.03 | **2.56** | 0.08 | 2.29 | 0.11 | 7.63 | 0.07 | 0.76 | 0.03 | **3.33** | 0.17 |
| **E56** | **2.61** | 0.07 | **6.89** | 0.14 | **0.72** | 0.02 | **2.64** | 0.09 | 2.34 | 0.09 | 7.69 | 0.10 | 0.72 | 0.04 | **3.28** | 0.13 |
| **W43sc** | **2.11** | 0.05 | **6.62** | 0.22 | **0.74** | 0.03 | **3.13** | 0.12 | 1.82 | 0.05 | 7.76 | 0.17 | 0.73 | 0.06 | **4.27** | 0.14 |

**1.5 M [Ch][Glu]**

| | $R_1$ ($s^{-1}$) | SE | $R_2$ ($s^{-1}$) | SE | NOE | SE | $R_2/R_1$ | SE |
|---|---|---|---|---|---|---|---|---|
| **Q2** | **1.64** | 0.03 | **12.19** | 0.16 | **0.77** | 0.06 | **7.42** | 0.18 |
| **Y3** | **1.62** | 0.08 | **12.01** | 0.50 | **0.74** | 0.05 | **7.41** | 0.47 |
| **K4** | **1.70** | 0.06 | **13.65** | 1.05 | **0.89** | 0.07 | **8.04** | 0.68 |
| **L5** | **1.70** | 0.05 | **12.52** | 0.18 | **0.92** | 0.06 | **7.37** | 0.26 |
| **I6** | **1.70** | 0.08 | **12.52** | 0.32 | **0.73** | 0.06 | **7.36** | 0.38 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **L7** | **1.59** | 0.05 | **12.22** | 0.22 | **0.83** | 0.07 | **7.71** | 0.29 |
| **N8** | **1.52** | 0.03 | **12.11** | 0.34 | **0.86** | 0.06 | **7.99** | 0.28 |
| **G9** | **1.55** | 0.04 | **12.17** | 0.19 | **0.67** | 0.05 | **7.87** | 0.23 |
| **K10** | **1.64** | 0.02 | **11.87** | 0.15 | **0.68** | 0.06 | **7.24** | 0.14 |
| **T11** | **1.61** | 0.02 | **11.28** | 0.19 | **0.63** | 0.05 | **6.99** | 0.15 |
| **L12** | **1.45** | 0.04 | **9.12** | 0.06 | **0.49** | 0.04 | **6.28** | 0.17 |
| **K13** | **1.44** | 0.03 | **11.02** | 0.19 | **0.65** | 0.05 | **7.63** | 0.19 |
| **G14** | **1.33** | 0.02 | **11.31** | 0.10 | **0.63** | 0.04 | **8.48** | 0.17 |
| **E15** | **1.53** | 0.03 | **11.76** | 0.10 | **0.69** | 0.04 | **7.67** | 0.16 |
| **T16** | **1.51** | 0.05 | **12.62** | 0.25 | **0.82** | 0.05 | **8.37** | 0.30 |
| **T17** | **1.56** | 0.05 | **11.55** | 0.11 | **0.81** | 0.05 | **7.41** | 0.25 |
| **T18** | **1.60** | 0.04 | **12.30** | 0.33 | **0.81** | 0.06 | **7.69** | 0.30 |
| **E19** | **1.54** | 0.03 | **11.73** | 0.18 | **0.74** | 0.05 | **7.61** | 0.21 |
| **A20** | **1.56** | 0.06 | **12.40** | 0.23 | **0.74** | 0.06 | **7.96** | 0.35 |
| **V21** | **1.66** | 0.03 | **13.03** | 0.25 | **0.70** | 0.06 | **7.86** | 0.19 |
| **D22** | **1.51** | 0.04 | **12.97** | 0.18 | **0.73** | 0.04 | **8.56** | 0.25 |
| **A23** | **1.61** | 0.06 | **14.40** | 0.54 | **0.87** | 0.07 | **8.96** | 0.45 |
| **A24** | **1.35** | 0.03 | **14.32** | 0.12 | **0.76** | 0.04 | **10.60** | 0.26 |
| **T25** | **1.48** | 0.04 | **15.34** | 0.40 | **0.86** | 0.08 | **10.34** | 0.39 |
| **A26** | **1.59** | 0.05 | **14.22** | 0.19 | **0.78** | 0.05 | **8.97** | 0.30 |
| **E27** | **1.46** | 0.04 | **13.60** | 0.81 | **0.75** | 0.03 | **9.33** | 0.61 |
| **K28** | **1.38** | 0.03 | **14.95** | 0.38 | **0.80** | 0.06 | **10.81** | 0.37 |
| **V29** | **1.44** | 0.04 | **14.76** | 0.34 | **0.72** | 0.06 | **10.28** | 0.36 |
| **F30** | **1.50** | 0.04 | **15.39** | 0.17 | **0.79** | 0.05 | **10.25** | 0.30 |
| **K31** | **1.47** | 0.05 | **15.02** | 0.46 | **0.79** | 0.06 | **10.20** | 0.47 |
| **Q32** | **1.38** | 0.03 | **15.46** | 0.33 | **0.80** | 0.05 | **11.20** | 0.36 |
| **Y33** | **1.47** | 0.04 | **15.01** | 0.52 | **0.75** | 0.05 | **10.24** | 0.43 |
| **A34** | **1.52** | 0.03 | **14.90** | 0.29 | **0.76** | 0.06 | **9.80** | 0.27 |
| **N35** | **1.37** | 0.03 | **15.02** | 0.51 | **0.72** | 0.05 | **10.94** | 0.46 |
| **D36** | **1.43** | 0.03 | **14.86** | 0.15 | **0.87** | 0.05 | **10.42** | 0.27 |
| **N37** | **1.43** | 0.04 | **12.67** | 0.19 | **0.76** | 0.04 | **8.85** | 0.26 |
| **G38** | **1.49** | 0.03 | **14.37** | 0.23 | **0.79** | 0.05 | **9.66** | 0.25 |
| **V39** | **1.42** | 0.02 | **14.11** | 0.16 | **0.82** | 0.05 | **9.90** | 0.19 |
| **D40** | **1.28** | 0.03 | **12.56** | 0.24 | **0.62** | 0.05 | **9.84** | 0.28 |
| **G41** | **1.15** | 0.01 | **9.78** | 0.16 | **0.37** | 0.04 | **8.51** | 0.17 |
| **E42** | **1.47** | 0.04 | **13.25** | 0.24 | **0.72** | 0.05 | **9.04** | 0.29 |
| **W43** | **1.58** | 0.05 | **12.74** | 0.29 | **0.78** | 0.07 | **8.05** | 0.31 |
| **T44** | **1.62** | 0.05 | **12.38** | 0.16 | **0.79** | 0.05 | **7.62** | 0.27 |
| **Y45** | **1.72** | 0.04 | **12.58** | 0.20 | **0.81** | 0.05 | **7.29** | 0.21 |
| **D46** | **1.67** | 0.06 | **12.98** | 0.22 | **0.83** | 0.05 | **7.78** | 0.31 |
| **D47** | **1.74** | 0.04 | **11.96** | 0.24 | **0.81** | 0.06 | **6.86** | 0.22 |
| **A48** | **1.39** | 0.03 | **12.44** | 0.11 | **0.66** | 0.04 | **8.92** | 0.23 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **T49** | **1.45** | 0.05 | **11.86** | 0.18 | **0.76** | 0.05 | **8.18** | 0.30 |
| **K50** | **1.75** | 0.06 | **12.68** | 0.16 | **0.77** | 0.06 | **7.26** | 0.28 |
| **T51** | **1.60** | 0.05 | **12.18** | 0.23 | **0.86** | 0.05 | **7.63** | 0.28 |
| **F52** | **1.76** | 0.09 | **13.53** | 0.56 | **0.90** | 0.09 | **7.69** | 0.52 |
| **T53** | **1.63** | 0.07 | **12.40** | 0.30 | **0.86** | 0.06 | **7.62** | 0.36 |
| **V54** | **1.71** | 0.06 | **13.53** | 0.45 | **0.85** | 0.06 | **7.89** | 0.40 |
| **T55** | **1.57** | 0.06 | **12.40** | 0.16 | **0.78** | 0.04 | **7.88** | 0.31 |
| **E56** | **1.59** | 0.06 | **12.25** | 0.19 | **0.73** | 0.05 | **7.69** | 0.32 |
| **W43sc** | **1.22** | 0.03 | **12.33** | 0.70 | **0.65** | 0.06 | **10.08** | 0.62 |

Longitudinal ($R_1$) and transverse relaxation ($R_2$) rate constants and {$^{1}$H}-$^{15}$N steady-state NOE values for 0.6 mM $^{13}$C$^{15}$N GB1 for water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentrations, at temperature of 298.2 K and $^{1}$H Larmor precession frequency of 600.13 MHz. Peak volumes were analysed with PINT and SE (standard error) is estimated from the fitting.

