UNIVERSITE MONTPELLIER II
SCIENCES ET TECHNIQUES DU LANGUEDOC

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER II

Discipline : Biophysique

Ecole Doctorale : CBS2 - Sciences chimiques et Biologiques pour la Santé

présentée et soutenue publiquement par

Jean-Baptiste ROUGET

le 13 Décembre 2010

Titre :

STUDY OF THE PHYSICAL BASIS OF PRESSURE EFFECTS ON PROTEINS USING MODEL ANKYRIN REPEAT CONSTRUCTS

JURY

M. Christian ROUMESTAND	Président du Jury
Mme Catherine ROYER	Directeur de thèse
M. Martin BLACKLEDGE	Rapporteur
M. Thomas KIEFHABER	Rapporteur
M. Bruno ROBERT	Examinateur
M. Martin WEIK	Examinateur
M. Doug BARRICK	Invité

I'd like to thank the whole "Centre de Biochimie Structurale", where I spent most of my time during the last three years. It was a great pleasure to work with everyone, and specially with Cathy because of her constant happiness and curiosity about everything. I want to thank too Doug Barrick, Rolland Winter, Kazuyuki Akasaka for our collaborations.

I don't forget of course my parents who supported me during all my studies and during all the changes this work made into my life. They are always here and positive for me, whatever the moments or the circumstances, and I'll never forget it.

Greg, Bastien, Cyril, thank you for your unconditional support and friendship.

Finally, a very special thank to Yamna for her comprehension, support and time, and I apologize sincerely for all the stress I brought to her during the redaction of this work, for all the weekends I didn't spend with her but in the lab, for all the times I wasn't able to help her more.

Contents

I Introduction 1

1 General introduction to protein folding 3
1.1 Protein Folding and Human health . . . 3
1.2 A short overview of protein folding . . . 4
1.3 Protein folding in the pressure-temperature plane . . . 6
1.3.1 Temperature effects on protein stability . . . 9
1.3.2 Pressure effects on protein stability . . . 12
At constant temperature . . . 12
With varying temperature . . . 14
1.4 ΔV^{u}_{eq} : an incomplete understanding of pressure effects on protein stability 18
1.5 Objectives . . . 20

2 A model for probing pressure effects on protein folding 21
2.1 Advantages of repeat proteins for studies on protein folding . . . 21
2.2 Ankyrin repeat proteins . . . 23
2.2.1 Background of the ankyrin repeat proteins . . . 23
2.2.2 Sequences of ankyrin motifs . . . 23
2.2.3 Structure of the ankyrin motif and protein . . . 24
2.3 Following the unfolding by fluorescence spectroscopy . . . 26
2.3.1 Existing mutants used in this work . . . 28
$Nank1-7^{*}$ as a reference for pressure-induced unfolding studies 28
N-terminal truncated constructs . . . 28
C-terminal truncated constructs . . . 29
A to G mutants based on the complete protein $Nank1-7\Delta$ . . . 30
What is known about these variants . . . 30
2.3.2 A well known folding mechanism . . . 35

II Materials and methods 39
2.4 Production and purification of Ankyrin proteins and mutants . . . 41
2.5 High-pressure Fluorescence Spectroscopy . . . 42
2.6 Small Angle X-ray Scattering (SAXS) under pressure . . . 45
2.7 Fourier Transform Infra Red Spectroscopy (FTIR) under pressure . . . 47

2.8 Densitometry 48
2.9 Differential Scanning Calorimetry (DSC) and Pressure Perturbation Calorimetry (PPC) 49
2.9.1 DSC 49
2.9.2 PPC 50
2.10 Molecular Dynamics Simulations 53

III Results **55**

3 $Nank1-7^{\star}$ **57**
3.1 Pressure-induced equilibrium unfolding at various temperature 57
3.2 Chain condensation in the pressure unfolded state 65
3.3 Secondary structure in the pressure-unfolded state 67
3.4 Pressure-jump kinetics 70
3.5 Pressure effects on $Nank1-7^{\star}$ 75
3.6 Short discussion 78