**Table A.8. Relaxation data for GB1 in water and in the presence of [Ch][Glu], at 310.2 K and 600 MHz.**

| | Water | | | | | | | | 0.1 M [Ch][Glu] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $R_1$ ($s^{-1}$) | SE | $R_2$ ($s^{-1}$) | SE | NOE | SE | $R_2/R_1$ | SE | $R_1$ ($s^{-1}$) | SE | $R_2$ ($s^{-1}$) | SE | NOE | SE | $R_2/R_1$ | SE |
| **Q2** | **3.02** | 0.65 | **3.60** | 0.55 | **0.64** | 0.03 | **1.19** | 0.31 | **3.42** | 0.26 | **4.08** | 0.53 | **0.65** | 0.04 | **1.19** | 0.18 |
| **Y3** | **2.96** | 0.09 | **3.43** | 0.02 | **0.58** | 0.01 | **1.16** | 0.04 | **3.01** | 0.07 | **3.63** | 0.02 | **0.63** | 0.01 | **1.20** | 0.03 |
| **K4** | **3.13** | 0.06 | **3.79** | 0.09 | **0.55** | 0.01 | **1.21** | 0.04 | **3.16** | 0.07 | **3.96** | 0.09 | **0.66** | 0.02 | **1.25** | 0.04 |
| **L5** | **3.09** | 0.07 | **3.76** | 0.10 | **0.64** | 0.01 | **1.22** | 0.04 | **3.18** | 0.07 | **3.96** | 0.13 | **0.62** | 0.01 | **1.25** | 0.05 |
| **I6** | **3.06** | 0.07 | **3.66** | 0.03 | **0.54** | 0.01 | **1.19** | 0.03 | **3.12** | 0.07 | **3.88** | 0.03 | **0.62** | 0.01 | **1.24** | 0.03 |
| **L7** | **2.96** | 0.07 | **3.57** | 0.02 | **0.62** | 0.01 | **1.20** | 0.03 | **3.04** | 0.06 | **3.75** | 0.01 | **0.61** | 0.01 | **1.24** | 0.03 |
| **N8** | **2.82** | 0.07 | **3.57** | 0.12 | **0.57** | 0.01 | **1.26** | 0.05 | **2.88** | 0.06 | **3.79** | 0.14 | **0.54** | 0.01 | **1.31** | 0.05 |
| **G9** | **2.94** | 0.09 | **3.53** | 0.04 | **0.57** | 0.01 | **1.20** | 0.04 | **2.96** | 0.07 | **3.74** | 0.04 | **0.59** | 0.01 | **1.26** | 0.03 |
| **K10** | **2.92** | 0.14 | **3.53** | 0.13 | **0.50** | 0.01 | **1.21** | 0.07 | **2.87** | 0.09 | **3.71** | 0.08 | **0.50** | 0.02 | **1.29** | 0.05 |
| **T11** | **3.36** | 0.48 | **3.85** | 0.74 | **0.55** | 0.01 | **1.15** | 0.28 | **2.87** | 0.18 | **4.14** | 0.30 | **0.52** | 0.02 | **1.44** | 0.14 |
| **L12** | **2.64** | 0.24 | **2.76** | 0.23 | **0.61** | 0.01 | **1.05** | 0.13 | **2.35** | 0.11 | **2.94** | 0.10 | **0.60** | 0.01 | **1.25** | 0.07 |
| **K13** | **2.70** | 0.10 | **3.25** | 0.07 | **0.62** | 0.01 | **1.20** | 0.05 | **2.70** | 0.07 | **3.46** | 0.07 | **0.66** | 0.01 | **1.28** | 0.04 |
| **G14** | **2.55** | 0.06 | **3.17** | 0.04 | **0.54** | 0.01 | **1.24** | 0.03 | **2.57** | 0.05 | **3.34** | 0.04 | **0.64** | 0.01 | **1.30** | 0.03 |
| **E15** | **3.32** | 0.45 | **3.48** | 0.45 | **0.69** | 0.01 | **1.05** | 0.20 | **2.97** | 0.18 | **3.70** | 0.22 | **0.69** | 0.02 | **1.24** | 0.11 |
| **T16** | **2.95** | 0.07 | **3.58** | 0.03 | **0.55** | 0.01 | **1.21** | 0.03 | **2.99** | 0.06 | **3.77** | 0.03 | **0.60** | 0.01 | **1.26** | 0.03 |
| **T17** | **3.44** | 0.47 | **3.54** | 0.40 | **0.67** | 0.01 | **1.03** | 0.18 | **3.12** | 0.19 | **3.76** | 0.17 | **0.72** | 0.01 | **1.21** | 0.09 |
| **T18** | **3.07** | 0.10 | **3.64** | 0.04 | **0.68** | 0.01 | **1.19** | 0.04 | **3.09** | 0.08 | **3.87** | 0.04 | **0.69** | 0.02 | **1.25** | 0.04 |
| **E19** | **3.52** | 0.51 | **3.57** | 0.56 | **0.66** | 0.01 | **1.01** | 0.22 | **3.12** | 0.20 | **3.84** | 0.29 | **0.63** | 0.02 | **1.23** | 0.12 |
| **A20** | **3.00** | 0.12 | **4.13** | 0.22 | **0.66** | 0.01 | **1.38** | 0.09 | **2.98** | 0.09 | **4.56** | 0.18 | **0.63** | 0.02 | **1.53** | 0.08 |
| **V21** | **3.78** | 0.58 | **3.96** | 0.65 | **0.65** | 0.02 | **1.05** | 0.23 | **3.37** | 0.23 | **4.06** | 0.38 | **0.65** | 0.03 | **1.21** | 0.14 |
| **D22** | **3.27** | 0.22 | **3.76** | 0.13 | **0.73** | 0.01 | **1.15** | 0.09 | **3.17** | 0.12 | **4.01** | 0.07 | **0.75** | 0.01 | **1.27** | 0.05 |
| **A23** | **3.97** | 0.58 | **4.20** | 0.35 | **0.71** | 0.02 | **1.06** | 0.18 | **3.52** | 0.22 | **4.33** | 0.28 | **0.68** | 0.02 | **1.23** | 0.11 |
| **A24** | **3.18** | 0.15 | **3.97** | 0.10 | **0.79** | 0.01 | **1.25** | 0.07 | **3.13** | 0.09 | **4.21** | 0.04 | **0.79** | 0.02 | **1.35** | 0.04 |
| **T25** | **3.30** | 0.20 | **3.99** | 0.12 | **0.72** | 0.01 | **1.21** | 0.08 | **3.23** | 0.12 | **4.18** | 0.07 | **0.69** | 0.02 | **1.29** | 0.05 |
| **A26** | **3.28** | 0.11 | **3.94** | 0.03 | **0.60** | 0.01 | **1.20** | 0.04 | **3.32** | 0.09 | **4.18** | 0.04 | **0.65** | 0.02 | **1.26** | 0.04 |
| **E27** | **3.32** | 0.08 | **4.14** | 0.01 | **0.71** | 0.01 | **1.25** | 0.03 | **3.36** | 0.08 | **4.38** | 0.02 | **0.70** | 0.01 | **1.30** | 0.03 |
| **K28** | **3.21** | 0.07 | **4.17** | 0.02 | **0.73** | 0.01 | **1.30** | 0.03 | **3.29** | 0.07 | **4.39** | 0.02 | **0.71** | 0.02 | **1.33** | 0.03 |
| **V29** | **3.18** | 0.07 | **4.03** | 0.03 | **0.72** | 0.01 | **1.27** | 0.03 | **3.24** | 0.06 | **4.24** | 0.01 | **0.73** | 0.02 | **1.31** | 0.03 |
| **F30** | **3.35** | 0.08 | **4.13** | 0.03 | **0.61** | 0.01 | **1.23** | 0.03 | **3.41** | 0.08 | **4.37** | 0.04 | **0.65** | 0.01 | **1.28** | 0.03 |
| **K31** | **3.32** | 0.08 | **4.14** | 0.03 | **0.67** | 0.01 | **1.25** | 0.03 | **3.39** | 0.09 | **4.41** | 0.04 | **0.71** | 0.02 | **1.30** | 0.04 |
| **Q32** | **3.19** | 0.07 | **4.22** | 0.02 | **0.64** | 0.01 | **1.32** | 0.03 | **3.25** | 0.07 | **4.45** | 0.04 | **0.68** | 0.01 | **1.37** | 0.03 |
| **Y33** | **3.27** | 0.07 | **4.18** | 0.01 | **0.66** | 0.01 | **1.28** | 0.03 | **3.34** | 0.07 | **4.42** | 0.04 | **0.63** | 0.01 | **1.32** | 0.03 |
| **A34** | **3.33** | 0.09 | **4.15** | 0.03 | **0.59** | 0.01 | **1.25** | 0.03 | **3.40** | 0.08 | **4.36** | 0.04 | **0.63** | 0.02 | **1.28** | 0.03 |
| **N35** | **3.15** | 0.08 | **4.17** | 0.04 | **0.69** | 0.01 | **1.32** | 0.04 | **3.18** | 0.07 | **4.41** | 0.04 | **0.69** | 0.01 | **1.38** | 0.03 |
| **D36** | **3.26** | 0.10 | **4.10** | 0.03 | **0.71** | 0.01 | **1.26** | 0.04 | **3.30** | 0.09 | **4.32** | 0.03 | **0.67** | 0.01 | **1.31** | 0.04 |
| **N37** | **3.07** | 0.09 | **3.69** | 0.06 | **0.63** | 0.01 | **1.20** | 0.04 | **3.09** | 0.08 | **3.93** | 0.05 | **0.70** | 0.01 | **1.27** | 0.04 |
| **G38** | **3.18** | 0.11 | **4.19** | 0.15 | **0.57** | 0.01 | **1.32** | 0.07 | **3.21** | 0.09 | **4.64** | 0.18 | **0.60** | 0.01 | **1.45** | 0.07 |
| **V39** | **3.02** | 0.09 | **3.86** | 0.02 | **0.63** | 0.01 | **1.28** | 0.04 | **3.09** | 0.08 | **4.07** | 0.04 | **0.70** | 0.02 | **1.32** | 0.04 |
| **D40** | **2.56** | 0.06 | **4.23** | 0.16 | **0.48** | 0.01 | **1.65** | 0.07 | **2.56** | 0.05 | **5.14** | 0.18 | **0.50** | 0.01 | **2.01** | 0.08 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **G41** | **1.97** | 0.05 | **3.30** | 0.13 | **0.35** | 0.01 | **1.68** | 0.08 | **1.91** | 0.03 | **3.93** | 0.17 | **0.47** | 0.01 | **2.06** | 0.09 |
| **E42** | **2.85** | 0.07 | **3.62** | 0.03 | **0.62** | 0.01 | **1.27** | 0.03 | **2.91** | 0.07 | **3.84** | 0.05 | **0.53** | 0.01 | **1.32** | 0.04 |
| **W43** | **3.02** | 0.09 | **4.84** | 0.19 | **0.63** | 0.01 | **1.60** | 0.08 | **3.03** | 0.08 | **5.61** | 0.17 | **0.62** | 0.01 | **1.85** | 0.07 |
| **T44** | **3.10** | 0.09 | **3.64** | 0.04 | **0.62** | 0.01 | **1.17** | 0.04 | **3.11** | 0.08 | **3.82** | 0.02 | **0.62** | 0.01 | **1.23** | 0.03 |
| **Y45** | **3.62** | 0.51 | **3.72** | 0.46 | **0.60** | 0.01 | **1.03** | 0.19 | **3.19** | 0.21 | **3.84** | 0.18 | **0.61** | 0.02 | **1.20** | 0.10 |
| **D46** | **3.15** | 0.11 | **4.86** | 0.21 | **0.68** | 0.01 | **1.54** | 0.09 | **3.14** | 0.09 | **5.59** | 0.22 | **0.69** | 0.01 | **1.78** | 0.08 |
| **D47** | **3.83** | 0.61 | **3.57** | 0.60 | **0.64** | 0.02 | **0.93** | 0.22 | **3.36** | 0.24 | **3.86** | 0.38 | **0.65** | 0.03 | **1.15** | 0.14 |
| **A48** | **3.27** | 0.37 | **3.68** | 0.36 | **0.74** | 0.01 | **1.13** | 0.17 | **3.00** | 0.16 | **3.90** | 0.16 | **0.71** | 0.02 | **1.30** | 0.09 |
| **T49** | **3.08** | 0.20 | **3.55** | 0.10 | **0.66** | 0.01 | **1.15** | 0.08 | **3.01** | 0.12 | **3.83** | 0.05 | **0.73** | 0.01 | **1.27** | 0.05 |
| **K50** | **3.21** | 0.14 | **3.70** | 0.10 | **0.57** | 0.01 | **1.15** | 0.06 | **3.25** | 0.10 | **3.93** | 0.08 | **0.60** | 0.02 | **1.21** | 0.05 |
| **T51** | **3.19** | 0.15 | **3.54** | 0.07 | **0.66** | 0.01 | **1.11** | 0.06 | **3.18** | 0.11 | **3.74** | 0.03 | **0.68** | 0.01 | **1.17** | 0.04 |
| **F52** | **3.22** | 0.08 | **4.10** | 0.08 | **0.61** | 0.01 | **1.27** | 0.04 | **3.24** | 0.07 | **4.76** | 0.11 | **0.67** | 0.02 | **1.47** | 0.05 |
| **T53** | **3.13** | 0.10 | **3.67** | 0.05 | **0.57** | 0.01 | **1.17** | 0.04 | **3.15** | 0.08 | **3.88** | 0.05 | **0.67** | 0.01 | **1.23** | 0.04 |
| **V54** | **3.10** | 0.09 | **3.71** | 0.05 | **0.52** | 0.01 | **1.20** | 0.04 | **3.19** | 0.08 | **3.93** | 0.06 | **0.60** | 0.01 | **1.23** | 0.04 |
| **T55** | **3.00** | 0.10 | **3.56** | 0.04 | **0.62** | 0.01 | **1.18** | 0.04 | **3.04** | 0.08 | **3.82** | 0.03 | **0.69** | 0.01 | **1.26** | 0.04 |
| **E56** | **2.86** | 0.07 | **3.65** | 0.05 | **0.48** | 0.01 | **1.28** | 0.04 | **2.95** | 0.06 | **3.79** | 0.03 | **0.62** | 0.01 | **1.28** | 0.03 |
| **W43sc** | **2.74** | 0.12 | **3.87** | 0.25 | **0.50** | 0.01 | **1.41** | 0.11 | **2.65** | 0.08 | **4.54** | 0.23 | **0.53** | 0.02 | **1.71** | 0.10 |
| | | | **0.75 M [Ch][Glu]** | | | | | | | | **1.0 M [Ch][Glu]** | | | | | |
| | **$R_1$ ($s^{-1}$)** | SE | **$R_2$ ($s^{-1}$)** | SE | **NOE** | SE | **$R_2/R_1$** | SE | **$R_1$ ($s^{-1}$)** | SE | **$R_2$ ($s^{-1}$)** | SE | **NOE** | SE | **$R_2/R_1$** | SE |
| **Q2** | **-** | - | **-** | - | **-** | - | **-** | - | - | - | - | - | - | - | **-** | - |
| **Y3** | **2.91** | 0.10 | **5.30** | 0.05 | **0.75** | 0.02 | **1.82** | 0.06 | **2.82** | 0.14 | **5.81** | 0.08 | **0.74** | 0.04 | **2.06** | 0.11 |
| **K4** | **3.01** | 0.10 | **5.90** | 0.25 | **0.77** | 0.02 | **1.96** | 0.10 | **2.94** | 0.14 | **6.37** | 0.30 | **0.80** | 0.05 | **2.17** | 0.14 |
| **L5** | **3.02** | 0.09 | **5.76** | 0.09 | **0.73** | 0.02 | **1.91** | 0.07 | **2.93** | 0.13 | **6.11** | 0.06 | **0.76** | 0.04 | **2.08** | 0.09 |
| **I6** | **2.93** | 0.10 | **5.57** | 0.07 | **0.73** | 0.02 | **1.90** | 0.07 | **2.95** | 0.13 | **6.02** | 0.09 | **0.74** | 0.05 | **2.04** | 0.10 |
| **L7** | **2.89** | 0.09 | **5.46** | 0.08 | **0.68** | 0.02 | **1.89** | 0.06 | **2.85** | 0.08 | **5.91** | 0.10 | **0.85** | 0.05 | **2.07** | 0.07 |
| **N8** | **2.73** | 0.08 | **5.45** | 0.06 | **0.75** | 0.02 | **2.00** | 0.06 | **2.66** | 0.10 | **5.80** | 0.06 | **0.71** | 0.04 | **2.18** | 0.09 |
| **G9** | **2.82** | 0.08 | **5.42** | 0.05 | **0.73** | 0.02 | **1.92** | 0.06 | **2.73** | 0.12 | **5.86** | 0.12 | **0.78** | 0.03 | **2.15** | 0.10 |
| **K10** | **2.84** | 0.07 | **5.41** | 0.08 | **0.72** | 0.02 | **1.90** | 0.06 | **2.73** | 0.11 | **5.93** | 0.07 | **0.71** | 0.05 | **2.17** | 0.09 |
| **T11** | **2.90** | 0.11 | **5.72** | 0.23 | **0.68** | 0.02 | **1.97** | 0.11 | **2.70** | 0.11 | **6.03** | 0.32 | **0.59** | 0.04 | **2.23** | 0.15 |
| **L12** | **2.41** | 0.08 | **4.16** | 0.07 | **0.61** | 0.02 | **1.73** | 0.06 | **2.27** | 0.11 | **4.50** | 0.05 | **0.61** | 0.03 | **1.99** | 0.10 |
| **K13** | **2.64** | 0.08 | **4.93** | 0.11 | **0.65** | 0.02 | **1.86** | 0.07 | **2.50** | 0.07 | **5.44** | 0.07 | **0.64** | 0.03 | **2.17** | 0.07 |
| **G14** | **2.46** | 0.06 | **4.92** | 0.07 | **0.68** | 0.01 | **2.00** | 0.05 | **2.35** | 0.07 | **5.29** | 0.10 | **0.67** | 0.03 | **2.25** | 0.08 |
| **E15** | **3.00** | 0.14 | **5.44** | 0.11 | **0.70** | 0.02 | **1.81** | 0.09 | **2.80** | 0.15 | **5.87** | 0.31 | **0.65** | 0.04 | **2.10** | 0.15 |
| **T16** | **2.89** | 0.08 | **5.49** | 0.03 | **0.73** | 0.02 | **1.90** | 0.06 | **2.72** | 0.11 | **5.94** | 0.06 | **0.75** | 0.04 | **2.18** | 0.09 |
| **T17** | **3.12** | 0.16 | **5.37** | 0.09 | **0.74** | 0.02 | **1.72** | 0.09 | **3.02** | 0.19 | **5.90** | 0.20 | **0.75** | 0.04 | **1.95** | 0.14 |
| **T18** | **2.96** | 0.11 | **5.70** | 0.07 | **0.75** | 0.02 | **1.92** | 0.07 | **2.85** | 0.13 | **5.93** | 0.11 | **0.79** | 0.04 | **2.08** | 0.10 |
| **E19** | **3.10** | 0.13 | **5.49** | 0.12 | **0.75** | 0.02 | **1.77** | 0.08 | **2.87** | 0.16 | **5.85** | 0.34 | **0.78** | 0.04 | **2.04** | 0.16 |
| **A20** | **2.85** | 0.11 | **5.75** | 0.11 | **0.76** | 0.02 | **2.01** | 0.09 | **2.83** | 0.12 | **6.30** | 0.14 | **0.73** | 0.04 | **2.23** | 0.11 |
| **V21** | **3.32** | 0.18 | **6.09** | 0.15 | **0.79** | 0.02 | **1.83** | 0.11 | **3.09** | 0.19 | **6.15** | 0.39 | **0.80** | 0.06 | **1.99** | 0.18 |
| **D22** | **2.94** | 0.12 | **5.82** | 0.05 | **0.73** | 0.02 | **1.98** | 0.08 | **2.78** | 0.13 | **6.27** | 0.10 | **0.72** | 0.03 | **2.26** | 0.11 |
| **A23** | **3.39** | 0.14 | **6.41** | 0.08 | **0.78** | 0.02 | **1.89** | 0.08 | **3.19** | 0.17 | **6.65** | 0.18 | **0.84** | 0.05 | **2.08** | 0.13 |
| **A24** | **2.77** | 0.09 | **6.19** | 0.10 | **0.73** | 0.02 | **2.24** | 0.08 | **2.59** | 0.10 | **6.75** | 0.16 | **0.71** | 0.03 | **2.61** | 0.12 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **T25** | **2.96** | 0.14 | **7.33** | 0.35 | **0.64** | 0.04 | **2.47** | 0.16 | **2.79** | 0.14 | **6.76** | 0.09 | **0.79** | 0.05 | **2.42** | 0.12 |
| **A26** | **3.08** | 0.10 | **6.14** | 0.08 | **0.82** | 0.02 | **1.99** | 0.07 | **2.88** | 0.14 | **6.65** | 0.09 | **0.79** | 0.04 | **2.31** | 0.12 |
| **E27** | **2.95** | 0.09 | **6.28** | 0.20 | **0.81** | 0.01 | **2.12** | 0.10 | **2.83** | 0.10 | **6.93** | 0.04 | **0.77** | 0.03 | **2.45** | 0.09 |
| **K28** | **2.84** | 0.08 | **6.55** | 0.03 | **0.76** | 0.02 | **2.30** | 0.06 | **2.69** | 0.08 | **7.17** | 0.04 | **0.81** | 0.04 | **2.66** | 0.08 |
| **V29** | **2.87** | 0.07 | **6.33** | 0.05 | **0.75** | 0.02 | **2.21** | 0.06 | **2.73** | 0.08 | **6.90** | 0.10 | **0.77** | 0.04 | **2.53** | 0.08 |
| **F30** | **2.97** | 0.09 | **6.50** | 0.10 | **0.79** | 0.02 | **2.19** | 0.07 | **2.85** | 0.09 | **7.09** | 0.08 | **0.83** | 0.04 | **2.49** | 0.08 |
| **K31** | **2.93** | 0.08 | **6.38** | 0.10 | **0.74** | 0.02 | **2.18** | 0.07 | **2.78** | 0.10 | **7.02** | 0.12 | **0.82** | 0.05 | **2.53** | 0.10 |
| **Q32** | **2.78** | 0.08 | **6.50** | 0.04 | **0.77** | 0.02 | **2.34** | 0.06 | **2.63** | 0.08 | **7.16** | 0.06 | **0.75** | 0.03 | **2.72** | 0.08 |
| **Y33** | **2.91** | 0.08 | **6.29** | 0.59 | **0.76** | 0.01 | **2.16** | 0.21 | **2.82** | 0.10 | **6.93** | 0.07 | **0.82** | 0.03 | **2.45** | 0.09 |
| **A34** | **3.02** | 0.08 | **6.48** | 0.09 | **0.78** | 0.02 | **2.14** | 0.07 | **2.80** | 0.09 | **6.98** | 0.09 | **0.83** | 0.05 | **2.49** | 0.09 |
| **N35** | **2.76** | 0.08 | **6.38** | 0.07 | **0.75** | 0.02 | **2.31** | 0.07 | **2.62** | 0.09 | **7.00** | 0.11 | **0.75** | 0.04 | **2.67** | 0.10 |
| **D36** | **2.91** | 0.09 | **6.34** | 0.03 | **0.80** | 0.02 | **2.18** | 0.07 | **2.78** | 0.11 | **6.88** | 0.06 | **0.80** | 0.04 | **2.48** | 0.10 |
| **N37** | **2.76** | 0.10 | **5.69** | 0.06 | **0.75** | 0.01 | **2.06** | 0.07 | **2.69** | 0.11 | **6.10** | 0.08 | **0.78** | 0.03 | **2.27** | 0.10 |
| **G38** | **2.91** | 0.10 | **6.32** | 0.11 | **0.78** | 0.02 | **2.18** | 0.09 | **2.83** | 0.12 | **7.00** | 0.19 | **0.80** | 0.03 | **2.48** | 0.13 |
| **V39** | **2.76** | 0.07 | **5.97** | 0.08 | **0.72** | 0.02 | **2.16** | 0.06 | **2.62** | 0.07 | **6.34** | 0.07 | **0.82** | 0.04 | **2.42** | 0.07 |
| **D40** | **2.36** | 0.04 | **5.81** | 0.09 | **0.64** | 0.01 | **2.47** | 0.06 | **2.21** | 0.05 | **6.68** | 0.09 | **0.62** | 0.03 | **3.03** | 0.07 |
| **G41** | **1.89** | 0.03 | **4.39** | 0.11 | **0.42** | 0.01 | **2.32** | 0.07 | **1.75** | 0.04 | **5.04** | 0.09 | **0.43** | 0.03 | **2.89** | 0.08 |
| **E42** | **2.71** | 0.06 | **5.95** | 0.19 | **0.82** | 0.02 | **2.20** | 0.09 | **2.61** | 0.08 | **6.30** | 0.09 | **0.79** | 0.05 | **2.42** | 0.08 |
| **W43** | **2.92** | 0.11 | **6.24** | 0.13 | **0.75** | 0.02 | **2.14** | 0.09 | **2.88** | 0.14 | **7.23** | 0.26 | **0.76** | 0.04 | **2.52** | 0.15 |
| **T44** | **2.96** | 0.12 | **5.64** | 0.08 | **0.77** | 0.02 | **1.91** | 0.08 | **2.89** | 0.15 | **6.01** | 0.07 | **0.80** | 0.04 | **2.08** | 0.11 |
| **Y45** | **3.37** | 0.15 | **5.72** | 0.10 | **0.78** | 0.02 | **1.69** | 0.08 | **3.16** | 0.21 | **6.19** | 0.41 | **0.85** | 0.05 | **1.96** | 0.19 |
| **D46** | **3.04** | 0.11 | **6.27** | 0.11 | **0.72** | 0.02 | **2.06** | 0.08 | **3.03** | 0.17 | **7.20** | 0.10 | **0.76** | 0.04 | **2.37** | 0.13 |
| **D47** | **3.44** | 0.22 | **5.69** | 0.33 | **0.78** | 0.02 | **1.65** | 0.14 | **3.30** | 0.29 | **5.95** | 0.21 | **0.79** | 0.06 | **1.80** | 0.17 |
| **A48** | **2.89** | 0.14 | **5.65** | 0.11 | **0.74** | 0.02 | **1.96** | 0.10 | **2.71** | 0.15 | **6.07** | 0.18 | **0.78** | 0.04 | **2.24** | 0.14 |
| **T49** | **2.93** | 0.09 | **5.39** | 0.07 | **0.73** | 0.02 | **1.84** | 0.06 | **2.76** | 0.17 | **5.79** | 0.08 | **0.80** | 0.03 | **2.10** | 0.13 |
| **K50** | **3.15** | 0.12 | **5.71** | 0.08 | **0.76** | 0.02 | **1.81** | 0.08 | **3.05** | 0.16 | **6.18** | 0.09 | **0.84** | 0.04 | **2.02** | 0.11 |
| **T51** | **3.05** | 0.14 | **5.51** | 0.06 | **0.73** | 0.02 | **1.81** | 0.09 | **2.95** | 0.18 | **5.89** | 0.14 | **0.75** | 0.04 | **2.00** | 0.13 |
| **F52** | **3.12** | 0.10 | **5.96** | 0.06 | **0.77** | 0.02 | **1.91** | 0.06 | **3.19** | 0.16 | **6.93** | 0.18 | **0.77** | 0.05 | **2.17** | 0.12 |
| **T53** | **2.99** | 0.09 | **5.69** | 0.03 | **0.73** | 0.02 | **1.90** | 0.06 | **2.93** | 0.16 | **6.19** | 0.16 | **0.79** | 0.04 | **2.11** | 0.13 |
| **V54** | **3.07** | 0.10 | **5.65** | 0.08 | **0.77** | 0.02 | **1.84** | 0.07 | **3.01** | 0.15 | **6.14** | 0.08 | **0.79** | 0.04 | **2.04** | 0.10 |
| **T55** | **3.01** | 0.11 | **5.59** | 0.09 | **0.72** | 0.01 | **1.86** | 0.07 | **2.86** | 0.13 | **6.06** | 0.10 | **0.72** | 0.02 | **2.12** | 0.11 |
| **E56** | **2.89** | 0.09 | **5.56** | 0.12 | **0.74** | 0.02 | **1.92** | 0.07 | **2.77** | 0.10 | **5.84** | 0.03 | **0.70** | 0.04 | **2.11** | 0.08 |
| **W43sc** | **2.50** | 0.08 | **5.52** | 0.16 | **0.74** | 0.02 | **2.20** | 0.09 | **2.36** | 0.11 | **6.18** | 0.16 | **0.75** | 0.06 | **2.62** | 0.14 |

**1.5 M [Ch][Glu]**

| | $R_1$ ($s^{-1}$) | SE | $R_2$ ($s^{-1}$) | SE | NOE | SE | $R_2/R_1$ | SE |
|---|---|---|---|---|---|---|---|---|
| **Q2** | **-** | - | **-** | - | **-** | - | **-** | - |
| **Y3** | **2.17** | 0.08 | **8.81** | 0.20 | **0.73** | 0.05 | **4.06** | 0.18 |
| **K4** | **2.16** | 0.09 | **9.71** | 0.43 | **0.76** | 0.07 | **4.49** | 0.27 |
| **L5** | **2.27** | 0.10 | **9.15** | 0.19 | **0.80** | 0.06 | **4.04** | 0.20 |
| **I6** | **2.24** | 0.10 | **8.94** | 0.17 | **0.80** | 0.06 | **4.00** | 0.19 |
| **L7** | **2.12** | 0.10 | **8.91** | 0.07 | **0.79** | 0.06 | **4.20** | 0.19 |
| **N8** | **2.01** | 0.06 | **8.70** | 0.10 | **0.79** | 0.05 | **4.33** | 0.13 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **G9** | **2.02** | 0.07 | **8.72** | 0.16 | **0.72** | 0.05 | **4.31** | 0.17 |
| **K10** | **2.11** | 0.06 | **8.54** | 0.10 | **0.70** | 0.06 | **4.04** | 0.13 |
| **T11** | **2.09** | 0.05 | **8.58** | 0.11 | **0.58** | 0.05 | **4.10** | 0.12 |
| **L12** | **1.85** | 0.05 | **6.71** | 0.04 | **0.60** | 0.04 | **3.63** | 0.10 |
| **K13** | **1.89** | 0.05 | **8.07** | 0.06 | **0.74** | 0.05 | **4.28** | 0.13 |
| **G14** | **1.78** | 0.02 | **8.12** | 0.11 | **0.65** | 0.04 | **4.57** | 0.09 |
| **E15** | **2.02** | 0.05 | **8.84** | 0.16 | **0.72** | 0.05 | **4.37** | 0.14 |
| **T16** | **2.07** | 0.06 | **8.90** | 0.14 | **0.77** | 0.05 | **4.30** | 0.14 |
| **T17** | **2.17** | 0.11 | **8.64** | 0.09 | **0.77** | 0.04 | **3.97** | 0.21 |
| **T18** | **2.16** | 0.09 | **9.04** | 0.21 | **0.83** | 0.05 | **4.20** | 0.20 |
| **E19** | **2.12** | 0.08 | **8.42** | 0.14 | **0.75** | 0.05 | **3.97** | 0.17 |
| **A20** | **2.05** | 0.08 | **8.71** | 0.14 | **0.84** | 0.06 | **4.26** | 0.18 |
| **V21** | **2.28** | 0.07 | **9.10** | 0.24 | **0.78** | 0.07 | **4.00** | 0.16 |
| **D22** | **2.08** | 0.09 | **9.44** | 0.11 | **0.80** | 0.04 | **4.54** | 0.20 |
| **A23** | **2.35** | 0.09 | **10.12** | 0.16 | **0.84** | 0.07 | **4.31** | 0.18 |
| **A24** | **1.81** | 0.06 | **10.31** | 0.10 | **0.80** | 0.05 | **5.69** | 0.19 |
| **T25** | **2.05** | 0.08 | **10.36** | 0.22 | **0.82** | 0.05 | **5.06** | 0.22 |
| **A26** | **2.08** | 0.09 | **9.97** | 0.16 | **0.80** | 0.05 | **4.78** | 0.22 |
| **E27** | **2.05** | 0.07 | **10.59** | 0.15 | **0.81** | 0.04 | **5.18** | 0.18 |
| **K28** | **1.90** | 0.05 | **10.99** | 0.15 | **0.84** | 0.05 | **5.77** | 0.18 |
| **V29** | **1.99** | 0.06 | **10.46** | 0.11 | **0.78** | 0.05 | **5.26** | 0.16 |
| **F30** | **2.05** | 0.06 | **10.72** | 0.10 | **0.82** | 0.05 | **5.22** | 0.17 |
| **K31** | **1.99** | 0.05 | **10.75** | 0.13 | **0.82** | 0.06 | **5.40** | 0.15 |
| **Q32** | **1.87** | 0.04 | **10.87** | 0.18 | **0.83** | 0.05 | **5.82** | 0.16 |
| **Y33** | **1.99** | 0.06 | **10.85** | 0.23 | **0.80** | 0.05 | **5.45** | 0.21 |
| **A34** | **2.05** | 0.05 | **10.86** | 0.14 | **0.73** | 0.05 | **5.30** | 0.15 |
| **N35** | **1.88** | 0.04 | **10.85** | 0.09 | **0.83** | 0.05 | **5.79** | 0.13 |
| **D36** | **1.95** | 0.06 | **10.66** | 0.11 | **0.79** | 0.05 | **5.46** | 0.18 |
| **N37** | **1.96** | 0.04 | **9.31** | 0.15 | **0.75** | 0.04 | **4.75** | 0.13 |
| **G38** | **2.04** | 0.06 | **10.20** | 0.10 | **0.78** | 0.05 | **5.01** | 0.15 |
| **V39** | **1.89** | 0.06 | **9.75** | 0.10 | **0.79** | 0.04 | **5.15** | 0.16 |
| **D40** | **1.61** | 0.03 | **9.42** | 0.10 | **0.71** | 0.04 | **5.85** | 0.13 |
| **G41** | **1.37** | 0.01 | **7.13** | 0.14 | **0.45** | 0.04 | **5.21** | 0.11 |
| **E42** | **1.94** | 0.05 | **9.47** | 0.16 | **0.76** | 0.04 | **4.88** | 0.16 |
| **W43** | **2.13** | 0.06 | **9.78** | 0.13 | **0.82** | 0.06 | **4.59** | 0.13 |
| **T44** | **2.15** | 0.09 | **8.83** | 0.11 | **0.79** | 0.05 | **4.12** | 0.18 |
| **Y45** | **2.31** | 0.07 | **9.39** | 0.17 | **0.84** | 0.05 | **4.07** | 0.14 |
| **D46** | **2.34** | 0.12 | **10.02** | 0.15 | **0.90** | 0.05 | **4.29** | 0.23 |
| **D47** | **2.33** | 0.12 | **8.61** | 0.36 | **0.77** | 0.06 | **3.70** | 0.25 |
| **A48** | **1.96** | 0.06 | **9.06** | 0.12 | **0.78** | 0.04 | **4.62** | 0.16 |
| **T49** | **2.05** | 0.09 | **8.36** | 0.15 | **0.82** | 0.05 | **4.08** | 0.19 |
| **K50** | **2.44** | 0.12 | **9.56** | 0.15 | **0.74** | 0.05 | **3.93** | 0.21 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **T51** | **2.21** | 0.11 | **8.72** | 0.10 | **0.79** | 0.04 | **3.95** | 0.20 |
| **F52** | **2.38** | 0.15 | **10.07** | 0.38 | **0.73** | 0.06 | **4.24** | 0.31 |
| **T53** | **2.24** | 0.10 | **9.21** | 0.16 | **0.81** | 0.05 | **4.12** | 0.20 |
| **V54** | **2.29** | 0.10 | **9.17** | 0.09 | **0.78** | 0.05 | **4.01** | 0.18 |
| **T55** | **2.16** | 0.07 | **9.13** | 0.10 | **0.75** | 0.02 | **4.23** | 0.15 |
| **E56** | **2.07** | 0.07 | **8.91** | 0.14 | **0.71** | 0.05 | **4.31** | 0.15 |
| **W43sc** | **1.65** | 0.04 | **9.19** | 0.21 | **0.76** | 0.08 | **5.57** | 0.18 |

Longitudinal ($R_1$) and transverse relaxation ($R_2$) rate constants and $\{^1H\}$-$^{15}N$ steady-state NOE values for 0.6 mM $^{13}C^{15}N$ GB1 for water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentrations, at temperature of 310.2 K and $^1H$ Larmor precession frequency of 600.13 MHz. Peak volumes were analysed with PINT and SE (standard error) is estimated from the fitting.

**Table A.9. Model-free data for GB1 in water and in the presence of [Ch][Glu], at 298.2 K (500 and 600 MHz).**

| | Water | | | | | | | 0.1 M [Ch][Glu] | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | model | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | model |
| **Q2** | | | | | | | 'm0' | | | | | | | 'm0' |
| **Y3** | 0.941 | 0.005 | | | | | 'm1' | 0.979 | 0.008 | | | | | 'm1' |
| **K4** | 0.971 | 0.008 | | | | | 'm1' | | | | | | | 'm0' |
| **L5** | 0.956 | 0.003 | | | | | 'm1' | | | | | | | 'm0' |
| **I6** | 0.979 | 0.003 | | | | | 'm1' | | | | | 0.219 | 0.054 | 'm9' |
| **L7** | 0.964 | 0.004 | | | | | 'm1' | 0.987 | 0.003 | | | | | 'm1' |
| **N8** | 0.940 | 0.016 | | | 0.212 | 0.088 | 'm3' | 0.939 | 0.016 | | | 0.601 | 0.093 | 'm3' |
| **G9** | 0.914 | 0.003 | | | | | 'm1' | 0.973 | 0.004 | | | | | 'm1' |
| **K10** | 0.878 | 0.006 | 0.086 | 0.027 | | | 'm2' | 0.960 | 0.009 | | | | | 'm1' |
| **T11** | 0.894 | 0.010 | | | | | 'm1' | 0.944 | 0.015 | | | | | 'm1' |
| **L12** | 0.722 | 0.005 | | | | | 'm1' | 0.755 | 0.009 | 0.019 | 0.009 | | | 'm2' |
| **K13** | 0.834 | 0.003 | | | | | 'm1' | 0.859 | 0.006 | | | | | 'm1' |
| **G14** | 0.835 | 0.003 | | | | | 'm1' | 0.826 | 0.003 | | | | | 'm1' |
| **E15** | 0.872 | 0.009 | | | | | 'm1' | 0.919 | 0.011 | | | | | 'm1' |
| **T16** | 0.927 | 0.002 | | | | | 'm1' | 0.962 | 0.006 | | | | | 'm1' |
| **T17** | 0.904 | 0.007 | | | | | 'm1' | 0.949 | 0.006 | | | | | 'm1' |
| **T18** | 0.957 | 0.004 | | | | | 'm1' | | | | | | | 'm0' |
| **E19** | 0.910 | 0.007 | | | | | 'm1' | 0.933 | 0.008 | | | | | 'm1' |
| **A20** | 0.804 | 0.018 | | | 0.915 | 0.117 | 'm3' | | | | | | | 'm0' |
| **V21** | | | | | | | 'm0' | | | | | 0.121 | 0.073 | 'm9' |
| **D22** | 0.961 | 0.005 | | | | | 'm1' | | | | | | | 'm0' |
| **A23** | | | | | | | 'm0' | | | | | 0.280 | 0.066 | 'm9' |
| **A24** | 0.963 | 0.004 | | | | | 'm1' | 0.977 | 0.006 | | | | | 'm1' |
| **T25** | | | | | | | 'm0' | | | | | 0.194 | 0.051 | 'm9' |
| **A26** | 0.972 | 0.004 | | | | | 'm1' | | | | | 0.079 | 0.046 | 'm9' |
| **E27** | 0.988 | 0.001 | | | | | 'm1' | | | | | 0.057 | 0.022 | 'm9' |
| **K28** | | | | | | | 'm0' | | | | | 0.050 | 0.012 | 'm9' |
| **V29** | 0.973 | 0.002 | | | | | 'm1' | 0.981 | 0.003 | | | | | 'm1' |
| **F30** | | | | | | | 'm0' | | | | | 0.122 | 0.017 | 'm9' |
| **K31** | | | | | 0.046 | 0.010 | 'm9' | | | | | 0.173 | 0.013 | 'm9' |
| **Q32** | | | | | | | 'm0' | | | | | 0.107 | 0.016 | 'm9' |
| **Y33** | | | | | | | 'm0' | | | | | | | 'm0' |
| **A34** | | | | | 0.040 | 0.025 | 'm9' | | | | | 0.117 | 0.021 | 'm9' |
| **N35** | 0.992 | 0.003 | | | | | 'm1' | | | | | | | 'm0' |
| **D36** | | | | | | | 'm0' | | | | | | | 'm0' |
| **N37** | 0.920 | 0.003 | | | | | 'm1' | 0.935 | 0.005 | | | | | 'm1' |
| **G38** | | | | | 0.132 | 0.030 | 'm9' | 0.819 | 0.018 | | | 1.563 | 0.093 | 'm3' |
| **V39** | 0.958 | 0.002 | | | | | 'm1' | 0.982 | 0.003 | | | | | 'm1' |
| **D40** | 0.796 | 0.015 | 0.053 | 0.011 | 0.409 | 0.095 | 'm4' | 0.679 | 0.012 | 0.024 | 0.007 | 2.008 | 0.065 | 'm4' |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **G41** | 0.657 | 0.008 | 0.037 | 0.005 | 0.258 | 0.064 | 'm4' | 0.646 | 0.007 | 0.050 | 0.006 | 0.815 | 0.042 | 'm4' |
| **E42** | 0.946 | 0.004 | | | | | 'm1' | | | | | | | 'm0' |
| **W43** | | | | | | | 'm0' | 0.927 | 0.020 | | | 1.159 | 0.108 | 'm3' |
| **T44** | 0.953 | 0.005 | | | | | 'm1' | | | | | | | 'm0' |
| **Y45** | 0.943 | 0.006 | | | | | 'm1' | | | | | | | 'm0' |
| **D46** | | | | | 0.120 | 0.044 | 'm9' | | | | | 0.122 | 0.077 | 'm9' |
| **D47** | 0.945 | 0.016 | | | | | 'm1' | | | | | | | 'm0' |
| **A48** | 0.882 | 0.007 | | | | | 'm1' | 0.924 | 0.013 | | | | | 'm1' |
| **T49** | 0.930 | 0.010 | | | | | 'm1' | 0.974 | 0.008 | | | | | 'm1' |
| **K50** | 0.966 | 0.006 | | | | | 'm1' | | | | | 0.176 | 0.067 | 'm9' |
| **T51** | 0.959 | 0.007 | | | | | 'm1' | | | | | | | 'm0' |
| **F52** | | | | | | | 'm0' | | | | | 0.207 | 0.039 | 'm9' |
| **T53** | 0.948 | 0.005 | | | | | 'm1' | | | | | | | 'm0' |
| **V54** | 0.982 | 0.004 | | | | | 'm1' | | | | | 0.086 | 0.034 | 'm9' |
| **T55** | 0.955 | 0.013 | | | | | 'm1' | | | | | | | 'm0' |
| **E56** | | | | | | | 'm0' | | | | | | | 'm0' |
| **W43sc** | 0.854 | 0.004 | | | | | 'm1' | 0.962 | 0.008 | | | | | 'm1' |

| | **0.75 M [Ch][Glu]** | | | | | | | **1.0 M [Ch][Glu]** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | **model** | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | **model** |
| **Q2** | | | | | | | 'm0' | 0.620 | 0.039 | | | | | 'm5' |
| **Y3** | 0.978 | 0.007 | | | | | 'm1' | 0.708 | 0.024 | | | | | 'm5' |
| **K4** | | | | | 0.442 | 0.164 | 'm9' | 0.885 | 0.009 | | | | | 'm1' |
| **L5** | | | | | | | 'm0' | 0.875 | 0.005 | | | | | 'm1' |
| **I6** | | | | | | | 'm0' | 0.876 | 0.006 | | | | | 'm1' |
| **L7** | 0.976 | 0.004 | | | | | 'm1' | 0.841 | 0.005 | | | | | 'm1' |
| **N8** | 0.937 | 0.007 | | | | | 'm1' | 0.826 | 0.004 | 0.019 | 0.012 | | | 'm2' |
| **G9** | 0.958 | 0.008 | | | | | 'm1' | 0.593 | 0.027 | | | | | 'm5' |
| **K10** | 0.914 | 0.022 | 0.136 | 0.220 | 0.238 | 0.136 | 'm4' | 0.686 | 0.011 | | | | | 'm5' |
| **T11** | 0.285 | 0.169 | | | 1.766 | 0.406 | 'm7' | 0.507 | 0.074 | | | 1.204 | 0.528 | 'm7' |
| **L12** | 0.754 | 0.006 | 0.040 | 0.007 | | | 'm2' | 0.558 | 0.007 | | | | | 'm5' |
| **K13** | 0.872 | 0.007 | 0.061 | 0.019 | | | 'm2' | 0.679 | 0.011 | | | | | 'm5' |
| **G14** | 0.820 | 0.012 | 0.048 | 0.011 | 0.319 | 0.110 | 'm4' | 0.693 | 0.007 | | | | | 'm5' |
| **E15** | 0.925 | 0.008 | 0.089 | 0.059 | | | 'm2' | 0.774 | 0.008 | | | | | 'm5' |
| **T16** | 0.974 | 0.014 | | | 0.324 | 0.118 | 'm3' | 0.826 | 0.004 | 0.029 | 0.013 | | | 'm2' |
| **T17** | 0.963 | 0.003 | | | | | 'm1' | 0.715 | 0.018 | | | | | 'm5' |
| **T18** | | | | | 0.186 | 0.050 | 'm9' | 0.799 | 0.016 | | | | | 'm5' |
| **E19** | 0.957 | 0.009 | | | | | 'm1' | 0.682 | 0.018 | | | | | 'm5' |
| **A20** | 0.935 | 0.021 | | | 0.590 | 0.179 | 'm3' | 0.489 | 0.117 | | | 1.291 | 0.674 | 'm7' |
| **V21** | | | | | 0.138 | 0.070 | 'm9' | 0.656 | 0.030 | | | | | 'm5' |
| **D22** | | | | | | | 'm0' | 0.756 | 0.013 | | | | | 'm5' |
| **A23** | | | | | 0.428 | 0.098 | 'm9' | 0.831 | 0.017 | | | | | 'm5' |
| **A24** | 0.973 | 0.016 | | | | | 'm1' | 0.814 | 0.012 | | | | | 'm5' |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **T25** | | | | | 0.428 | 0.105 | 'm9' | | | | | 4.777 | 1.170 | 'm7' |
| **A26** | | | | | | | 'm0' | 0.784 | 0.015 | | | | | 'm5' |
| **E27** | | | | | | | 'm0' | 0.859 | 0.008 | 0.056 | 0.006 | | | 'm2' |
| **K28** | | | | | 0.120 | 0.039 | 'm9' | 0.879 | 0.004 | | | | | 'm1' |
| **V29** | | | | | | | 'm0' | 0.815 | 0.011 | | | | | 'm5' |
| **F30** | | | | | 0.276 | 0.092 | 'm9' | 0.893 | 0.007 | | | 0.217 | 0.107 | 'm3' |
| **K31** | | | | | 0.352 | 0.069 | 'm9' | 0.901 | 0.005 | | | | | 'm1' |
| **Q32** | 0.978 | 0.013 | | | 0.513 | 0.128 | 'm3' | 0.879 | 0.005 | | | 0.298 | 0.082 | 'm3' |
| **Y33** | | | | | | | 'm0' | 0.889 | 0.006 | | | | | 'm1' |
| **A34** | | | | | 0.223 | 0.073 | 'm9' | 0.899 | 0.005 | | | | | 'm1' |
| **N35** | | | | | | | 'm0' | 0.871 | 0.005 | | | | | 'm1' |
| **D36** | | | | | 0.187 | 0.063 | 'm9' | 0.882 | 0.004 | | | | | 'm1' |
| **N37** | 0.961 | 0.005 | | | | | 'm1' | 0.815 | 0.004 | | | | | 'm1' |
| **G38** | 0.924 | 0.017 | | | 1.362 | 0.149 | 'm3' | 0.500 | 0.084 | | | 2.132 | 0.493 | 'm7' |
| **V39** | 0.928 | 0.015 | 0.048 | 0.029 | | | 'm2' | 0.741 | 0.012 | | | | | 'm5' |
| **D40** | 0.234 | 0.108 | | | 2.808 | 0.324 | 'm7' | 0.218 | 0.066 | | | 4.180 | 0.402 | 'm7' |
| **G41** | 0.457 | 0.070 | | | 1.377 | 0.293 | 'm8' | 0.362 | 0.026 | | | 1.524 | 0.247 | 'm7' |
| **E42** | 0.935 | 0.013 | | | 0.737 | 0.156 | 'm3' | 0.836 | 0.008 | 0.020 | 0.012 | | | 'm2' |
| **W43** | 0.961 | 0.017 | | | 0.695 | 0.170 | 'm3' | 0.798 | 0.008 | 0.024 | 0.012 | 0.854 | 0.097 | 'm4' |
| **T44** | | | | | 0.124 | 0.053 | 'm9' | 0.839 | 0.007 | | | 0.793 | 0.113 | 'm3' |
| **Y45** | | | | | 0.190 | 0.059 | 'm9' | 0.839 | 0.021 | | | | | 'm5' |
| **D46** | 0.951 | 0.022 | | | 0.934 | 0.178 | 'm3' | 0.861 | 0.005 | | | | | 'm1' |
| **D47** | | | | | | | 'm0' | 0.776 | 0.018 | | | | | 'm5' |
| **A48** | 0.919 | 0.005 | | | | | 'm1' | 0.751 | 0.011 | | | | | 'm5' |
| **T49** | 0.955 | 0.010 | | | | | 'm1' | 0.801 | 0.003 | | | | | 'm1' |
| **K50** | | | | | 0.230 | 0.056 | 'm9' | 0.885 | 0.008 | | | | | 'm1' |
| **T51** | 0.992 | 0.005 | | | | | 'm1' | 0.833 | 0.006 | | | | | 'm1' |
| **F52** | | | | | 0.504 | 0.152 | 'm9' | 0.869 | 0.014 | | | 0.289 | 0.178 | 'm3' |
| **T53** | | | | | 0.205 | 0.025 | 'm9' | 0.825 | 0.013 | | | 0.871 | 0.156 | 'm3' |
| **V54** | | | | | 0.285 | 0.058 | 'm9' | 0.859 | 0.007 | | | 0.680 | 0.126 | 'm3' |
| **T55** | 0.979 | 0.007 | | | | | 'm1' | 0.832 | 0.006 | | | | | 'm1' |
| **E56** | | | | | | | 'm0' | 0.829 | 0.008 | 0.027 | 0.013 | 0.293 | 0.110 | 'm4' |
| **W43sc** | 0.834 | 0.014 | | | 1.101 | 0.184 | 'm3' | 0.750 | 0.007 | | | 1.090 | 0.129 | 'm3' |

**1.5 M [Ch][Glu]**

| | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | model |
|---|---|---|---|---|---|---|---|
| **Q2** | 0.911 | 0.013 | 1.389 | 0.440 | | | 'm2' |
| **Y3** | | | | | | | 'm0' |
| **K4** | | | | | 1.299 | 0.675 | 'm9' |
| **L5** | | | | | | | 'm0' |
| **I6** | | | | | | | 'm0' |
| **L7** | 0.975 | 0.013 | | | | | 'm1' |
| **N8** | 0.913 | 0.016 | | | 0.820 | 0.343 | 'm3' |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **G9** | 0.909 | 0.019 | 0.112 | 0.065 | 0.698 | 0.310 | 'm4' |
| **K10** | 0.944 | 0.012 | 0.262 | 0.202 | | | 'm2' |
| **T11** | 0.903 | 0.011 | 0.280 | 0.102 | | | 'm2' |
| **L12** | 0.709 | 0.006 | | | | | 'm5' |
| **K13** | 0.867 | 0.009 | 0.080 | 0.023 | | | 'm2' |
| **G14** | 0.778 | 0.014 | 0.051 | 0.009 | 1.318 | 0.191 | 'm4' |
| **E15** | 0.890 | 0.014 | 0.095 | 0.027 | 0.440 | 0.206 | 'm4' |
| **T16** | 0.920 | 0.017 | | | 1.288 | 0.313 | 'm3' |
| **T17** | 0.925 | 0.007 | | | | | 'm1' |
| **T18** | 0.948 | 0.016 | | | 0.731 | 0.333 | 'm3' |
| **E19** | 0.925 | 0.011 | 0.054 | 0.047 | | | 'm2' |
| **A20** | 0.925 | 0.025 | 0.072 | 0.175 | 0.810 | 0.394 | 'm4' |
| **V21** | 0.934 | 0.031 | 1.365 | 1.394 | 0.979 | 0.418 | 'm4' |
| **D22** | 0.973 | 0.009 | 0.220 | 0.472 | | | 'm2' |
| **A23** | | | | | 0.632 | 0.302 | 'm9' |
| **A24** | 0.966 | 0.007 | | | | | 'm1' |
| **T25** | | | | | 1.201 | 0.376 | 'm9' |
| **A26** | 0.918 | 0.034 | 5.267 | 2.797 | 0.814 | 0.381 | 'm4' |
| **E27** | 0.931 | 0.022 | 1.277 | 0.490 | | | 'm2' |
| **K28** | | | | | | | 'm0' |
| **V29** | | | | | | | 'm0' |
| **F30** | 0.939 | 0.028 | 1.644 | 1.073 | 1.593 | 0.365 | 'm4' |
| **K31** | | | | | | | 'm0' |
| **Q32** | | | | | | | 'm0' |
| **Y33** | | | | | | | 'm0' |
| **A34** | 0.943 | 0.023 | 1.975 | 2.292 | 1.156 | 0.376 | 'm4' |
| **N35** | | | | | | | 'm0' |
| **D36** | | | | | 0.335 | 0.133 | 'm9' |
| **N37** | 0.939 | 0.010 | 0.066 | 0.061 | | | 'm2' |
| **G38** | 0.885 | 0.015 | | | 3.239 | 0.299 | 'm3' |
| **V39** | | | | | | | 'm0' |
| **D40** | 0.781 | 0.016 | 0.050 | 0.009 | 1.649 | 0.292 | 'm4' |
| **G41** | 0.510 | 0.044 | | | 2.269 | 0.603 | 'm7' |
| **E42** | 0.868 | 0.021 | 0.046 | 0.020 | 2.455 | 0.357 | 'm4' |
| **W43** | 0.960 | 0.021 | | | 0.891 | 0.373 | 'm3' |
| **T44** | | | | | | | 'm0' |
| **Y45** | | | | | 0.399 | 0.134 | 'm9' |
| **D46** | | | | | 0.459 | 0.212 | 'm9' |
| **D47** | 0.933 | 0.021 | 2.180 | 1.909 | | | 'm2' |
| **A48** | 0.897 | 0.008 | 0.104 | 0.028 | | | 'm2' |
| **T49** | 0.902 | 0.012 | | | | | 'm1' |
| **K50** | | | | | | | 'm0' |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **T51** | 0.972 | 0.013 | | | | | 'm1' |
| **F52** | | | | | 0.945 | 0.481 | 'm9' |
| **T53** | | | | | | | 'm0' |
| **V54** | | | | | 1.287 | 0.362 | 'm9' |
| **T55** | 0.980 | 0.010 | | | | | 'm1' |
| **E56** | 0.979 | 0.012 | | | | | 'm1' |
| **W43sc** | 0.764 | 0.018 | 0.031 | 0.011 | 2.543 | 0.549 | 'm4' |

Model-free parameters obtained for the relaxation dataset of water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentrations, at temperature of 298.2 K and both $^{1}H$ Larmor precession frequency of 600.13 MHz and 500.34 MHz. Analysis performed with the automated model-free analysis within relax software. For clarity, the blank spaces are maintained to indicate those residues that cannot be measured for some parameters.

**Table A.10. Model-free data for GB1 in water and in the presence of [Ch][Glu], at 310.2 K (500 and 600 MHz).**

| | Water | | | | | | | 0.1 M [Ch][Glu] | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | model | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | model |
| **Q2** | 0.898 | 0.014 | | | | | 'm1' | | | | | | | 'm0' |
| **Y3** | 0.799 | 0.004 | 0.010 | 0.005 | | | 'm2' | 0.855 | 0.003 | | | | | 'm1' |
| **K4** | 0.923 | 0.012 | 0.075 | 0.029 | | | 'm2' | | | | | | | 'm0' |
| **L5** | 0.871 | 0.007 | | | | | 'm1' | 0.927 | 0.006 | | | | | 'm1' |
| **I6** | 0.854 | 0.005 | 0.032 | 0.007 | | | 'm2' | 0.919 | 0.005 | | | | | 'm1' |
| **L7** | 0.830 | 0.003 | | | | | 'm1' | 0.889 | 0.002 | | | | | 'm1' |
| **N8** | 0.828 | 0.005 | | | | | 'm1' | 0.896 | 0.007 | 0.038 | 0.014 | | | 'm2' |
| **G9** | 0.804 | 0.005 | 0.016 | 0.004 | | | 'm2' | 0.869 | 0.005 | | | | | 'm1' |
| **K10** | 0.848 | 0.009 | 0.062 | 0.010 | | | 'm2' | 0.884 | 0.011 | 0.067 | 0.019 | | | 'm2' |
| **T11** | 0.912 | 0.011 | 0.045 | 0.022 | | | 'm2' | 0.883 | 0.016 | 0.042 | 0.019 | | | 'm2' |
| **L12** | 0.635 | 0.008 | | | | | 'm1' | 0.694 | 0.009 | | | | | 'm1' |
| **K13** | 0.758 | 0.004 | | | | | 'm1' | 0.810 | 0.006 | | | | | 'm1' |
| **G14** | 0.696 | 0.003 | 0.018 | 0.003 | | | 'm2' | 0.764 | 0.004 | | | | | 'm1' |
| **E15** | 0.794 | 0.018 | | | | | 'm1' | 0.850 | 0.017 | | | | | 'm1' |
| **T16** | 0.816 | 0.004 | 0.024 | 0.005 | | | 'm2' | 0.884 | 0.004 | | | | | 'm1' |
| **T17** | 0.826 | 0.007 | | | | | 'm1' | 0.884 | 0.011 | | | | | 'm1' |
| **T18** | 0.834 | 0.005 | | | | | 'm1' | 0.900 | 0.004 | | | | | 'm1' |
| **E19** | 0.920 | 0.015 | | | | | 'm1' | 0.907 | 0.021 | | | | | 'm1' |
| **A20** | 0.763 | 0.004 | | | | | 'm1' | 0.851 | 0.007 | | | | | 'm1' |
| **V21** | 0.945 | 0.014 | | | | | 'm1' | | | | | | | 'm0' |
| **D22** | 0.817 | 0.005 | | | | | 'm1' | 0.898 | 0.007 | | | | | 'm1' |
| **A23** | 0.909 | 0.016 | | | | | 'm1' | | | | | | | 'm0' |
| **A24** | 0.877 | 0.007 | | | | | 'm1' | 0.913 | 0.003 | | | | | 'm1' |
| **T25** | 0.849 | 0.003 | | | | | 'm1' | 0.945 | 0.007 | | | | | 'm1' |
| **A26** | 0.866 | 0.005 | 0.017 | 0.008 | | | 'm2' | 0.944 | 0.006 | | | | | 'm1' |
| **E27** | 0.915 | 0.002 | | | | | 'm1' | 0.972 | 0.004 | | | | | 'm1' |
| **K28** | 0.918 | 0.004 | | | | | 'm1' | 0.981 | 0.003 | | | | | 'm1' |
| **V29** | 0.874 | 0.004 | | | | | 'm1' | 0.947 | 0.003 | | | | | 'm1' |
| **F30** | 0.930 | 0.006 | | | | | 'm1' | | | | | | | 'm0' |
| **K31** | 0.957 | 0.006 | | | | | 'm1' | | | | | | | 'm0' |
| **Q32** | 0.908 | 0.003 | | | | | 'm1' | 0.976 | 0.003 | | | | | 'm1' |
| **Y33** | 0.921 | 0.003 | | | | | 'm1' | | | | | | | 'm0' |
| **A34** | 0.924 | 0.004 | 0.027 | 0.015 | | | 'm2' | | | | | | | 'm0' |
| **N35** | 0.921 | 0.006 | | | | | 'm1' | 0.977 | 0.006 | | | | | 'm1' |
| **D36** | 0.892 | 0.004 | | | | | 'm1' | 0.969 | 0.005 | | | | | 'm1' |
| **N37** | 0.846 | 0.005 | | | | | 'm1' | 0.902 | 0.007 | | | | | 'm1' |
| **G38** | 0.835 | 0.005 | 0.016 | 0.005 | | | 'm2' | 0.933 | 0.009 | | | | | 'm1' |
| **V39** | 0.874 | 0.003 | | | | | 'm1' | 0.919 | 0.005 | | | | | 'm1' |
| **D40** | 0.746 | 0.002 | 0.041 | 0.003 | | | 'm2' | 0.809 | 0.006 | 0.048 | 0.007 | | | 'm2' |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **G41** | 0.682 | 0.008 | 0.076 | 0.004 | | | 'm2' | 0.564 | 0.051 | | | | | 'm5' |
| **E42** | 0.872 | 0.005 | | | | | 'm1' | 0.943 | 0.009 | 0.071 | 0.056 | | | 'm2' |
| **W43** | 0.871 | 0.009 | | | | | 'm1' | | | | | | | 'm0' |
| **T44** | 0.855 | 0.006 | | | | | 'm1' | 0.901 | 0.003 | | | | | 'm1' |
| **Y45** | 0.902 | 0.015 | | | | | 'm1' | 0.914 | 0.018 | | | | | 'm1' |
| **D46** | 0.881 | 0.012 | | | | | 'm1' | 0.982 | 0.011 | | | | | 'm1' |
| **D47** | 0.902 | 0.010 | | | | | 'm1' | 0.944 | 0.015 | | | | | 'm1' |
| **A48** | 0.855 | 0.010 | | | | | 'm1' | 0.859 | 0.011 | | | | | 'm1' |
| **T49** | 0.825 | 0.008 | | | | | 'm1' | 0.907 | 0.010 | | | | | 'm1' |
| **K50** | 0.862 | 0.007 | 0.015 | 0.008 | | | 'm2' | 0.928 | 0.008 | | | | | 'm1' |
| **T51** | 0.841 | 0.007 | | | | | 'm1' | 0.883 | 0.004 | | | | | 'm1' |
| **F52** | 0.872 | 0.012 | | | 0.286 | 0.084 | 'm3' | | | | | 0.363 | 0.086 | 'm9' |
| **T53** | 0.835 | 0.005 | 0.011 | 0.005 | | | 'm2' | 0.910 | 0.006 | | | | | 'm1' |
| **V54** | 0.879 | 0.007 | 0.065 | 0.011 | | | 'm2' | 0.932 | 0.005 | | | | | 'm1' |
| **T55** | 0.693 | 0.010 | | | 0.801 | 0.059 | 'm3' | 0.924 | 0.005 | | | | | 'm1' |
| **E56** | 0.867 | 0.009 | 0.091 | 0.013 | | | 'm2' | 0.921 | 0.005 | | | | | 'm1' |
| **W43sc** | 0.872 | 0.011 | 0.103 | 0.019 | | | 'm2' | 0.906 | 0.018 | 0.074 | 0.032 | 0.577 | 0.157 | 'm4' |
| | **0.75 M [Ch][Glu]** | | | | | | | **1.0 M [Ch][Glu]** | | | | | | |
| | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | **model** | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | **model** |
| **Q2** | | | | | | | | | | | | | | |
| **Y3** | 0.851 | 0.008 | | | 0.497 | 0.063 | 'm3' | | | | | | | 'm0' |
| **K4** | 0.890 | 0.011 | | | 0.671 | 0.167 | 'm3' | | | | | 0.580 | 0.203 | 'm9' |
| **L5** | 0.885 | 0.010 | | | 0.603 | 0.103 | 'm3' | | | | | 0.157 | 0.055 | 'm9' |
| **I6** | 0.876 | 0.010 | | | 0.546 | 0.077 | 'm3' | | | | | 0.148 | 0.066 | 'm9' |
| **L7** | 0.845 | 0.005 | 0.020 | 0.009 | 0.660 | 0.060 | 'm4' | | | | | | | 'm0' |
| **N8** | 0.831 | 0.007 | | | 0.793 | 0.057 | 'm3' | 0.954 | 0.021 | | | 0.359 | 0.140 | 'm3' |
| **G9** | 0.848 | 0.007 | | | 0.590 | 0.062 | 'm3' | | | | | | | 'm0' |
| **K10** | 0.879 | 0.014 | | | 0.400 | 0.113 | 'm3' | 0.943 | 0.026 | | | 0.498 | 0.170 | 'm3' |
| **T11** | 0.870 | 0.028 | 0.038 | 0.022 | 0.385 | 0.232 | 'm4' | 0.913 | 0.069 | 0.203 | 0.640 | 0.626 | 0.262 | 'm4' |
| **L12** | 0.719 | 0.008 | 0.032 | 0.005 | | | 'm2' | 0.755 | 0.008 | 0.044 | 0.012 | | | 'm2' |
| **K13** | 0.794 | 0.011 | 0.031 | 0.007 | 0.244 | 0.099 | 'm4' | 0.892 | 0.009 | 0.075 | 0.035 | | | 'm2' |
| **G14** | 0.742 | 0.006 | 0.014 | 0.004 | 0.367 | 0.056 | 'm4' | 0.836 | 0.007 | 0.038 | 0.018 | | | 'm2' |
| **E15** | 0.859 | 0.027 | | | 0.404 | 0.193 | 'm3' | | | | | | | 'm0' |
| **T16** | 0.842 | 0.005 | | | 0.672 | 0.044 | 'm3' | 0.988 | 0.005 | | | | | 'm1' |
| **T17** | 0.923 | 0.007 | | | | | 'm1' | | | | | | | 'm0' |
| **T18** | 0.851 | 0.006 | | | 0.727 | 0.065 | 'm3' | | | | | 0.128 | 0.066 | 'm9' |
| **E19** | 0.939 | 0.015 | | | | | 'm1' | | | | | | | 'm0' |
| **A20** | 0.824 | 0.010 | | | 0.538 | 0.092 | 'm3' | 0.945 | 0.026 | | | 0.403 | 0.215 | 'm3' |
| **V21** | 0.868 | 0.022 | | | 1.114 | 0.199 | 'm3' | | | | | | | 'm0' |
| **D22** | 0.846 | 0.014 | | | 0.822 | 0.104 | 'm3' | | | | | 0.153 | 0.070 | 'm9' |
| **A23** | 0.957 | 0.023 | | | 0.836 | 0.161 | 'm3' | | | | | 0.636 | 0.163 | 'm9' |
| **A24** | 0.823 | 0.007 | | | 1.268 | 0.088 | 'm3' | 0.954 | 0.023 | | | 0.816 | 0.207 | 'm3' |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **T25** | 0.867 | 0.005 | | | 1.142 | 0.088 | 'm3' | | | | | 0.551 | 0.075 | 'm9' |
| **A26** | 0.899 | 0.008 | | | 0.808 | 0.073 | 'm3' | | | | | 0.433 | 0.049 | 'm9' |
| **E27** | 0.890 | 0.017 | | | 0.816 | 0.147 | 'm3' | | | | | 0.793 | 0.035 | 'm9' |
| **K28** | 0.857 | 0.005 | | | 1.497 | 0.040 | 'm3' | | | | | 1.023 | 0.042 | 'm9' |
| **V29** | 0.865 | 0.007 | | | 1.125 | 0.062 | 'm3' | | | | | 0.609 | 0.073 | 'm9' |
| **F30** | 0.889 | 0.010 | | | 1.203 | 0.093 | 'm3' | | | | | 0.978 | 0.065 | 'm9' |
| **K31** | 0.893 | 0.010 | | | 1.286 | 0.099 | 'm3' | | | | | 1.084 | 0.095 | 'm9' |
| **Q32** | 0.846 | 0.004 | | | 1.486 | 0.036 | 'm3' | 0.954 | 0.017 | | | 1.222 | 0.128 | 'm3' |
| **Y33** | 0.869 | 0.007 | | | 1.017 | 0.090 | 'm3' | | | | | 0.800 | 0.054 | 'm9' |
| **A34** | 0.910 | 0.013 | | | 1.070 | 0.111 | 'm3' | | | | | 0.972 | 0.080 | 'm9' |
| **N35** | 0.842 | 0.004 | | | 1.347 | 0.055 | 'm3' | 0.954 | 0.021 | | | 1.124 | 0.178 | 'm3' |
| **D36** | 0.870 | 0.005 | | | 1.212 | 0.042 | 'm3' | | | | | 0.807 | 0.048 | 'm9' |
| **N37** | 0.809 | 0.007 | | | 0.913 | 0.062 | 'm3' | | | | | | | 'm0' |
| **G38** | 0.836 | 0.005 | | | 1.129 | 0.069 | 'm3' | | | | | 0.445 | 0.082 | 'm9' |
| **V39** | 0.828 | 0.004 | | | 1.156 | 0.044 | 'm3' | 0.943 | 0.019 | | | 0.780 | 0.131 | 'm3' |
| **D40** | 0.723 | 0.005 | 0.028 | 0.004 | 1.074 | 0.061 | 'm4' | 0.732 | 0.013 | 0.037 | 0.009 | 1.600 | 0.104 | 'm4' |
| **G41** | 0.586 | 0.008 | 0.054 | 0.003 | 0.675 | 0.079 | 'm4' | 0.573 | 0.010 | 0.049 | 0.005 | 1.328 | 0.086 | 'm4' |
| **E42** | 0.826 | 0.005 | | | 1.115 | 0.070 | 'm3' | 0.932 | 0.018 | | | 1.019 | 0.147 | 'm3' |
| **W43** | 0.849 | 0.011 | | | 1.038 | 0.115 | 'm3' | | | | | 0.236 | 0.110 | 'm9' |
| **T44** | 0.855 | 0.006 | | | 0.762 | 0.066 | 'm3' | | | | | | | 'm0' |
| **Y45** | | | | | | | 'm0' | | | | | | | 'm0' |
| **D46** | 0.849 | 0.008 | | | 1.016 | 0.100 | 'm3' | 0.937 | 0.035 | | | 1.287 | 0.247 | 'm3' |
| **D47** | 0.897 | 0.032 | | | 0.564 | 0.321 | 'm3' | | | | | | | 'm0' |
| **A48** | 0.853 | 0.027 | | | 0.680 | 0.191 | 'm3' | 0.969 | 0.017 | | | | | 'm1' |
| **T49** | 0.835 | 0.013 | | | 0.684 | 0.104 | 'm3' | 0.962 | 0.009 | | | | | 'm1' |
| **K50** | 0.871 | 0.011 | | | 0.684 | 0.091 | 'm3' | | | | | 0.218 | 0.082 | 'm9' |
| **T51** | 0.850 | 0.010 | | | 0.679 | 0.080 | 'm3' | | | | | | | 'm0' |
| **F52** | 0.864 | 0.006 | | | 1.106 | 0.067 | 'm3' | | | | | 0.928 | 0.141 | 'm9' |
| **T53** | 0.863 | 0.011 | | | 0.802 | 0.069 | 'm3' | | | | | | | 'm0' |
| **V54** | 0.891 | 0.006 | | | 0.583 | 0.072 | 'm3' | | | | | 0.194 | 0.060 | 'm9' |
| **T55** | 0.856 | 0.005 | | | 0.797 | 0.080 | 'm3' | | | | | 0.255 | 0.054 | 'm9' |
| **E56** | 0.852 | 0.007 | | | 0.801 | 0.112 | 'm3' | | | | | | | 'm0' |
| **W43sc** | 0.810 | 0.015 | | | 1.069 | 0.142 | 'm3' | 0.880 | 0.026 | | | 0.954 | 0.220 | 'm3' |

**1.5 M [Ch][Glu]**

| | $S^2$ | SE | $\tau_e$ (ns) | SE | $R_{ex}$ ($s^{-1}$) | SE | model |
|---|---|---|---|---|---|---|---|
| **Q2** | | | | | | | |
| **Y3** | 0.761 | 0.013 | | | 2.674 | 0.178 | 'm3' |
| **K4** | 0.783 | 0.012 | | | 3.490 | 0.297 | 'm3' |
| **L5** | 0.893 | 0.009 | | | | | 'm1' |
| **I6** | 0.790 | 0.013 | | | 2.651 | 0.179 | 'm3' |
| **L7** | 0.451 | 0.103 | | | 2.070 | 0.682 | 'm7' |
| **N8** | 0.778 | 0.008 | | | 1.911 | 0.115 | 'm3' |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **G9** | 0.793 | 0.010 | | | 1.944 | 0.163 | 'm3' |
| **K10** | 0.559 | 0.093 | | | 1.760 | 0.602 | 'm7' |
| **T11** | | | | | 5.353 | 0.688 | 'm7' |
| **L12** | 0.705 | 0.016 | 0.039 | 0.010 | 0.879 | 0.133 | 'm4' |
| **K13** | 0.761 | 0.010 | | | 1.767 | 0.102 | 'm3' |
| **G14** | 0.686 | 0.009 | 0.016 | 0.008 | 2.134 | 0.095 | 'm4' |
| **E15** | 0.804 | 0.018 | | | 2.195 | 0.209 | 'm3' |
| **T16** | 0.782 | 0.012 | | | 2.704 | 0.149 | 'm3' |
| **T17** | 0.837 | 0.024 | | | 1.648 | 0.236 | 'm3' |
| **T18** | 0.772 | 0.008 | | | 2.806 | 0.143 | 'm3' |
| **E19** | 0.790 | 0.021 | | | 1.842 | 0.223 | 'm3' |
| **A20** | 0.761 | 0.012 | | | 2.317 | 0.168 | 'm3' |
| **V21** | 0.872 | 0.017 | | | 1.522 | 0.241 | 'm3' |
| **D22** | 0.803 | 0.027 | | | 2.879 | 0.247 | 'm3' |
| **A23** | 0.879 | 0.025 | | | 2.857 | 0.271 | 'm3' |
| **A24** | 0.758 | 0.012 | | | 3.523 | 0.151 | 'm3' |
| **T25** | 0.872 | 0.006 | | | | | 'm1' |
| **A26** | 0.779 | 0.008 | | | 3.798 | 0.154 | 'm3' |
| **E27** | 0.735 | 0.005 | | | 4.761 | 0.124 | 'm3' |
| **K28** | 0.750 | 0.006 | | | 4.178 | 0.134 | 'm3' |
| **V29** | 0.774 | 0.006 | | | 3.604 | 0.111 | 'm3' |
| **F30** | 0.752 | 0.006 | | | 4.727 | 0.110 | 'm3' |
| **K31** | 0.746 | 0.010 | | | 4.775 | 0.143 | 'm3' |
| **Q32** | 0.822 | 0.006 | | | 2.277 | 0.130 | 'm3' |
| **Y33** | 0.743 | 0.009 | | | 4.875 | 0.136 | 'm3' |
| **A34** | 0.762 | 0.007 | | | 4.901 | 0.124 | 'm3' |
| **N35** | 0.737 | 0.007 | | | 4.833 | 0.097 | 'm3' |
| **D36** | 0.768 | 0.006 | | | 4.007 | 0.092 | 'm3' |
| **N37** | 0.758 | 0.009 | | | 3.367 | 0.135 | 'm3' |
| **G38** | 0.736 | 0.010 | | | 3.899 | 0.133 | 'm3' |
| **V39** | 0.723 | 0.008 | | | 3.785 | 0.115 | 'm3' |
| **D40** | 0.645 | 0.006 | | | 3.549 | 0.104 | 'm3' |
| **G41** | 0.506 | 0.033 | | | 2.677 | 0.221 | 'm7' |
| **E42** | 0.755 | 0.006 | | | 3.140 | 0.121 | 'm3' |
| **W43** | | | | | 5.705 | 0.527 | 'm7' |
| **T44** | 0.835 | 0.015 | | | 1.460 | 0.177 | 'm3' |
| **Y45** | 0.926 | 0.023 | | | 1.608 | 0.263 | 'm3' |
| **D46** | 0.780 | 0.021 | | | 3.156 | 0.229 | 'm3' |
| **D47** | 0.867 | 0.029 | | | 0.785 | 0.431 | 'm3' |
| **A48** | 0.289 | 0.126 | | | 2.608 | 0.936 | 'm7' |
| **T49** | 0.300 | 0.028 | | | | | 'm5' |
| **K50** | 0.818 | 0.020 | | | 2.903 | 0.217 | 'm3' |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **T51** | 0.810 | 0.019 | | | 2.301 | 0.193 | 'm3' |
| **F52** | 0.358 | 0.241 | | | 4.457 | 1.556 | 'm7' |
| **T53** | 0.820 | 0.018 | | | 2.102 | 0.225 | 'm3' |
| **V54** | 0.886 | 0.011 | | | 0.924 | 0.140 | 'm3' |
| **T55** | 0.785 | 0.010 | | | 2.830 | 0.129 | 'm3' |
| **E56** | 0.836 | 0.012 | 0.033 | 0.018 | 0.703 | 0.160 | 'm4' |
| **W43sc** | 0.387 | 0.078 | | | 1.723 | 0.951 | 'm7' |

Model-free parameters obtained for the relaxation dataset of water, 0.1 M, 0.75 M, 1.0 M, and 1.5 M of [Ch][Glu] concentrations, at temperature of 310.2 K and both $^{1}H$ Larmor precession frequency of 600.13 MHz and 500.34 MHz. Analysis performed with the automated model-free analysis within relax software. For clarity, the blank spaces are maintained to indicate those residues that cannot be measured for some parameters.

# Appendix B

**Table B.1. List of the affected residues for the interaction between [Ch][Glu], [Bmim][dca] and respective ionic salts.**

| aa | [Ch][Glu] | | | | aa | [Bmim][dca] | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0.35 M [Ch][Glu] | 0.35 M [NaCl] | 0.35 M [Ch][Cl] | 0.35 M [NaGlu] | | 0.15 M [Bmim][dca] | 0.15 M [NaCl] | 0.15 M [Bmim][Cl] | 0.15 M [Na][dca] |
| M1 | | | | | M1 | | | | |
| E2 | | | | | E2 | x | x | x | x |
| A3 | | | | | A3 | x | x | | x |
| I4 | | | x | | I4 | | | | |
| A5 | | | | | A5 | | | | |
| K6 | | x | x | | K6 | x | | | x |
| H7 | x | | | | H7 | x | x | x | |
| D8 | | x | x | | D8 | | x | x | |
| F9 | | | x | | F9 | x | | | |
| S10 | | | | | S10 | x | | | |
| A11 | | | | | A11 | | | | |
| T12 | | | | | T12 | x | x | x | x |
| A13 | | | | | A13 | x | | x | |
| D14 | | | | | D14 | | | | |
| D15 | | | | | D15 | x | | x | |
| E16 | | x | | | E16 | | | | |
| L17 | | | | | L17 | | x | x | x |
| S18 | | | | x | S18 | | | | |
| F19 | | | | | F19 | | | | |
| R20 | x | | | | R20 | | | | |
| K21 | | | | | K21 | | | | |
| T22 | | | | | T22 | x | | | |
| Q23 | | | x | | Q23 | x | | x | x |
| I24 | | | x | | I24 | | | x | |
| L25 | x | | | x | L25 | | | | |
| K26 | | | | | K26 | x | x | x | x |
| I27 | x | x | x | x | I27 | x | x | x | x |
| L28 | x | x | | x | L28 | x | x | x | x |
| N29 | | | | x | N29 | | | | x |
| M30 | | x | x | | M30 | | x | x | |
| E31 | | | | | E31 | x | | x | |
| D32 | | | | x | D32 | | | | |
| D33 | | | | | D33 | | | | |
| S34 | | | | | S34 | | | x | x |
| N35 | | | | | N35 | | | x | |
| W36 | x | | | | W36 | | | | x |
| Y37 | x | | x | | Y37 | x | | | |
| R38 | x | | x | | R38 | | | | |
| A39 | x | | | | A39 | | | | |
| E40 | | | | | E40 | x | x | x | |
| L41 | | | | | L41 | | | | |
| D42 | | | | | D42 | | | | |
| G43 | | | | | G43 | | | | |
| K44 | x | | | | K44 | | | | |
| E45 | | | x | | E45 | | x | | |
| G46 | | | | | G46 | | | | |
| L47 | | | | | L47 | | | | |
| I48 | x | | | | I48 | | | | |
| P49 | | | | | P49 | | | | |
| S50 | | | | | S50 | | | | |
| N51 | | | | | N51 | x | | x | |
| Y52 | | | x | | Y52 | | | | |
| I53 | x | | | x | I53 | | | | |
| E54 | x | | | | E54 | | | | |
| M55 | x | | | x | M55 | x | x | | x |
| K56 | x | | x | x | K56 | | | | |
| N57 | | | | | N57 | | | | |
| H58 | | | | | H58 | | | | |
| D59 | x | x | x | | D59 | x | x | x | x |

β1 β2 β3 β4 β5

Positive Negative Hydrophobic Polar, uncharged Special cases

The affected residues (above the determined cut-off line) are marked with an "x" and the primary sequence is coloured according to the amino acid chemistry (positive - blue; negative - red; hydrophobic – yellow; polar, uncharged – purple; special cases - grey). At the sides of the table it is depicted the secondary structure of the protein according to the 3D structures (folded and unfolded).

**Table B.2. Backbone assignment of drkN SH3 for the fully stabilised folded and unfolded states in ILs.**

| | Folded state, 1.0 M [Ch][Glu] IL | | | | | Unfolded state, 1.0 M [Bmim][dca] IL | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | δCα (± 0.03 ppm) | δCβ (± 0.02 ppm) | δCO (± 0.01 ppm) | $\delta H_N$ (± 0.002 ppm) | $\delta N_H$ (± 0.02 ppm) | δCα (± 0.02 ppm) | δCβ (± 0.02 ppm) | δCO (± 0.01 ppm) | $\delta H_N$ (± 0.002 ppm) | $\delta N_H$ (± 0.03 ppm) |
| **M1** | 55.56 | 34.21 | 173.32 | - | - | - | - | - | - | - |
| **E2** | 55.31 | 34.07 | 174.13 | 8.861 | 124.68 | 56.46 | 30.68 | 175.85 | - | - |
| **A3** | 50.38 | 25.07 | 174.40 | 8.973 | 122.09 | 52.53 | 19.36 | 177.40 | 8.437 | 124.91 |
| **I4** | 59.20 | 41.02 | 176.00 | 8.876 | 119.05 | 61.11 | 38.98 | 175.75 | 8.011 | 119.38 |
| **A5** | 52.92 | 21.38 | 178.87 | 8.795 | 128.18 | 52.31 | 19.46 | 177.16 | 8.273 | 127.53 |
| **K6** | 56.18 | 33.11 | 174.31 | 9.355 | 126.18 | 56.34 | 33.28 | 176.14 | 8.142 | 120.34 |
| **H7** | 53.80 | 34.97 | 172.95 | 7.511 | 114.17 | 56.12 | 31.21 | 174.66 | - | - |
| **D8** | 54.94 | 41.98 | 175.76 | 8.654 | 119.46 | 53.94 | 41.33 | 176.03 | 8.137 | 121.20 |
| **F9** | 57.31 | 42.76 | 173.63 | 8.521 | 121.14 | 58.18 | 39.24 | 175.89 | 8.325 | 121.15 |
| **S10** | 55.73 | 63.49 | 171.66 | 7.634 | 123.62 | 58.93 | 63.92 | 174.26 | 8.276 | 116.74 |
| **A11** | 52.85 | 20.96 | 178.42 | 8.267 | 127.52 | 52.72 | 19.45 | 177.75 | 8.211 | 125.59 |
| **T12** | 61.98 | 70.15 | 173.50 | 9.058 | 113.50 | 61.69 | 70.11 | 174.49 | 8.091 | 112.19 |
| **A13** | 50.69 | 22.34 | 178.13 | 7.556 | 123.97 | 52.86 | 19.40 | 177.62 | 8.306 | 125.71 |
| **D14** | 56.76 | 40.87 | 176.09 | 8.646 | 119.17 | 54.98 | 41.37 | 176.48 | 8.279 | 119.32 |
| **D15** | 53.37 | 40.20 | 176.20 | 8.363 | 114.61 | 54.98 | 41.28 | 176.94 | 8.247 | 120.38 |
| **E16** | 55.27 | 32.95 | 175.16 | 7.393 | 119.25 | 58.04 | 30.02 | 177.40 | 8.404 | 120.90 |
| **L17** | 53.59 | 45.49 | 174.38 | 8.466 | 124.65 | 56.36 | 41.82 | 178.15 | 8.199 | 120.95 |
| **S18** | 59.00 | 64.62 | 173.84 | 7.923 | 117.79 | 60.06 | 63.40 | 175.56 | 8.139 | 115.14 |
| **F19** | 56.86 | 40.65 | 173.88 | 8.660 | 114.63 | 59.49 | 39.15 | 176.67 | 8.159 | 121.76 |
| **R20** | 53.12 | 33.14 | 175.82 | 9.387 | 119.24 | 57.76 | 30.46 | 177.28 | 8.175 | 120.52 |
| **K21** | 58.21 | 33.25 | 177.31 | 9.233 | 120.54 | 58.26 | 32.71 | 177.53 | 8.173 | 119.94 |
| **T22** | 66.60 | 69.34 | 173.51 | 8.813 | 117.14 | 64.06 | 69.40 | 175.32 | 7.953 | 113.28 |
| **Q23** | 58.92 | 30.34 | 174.23 | 8.173 | 120.55 | 57.13 | 29.11 | 176.86 | 8.056 | 121.22 |
| **I24** | 59.59 | 37.83 | 175.73 | 8.209 | 120.46 | 62.81 | 38.27 | 176.70 | 7.963 | 120.19 |
| **L25** | 53.96 | 44.74 | 175.55 | 9.054 | 127.77 | 56.50 | 42.01 | 177.88 | 8.044 | 122.23 |
| **K26** | 55.53 | 33.17 | 175.97 | 8.272 | 122.01 | 57.73 | 32.83 | 177.44 | 7.922 | 119.91 |
| **I27** | 59.16 | 35.51 | 175.58 | 8.526 | 124.74 | 62.52 | 38.46 | 177.04 | 7.854 | 119.87 |
| **L28** | 55.54 | 42.74 | 177.28 | 9.198 | 128.45 | 55.85 | 42.33 | 177.20 | 8.133 | 122.47 |
| **N29** | 54.15 | 41.27 | 173.73 | 7.613 | 113.85 | 53.52 | 39.03 | 175.23 | 8.176 | 117.63 |
| **M30** | 55.15 | 33.23 | 174.65 | 8.915 | 123.75 | 56.08 | 32.99 | 176.24 | 8.199 | 119.94 |
| **E31** | 56.62 | 31.33 | 176.73 | 8.125 | 120.97 | 56.86 | 30.44 | 176.19 | 8.364 | 120.42 |
| **D32** | 56.58 | 41.08 | 176.38 | 8.377 | 120.95 | 54.47 | 41.62 | 176.13 | 8.246 | 120.77 |
| **D33** | 54.10 | 42.52 | 175.44 | 8.212 | 118.79 | 54.61 | 41.11 | 176.93 | 8.358 | 121.92 |
| **S34** | 59.45 | 64.09 | 174.94 | 8.260 | 116.09 | 59.83 | 63.62 | 174.97 | 8.393 | 115.84 |
| **N35** | 54.09 | 39.81 | 174.70 | 8.981 | 119.19 | 53.92 | 38.85 | 175.43 | 8.347 | 119.66 |
| **W36** | 56.15 | 31.56 | 174.85 | 8.005 | 121.63 | 58.01 | 29.50 | 175.99 | 7.944 | 120.47 |
| **Y37** | 52.89 | 39.70 | 175.43 | 8.933 | 120.66 | 58.38 | 38.47 | 175.48 | 7.746 | 119.31 |
| **R38** | 56.47 | 31.49 | 175.20 | 8.729 | 121.45 | 56.03 | 30.86 | 175.79 | 7.823 | 121.72 |
| **A39** | 50.84 | 26.11 | 173.34 | 9.561 | 131.54 | 53.10 | 19.26 | 177.91 | 8.145 | 124.65 |
| **E40** | 54.52 | 34.19 | 174.79 | 8.945 | 118.39 | 56.98 | 30.07 | 176.72 | 8.424 | 118.86 |
| **L41** | 54.78 | 45.94 | 175.88 | 8.918 | 125.91 | 55.31 | 42.31 | 177.20 | 8.086 | 121.56 |
| **D42** | 55.51 | 39.99 | 175.72 | 9.643 | 129.53 | 54.66 | 41.42 | 176.95 | 8.256 | 120.29 |
| **G43** | 45.55 | - | 173.80 | 9.150 | 104.61 | 45.78 | - | 174.68 | 8.381 | 109.15 |
| **K44** | 54.92 | 34.65 | 174.08 | 7.928 | 122.30 | 56.55 | 32.85 | 176.80 | 8.183 | 120.21 |
| **E45** | 53.88 | 33.90 | 176.89 | 8.474 | 120.49 | 57.35 | 30.03 | 176.97 | 8.573 | 120.85 |
| **G46** | 45.56 | - | 170.72 | 8.818 | 107.88 | 45.31 | - | 173.75 | 8.354 | 108.86 |
| **L47** | 54.62 | 44.65 | 177.58 | 9.255 | 120.46 | 55.18 | 42.59 | 176.82 | 7.968 | 120.91 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **I48** | 56.68 | 41.06 | 172.36 | 9.386 | 112.62 | 58.59 | 38.67 | 174.58 | 8.056 | 122.13 |
| **P49** | 60.75 | 30.03 | 178.79 | - | - | 63.52 | 32.20 | 176.98 | - | - |
| **S50** | 60.42 | 62.48 | 175.03 | 7.865 | 119.75 | 59.02 | 63.73 | 174.45 | 8.347 | 115.31 |
| **N51** | 53.67 | 36.41 | 175.68 | 8.403 | 115.23 | 53.35 | 38.81 | 174.82 | 8.319 | 119.42 |
| **Y52** | 58.46 | 39.04 | 174.89 | 7.704 | 118.91 | 58.60 | 38.78 | 175.49 | 8.033 | 119.85 |
| **I53** | 58.60 | 42.00 | 173.79 | 7.213 | 110.60 | 61.31 | 38.95 | 175.68 | 7.805 | 121.68 |
| **E54** | 54.45 | 33.15 | 175.70 | 8.736 | 120.22 | 56.67 | 30.45 | 176.33 | 8.259 | 123.81 |
| **M55** | 55.25 | 31.67 | 176.55 | 9.050 | 125.15 | 55.64 | 33.05 | 175.95 | 8.265 | 121.74 |
| **K56** | 56.26 | 33.20 | 175.91 | 8.660 | 123.63 | 56.23 | 33.17 | 175.97 | 8.223 | 121.73 |
| **N57** | 53.25 | 39.33 | 174.48 | 8.516 | 120.33 | 53.27 | 39.09 | 174.55 | 8.374 | 119.44 |
| **H58** | 55.64 | 30.51 | 173.76 | 8.276 | 118.99 | 55.72 | 30.68 | 173.82 | 8.102 | 118.99 |
| **D59** | 56.04 | 42.13 | 180.89 | 8.191 | 126.75 | 56.10 | 42.29 | 180.94 | 8.155 | 126.69 |
| **W36sc** | - | - | - | 10.226 | 129.06 | - | - | - | 10.204 | 129.49 |

The backbone and aliphatic $H^N$, $N^H$, C$\alpha$, C$\beta$ and CO chemical shifts are shown for drkN SH3 fully stabilised folded and unfolded states in 1.0 M [Ch][Glu] ($F_{[Ch][Glu]}$) or 1.0 M [Bmim][dca] ($U_{[Bmim][dca]}$), respectively. Spectra acquired at 298.2 K, 600.13 MHz. For the assignment, 2D $^{1}H$-$^{15}N$ HSQC and 3D triple resonance experiments (HNCACB, HNcoCACB, HNCO, HNcaCO spectra) were used. The errors are the average of the standard deviation per residue.

**Table B.3. Parameters for the drkN SH3 interconversion extracted from ZZex in water and aqueous-ILs.**

| | Water | | | | | | 0.35 M [Ch][Glu] IL | | | | | | 0.15 M [Bmim][dca] IL | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Residue** | $p_f$ | $p_u$ | $R_{1f}$ ($s^{-1}$) | $R_{1u}$ ($s^{-1}$) | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) | $p_f$ | $p_u$ | $R_{1f}$ ($s^{-1}$) | $R_{1u}$ ($s^{-1}$) | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) | $p_f$ | $p_u$ | $R_{1f}$ ($s^{-1}$) | $R_{1u}$ ($s^{-1}$) | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) |
| **A3** | 0.55 | 0.45 | 2.50 | 2.16 | 0.44 | 0.54 | 0.85 | 0.15 | 2.02 | 2.10 | 0.13 | - | 0.20 | 0.80 | 4.04 | 2.38 | 0.27 | 0.07 |
| **I4** | 0.51 | 0.49 | 2.50 | 2.00 | 0.86 | - | 0.82 | 0.18 | 2.08 | 2.85 | 0.14 | 0.64 | 0.10 | 0.90 | 1.30 | 1.84 | 1.55 | 0.18 |
| **A5** | 0.51 | 0.49 | 2.69 | 2.27 | 0.59 | 0.61 | 0.82 | 0.18 | 2.29 | 2.68 | - | - | - | - | - | - | - | - |
| **D8** | 0.59 | 0.41 | 3.36 | 2.34 | - | - | 0.88 | 0.12 | 1.93 | 0.50 | - | 0.64 | - | - | - | - | - | - |
| **S10** | 0.61 | 0.39 | 2.84 | 2.37 | 0.62 | 0.97 | 0.79 | 0.21 | 1.93 | 2.45 | 0.15 | 0.56 | 0.27 | 0.73 | 3.40 | 2.51 | 0.67 | 0.25 |
| **T12** | 0.45 | 0.55 | 2.36 | 2.46 | 0.78 | 0.63 | 0.88 | 0.12 | 1.88 | 2.00 | 0.15 | 1.12 | 0.16 | 0.84 | 2.69 | 2.39 | 1.20 | 0.22 |
| **A13** | 0.58 | 0.42 | 2.35 | 2.33 | 0.62 | 0.86 | 0.79 | 0.21 | 1.96 | 2.44 | 0.14 | 0.54 | 0.26 | 0.74 | 2.62 | 2.30 | 0.91 | 0.31 |
| **D15** | - | - | - | - | - | - | 0.59 | 0.41 | 1.96 | 1.87 | 0.24 | - | 0.15 | 0.85 | 2.23 | 2.39 | 1.29 | |
| **E16** | 0.46 | 0.54 | 2.50 | 2.52 | 0.47 | 0.40 | 0.87 | 0.13 | 2.05 | 2.46 | 0.16 | 1.04 | 0.16 | 0.84 | 2.85 | 2.33 | 0.54 | 0.11 |
| **S18** | 0.57 | 0.43 | 2.47 | 2.55 | 0.63 | 0.85 | 0.72 | 0.28 | 2.39 | 2.46 | 0.40 | 1.04 | 0.24 | 0.76 | 2.65 | 2.47 | 1.05 | 0.33 |
| **F19** | 0.72 | 0.28 | 3.11 | 2.08 | 0.51 | 1.31 | 0.68 | 0.32 | 2.62 | 2.38 | 0.36 | 0.79 | 0.27 | 0.73 | 4.20 | 3.16 | 0.93 | 0.34 |
| **R20** | 0.63 | 0.37 | 2.73 | 1.83 | 0.57 | 0.97 | 0.49 | 0.51 | 2.50 | 2.52 | - | 0.29 | 0.10 | 0.90 | 3.31 | 2.51 | 0.62 | 0.07 |
| **T22** | 0.59 | 0.41 | 2.71 | 2.50 | 0.67 | 0.98 | 0.80 | 0.20 | 2.09 | 2.63 | 0.18 | 0.72 | 0.23 | 0.77 | 2.57 | 2.55 | 1.26 | 0.37 |
| **L25** | 0.63 | 0.37 | 2.65 | 1.43 | 0.53 | 0.91 | 0.92 | 0.08 | 2.03 | 1.95 | 0.10 | 1.24 | 0.17 | 0.83 | 2.02 | 2.22 | 1.16 | 0.24 |
| **I27** | 0.53 | 0.47 | 2.39 | 2.14 | 0.77 | 0.87 | 0.64 | 0.36 | 1.82 | 1.88 | 0.20 | - | 0.14 | 0.86 | 1.70 | 2.00 | 1.04 | - |
| **L28** | 0.49 | 0.51 | 2.40 | 2.37 | 0.77 | 0.74 | 0.84 | 0.16 | 2.00 | 2.48 | 0.09 | 0.49 | 0.15 | 0.85 | 1.90 | 2.22 | 1.30 | 0.22 |
| **N29** | 0.66 | 0.34 | 2.60 | 2.07 | 0.54 | 1.07 | 0.82 | 0.18 | 2.06 | 1.51 | 0.13 | 0.58 | 0.30 | 0.70 | 2.60 | 2.33 | 0.98 | 0.42 |
| **M30** | 0.47 | 0.53 | 2.76 | 2.68 | 1.05 | 0.94 | - | - | - | - | - | - | - | - | - | - | - | - |
| **R38** | 0.55 | 0.45 | 2.50 | 2.51 | 0.49 | 0.60 | 0.84 | 0.16 | 2.00 | 2.48 | 0.09 | 0.49 | 0.18 | 0.82 | 2.57 | 2.37 | - | 0.17 |
| **A39** | 0.40 | 0.60 | 2.34 | 2.68 | 0.87 | 0.59 | 0.84 | 0.16 | 2.06 | 2.55 | 0.14 | 0.77 | 0.14 | 0.86 | 2.04 | 2.28 | 1.35 | 0.21 |
| **E40** | 0.46 | 0.54 | 2.27 | 2.41 | 0.87 | | 0.81 | 0.19 | 1.92 | 2.18 | 0.18 | 0.80 | 0.15 | 0.85 | 2.59 | 2.32 | 0.72 | 0.13 |
| **L41** | 0.60 | 0.40 | 2.58 | 3.10 | 0.57 | 0.84 | 0.85 | 0.15 | 1.94 | 1.94 | 0.18 | 1.06 | 0.18 | 0.82 | 2.33 | 2.22 | 0.84 | 0.18 |
| **D42** | 0.48 | 0.52 | 2.68 | 1.98 | 0.68 | 0.64 | 0.86 | 0.14 | 2.00 | 2.13 | 0.12 | 0.73 | 0.12 | 0.88 | 2.17 | 2.48 | 1.55 | 0.20 |
| **G43** | 0.55 | 0.45 | 2.46 | 2.37 | 0.69 | 0.84 | 0.74 | 0.26 | 2.13 | 2.61 | 0.01 | 0.02 | 0.20 | 0.80 | 2.44 | 2.32 | 1.12 | 0.27 |
| **G46** | 0.56 | 0.44 | 2.56 | 2.28 | 0.63 | 0.80 | 0.86 | 0.14 | 2.00 | 2.13 | 0.12 | 0.73 | 0.21 | 0.79 | 3.07 | 2.45 | 0.90 | 0.24 |
| **L47** | 0.37 | 0.63 | 2.42 | 2.57 | 0.54 | 0.31 | 0.74 | 0.26 | 2.13 | 2.61 | 0.01 | 0.02 | 0.12 | 0.88 | 2.07 | 2.21 | - | 0.16 |
| **I48** | 0.40 | 0.60 | 2.20 | 2.21 | 0.91 | 0.62 | 0.74 | 0.26 | 2.01 | 2.27 | 0.19 | 0.52 | 0.10 | 0.90 | 1.53 | 2.01 | - | 0.15 |
| **S50** | 0.61 | 0.39 | 2.79 | 2.40 | - | 1.03 | 0.88 | 0.12 | 2.09 | 1.98 | - | 1.18 | 0.26 | 0.74 | 2.69 | 2.50 | - | 0.41 |
| **Y52** | 0.50 | 0.50 | 2.69 | 2.23 | 0.57 | 0.56 | 0.80 | 0.20 | 2.16 | 1.65 | 0.06 | 0.25 | 0.16 | 0.84 | 2.34 | 2.12 | 1.08 | - |
| **I53** | 0.50 | 0.50 | 2.55 | 2.30 | 0.75 | 0.75 | 0.82 | 0.18 | 2.12 | 2.31 | 0.17 | 0.75 | 0.14 | 0.86 | 1.98 | 2.22 | 1.18 | 0.20 |
| **E54** | 0.42 | 0.58 | 2.40 | 2.52 | - | 0.44 | 0.76 | 0.24 | 2.05 | 2.77 | - | 0.36 | 0.12 | 0.88 | 2.61 | 2.29 | - | 0.10 |
| **M55** | 0.51 | 0.49 | 2.81 | 2.57 | 0.49 | 0.50 | 0.85 | 0.15 | 2.19 | 2.01 | 0.10 | 0.54 | 0.17 | 0.83 | 3.88 | 2.63 | 0.52 | - |
| **K56** | 0.62 | 0.38 | 2.69 | 2.30 | 0.62 | 1.02 | 0.86 | 0.14 | 2.22 | 2.22 | 0.20 | 1.26 | 0.25 | 0.75 | 3.70 | 2.44 | 0.48 | 0.16 |
| **W36sc** | 0.50 | 0.50 | 2.24 | 2.08 | 0.49 | 0.49 | 0.85 | 0.15 | 2.19 | 2.01 | 0.10 | 0.54 | 0.16 | 0.84 | 2.48 | 1.93 | 0.77 | 0.15 |

Parameters extracted by fitting the experimental data on both auto-peaks (*ff* and *uu*) and exchange cross peaks (*uf* and f*u*) simultaneously in MATLAB. Data extracted from a series of 2D ZZex spectra acquired at 293.2 K, 600.13 MHz with variable mixing time ranging from 10 to 750 ms.

**Table B.4. Rate constants of drkN SH3 interconversion.**

| | Water | | 0.35 M [Ch][Glu] | | 0.15 M [Bmim[dca] | |
|---|---|---|---|---|---|---|
| **Residue** | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) | $k_{fu}$ ($s^{-1}$) | $k_{uf}$ ($s^{-1}$) |
| **S10** | 0.62 | 0.97 | 0.15 | 0.56 | 0.67 | 0.25 |
| | | | 0.21 [a] | 0.77 [a] | 0.68 [a] | 0.26 [a] |
| **T12** | 0.78 | 0.63 | 0.15 | 1.12 | 1.20 | 0.22 |
| | | | 0.21 [a] | 1.56 [a] | 1.22 [a] | 0.23 [a] |
| **A13** | 0.62 | 0.86 | 0.14 | 0.54 | 0.91 | 0.31 |
| | | | 0.20 [a] | 0.75 [a] | 0.92 [a] | 0.32 [a] |
| **E16** | 0.47 | 0.40 | 0.16 | 1.04 | 0.54 | 0.11 |
| | | | 0.22 [a] | 1.45 [a] | 0.55 [a] | 0.11 [a] |
| **S18** | 0.63 | 0.85 | 0.40 | 1.04 | 1.05 | 0.33 |
| | | | 0.55 [a] | 1.45 [a] | 1.06 [a] | 0.33 [a] |
| **T22** | 0.67 | 0.98 | 0.18 | 0.72 | 1.26 | 0.37 |
| | | | 0.25 [a] | 0.99 [a] | 1.28 [a] | 0.38 [a] |
| **L25** | 0.53 | 0.91 | 0.10 | 1.24 | 1.16 | 0.24 |
| | | | 0.14 [a] | 1.72 [a] | 1.18 [a] | 0.25 [a] |
| **L28** | 0.77 | 0.74 | 0.09 | 0.49 | 1.30 | 0.22 |
| | | | 0.13 [a] | 0.68 [a] | 1.32 [a] | 0.22 [a] |
| **N29** | 0.54 | 1.07 | 0.13 | 0.58 | 0.98 | 0.42 |
| | | | 0.18 [a] | 0.80 [a] | 1.00 [a] | 0.43 [a] |
| **A39** | 0.87 | 0.59 | 0.14 | 0.77 | 1.35 | 0.21 |
| | | | 0.20 [a] | 1.07 [a] | 1.37 [a] | 0.22 [a] |
| **D42** | 0.68 | 0.64 | 0.12 | 0.73 | 1.55 | 0.20 |
| | | | 0.17 [a] | 1.01 [a] | 1.57 [a] | 0.21 [a] |
| **G46** | 0.63 | 0.80 | 0.12 | 0.73 | 0.90 | 0.24 |
| | | | 0.17 [a] | 1.01 [a] | 0.91 [a] | 0.25 [a] |
| **I53** | 0.75 | 0.75 | 0.17 | 0.75 | 1.18 | 0.20 |
| | | | 0.24 [a] | 1.04 [a] | 1.20 [a] | 0.20 [a] |
| **K56** | 0.62 | 1.02 | 0.20 | 1.26 | 0.48 | 0.16 |
| | | | 0.28 [a] | 1.75 [a] | 0.49 [a] | 0.16 [a] |
| **W36sc** | 0.49 | 0.49 | 0.10 | 0.54 | 0.77 | 0.15 |
| | | | 0.13 | 0.75 [a] | 0.78 [a] | 0.15 [a] |
| **Average** | **0.6 ± 0.1** | **0.8 ± 0.2** | **0.16 ± 0.07** | **0.8 ± 0.3** | **1.0 ± 0.3** | **0.24 ± 0.09** |
| | | | **0.2 ± 0.1[a]** | **1.1 ± 0.4[a]** | **1.0 ± 0.3[a]** | **0.25 ± 0.09[a]** |

Rate constants of accurately followed residues by ZZex experiments for water and aqueous-ILs conditions (0.35 M [Ch][Glu] and 0.15 M [Bmim][dca]) at 293.2 K, 600.13 MHz. Uncertainties represent the standard deviation from the average. [a] Viscosity-adjusted rates. The relative viscosity ($\eta_{rel}$) used in 0.35 M [ChGlu], and 0.15 M [Bmim][dca] are 1.39, and 1.02, respectively.

2021

MICAEL SIMÕES SILVA

# NMR INVESTIGATION OF ION-PAIR MODULATION OF PROTEIN STRUCTURE AND DYNAMICS AND THEIR RELATION TO PROTEIN STABILITY