4 Study of Ankyrin deletion constructs **83**
4.1 Study of the ΔV_{eq}^{f} dependence with size 83
4.1.1 Equilibrium unfolding at 20°C 83
N-terminal deletion constructs 83
Short discussion 87
C-terminal deletion constructs 87
Sequence mutated constructs : NAG2 and NAG5 89
Short discussion 90
Small constructs of two and three repeats 91
4.1.2 Kinetic experiments at 20°C 92
Short discussion 94
4.2 Study of the ΔV_{eq}^{f} dependence with temperature 97
4.2.1 At the equilibrium 97
$Nank2-7^{\star}$ 97
$Nank3-7^{\star}$ 98
4.2.2 Sequence mutated constructs : NAG2 and NAG5 101
NAG5 101
4.2.3 Mutants with C-terminal deletion 101
$Nank1-6^{\star}$ 101
$Nank1-5^{\star}$ 102
4.2.4 Kinetic experiments 103
Mutants with N-terminal deletion 103
$Nank2-7^{\star}$ 103
$Nank3-7^{\star}$ 105

4.2.5 Molecular Dynamic Simulations . . . 107
Cavity determination in the crystallographic structure of *Nank*1–7* . . . 107
Defining dynamics of the apparent cavities . . . 108
Exploring the voids . . . 112
Total amino acid hydration analysis . . . 114
4.3 Pressure perturbation calorimetry study . . . 117
4.3.1 Ankyrin repeat constructs . . . 117
4.3.2 PPC study on BPTI . . . 118

IV Discussion and conclusions 121

5 Discussion and conclusions 123
5.1 Discussion : Volume change for Ankyrin repeat protein folding . . . 123
5.1.1 How do size and sequence affect the volume change upon unfolding ? . . . 123
How does size contribute to ΔV_{eq}^{f} ? . . . 123
How does sequence contributes to ΔV_{eq}^{f} ? . . . 124
5.1.2 How do cavities contribute to ΔV_{eq}^{f} ? . . . 126
α-helices and ΔV_{eq}^{f} . . . 126
Localization of cavities in the structure thanks to experimental results and MD simulations . . . 127
Transition State Ensemble . . . 129
Analysis of the localization of cavities and amino acid sequence 130
5.1.3 Pressure effects on the pre-exponential factor . . . 130
5.2 Conclusions . . . 131
5.3 Perspectives . . . 132
5.3.1 Pressure perturbation calorimetry . . . 132
5.3.2 High pressure NMR . . . 134

V Appendix 137

A Developing new tools 139

VI Contributions 141

List of Tables

1.1 Rate of elementary steps in protein folding 6

2.1 Sequence of the 6 identified and the seventh putative Ankyrin Repeat form Drosophila melanogaster Notch 24
2.2 Urea-induced unfolding parameters of CD and fluorescence 28
2.3 Thermodynamic parameters for folding of deletion constructs of the Notch ankyrin domain 29
2.4 Urea-induced unfolding parameters of CD and fluorescence 30
2.5 Free energies calculated with Ising model 37
2.6 Vibrational bands of diagnostic value for high-pressure FTIR studies . . 47
2.7 Typical assignments after complete H/D exchange 48

3.1 Thermodynamic parameters (ΔG_u^0 and ΔV_{eq}^f) for the pressure-induced unfolding of Nank1-7 as a function of temperature and urea concentration 60
3.2 Kinetic parameters from the p-jump fluorescence studies on Nank1-7 . 71

4.1 Thermodynamic variables of $Nank2-7^\star$, $Nank3-7^\star$ and $Nank4-7^\star$, at 20°C, depending on urea concentration. 86
4.2 Thermodynamic variables of $Nank2-7^\star$, $Nank3-7^\star$ and $Nank4-7^\star$ at 20°C. 86
4.3 Thermodynamic variables of $Nank1-6^\star$, at 20°C, depending on urea concentration. 88
4.4 Thermodynamic variables of $Nank1-5^\star$, at 20°C, depending on urea concentration. 88
4.5 Thermodynamic variables of Nank NAG2, at 20°C, depending on urea concentration. 90
4.6 Thermodynamic variables of NAG5, at 20°C, depending on urea concentration. 90
4.7 Thermodynamic variables of all the mutants at 20°C, depending on urea concentration. 95
4.8 Volumes of each mutant, at 20°C. 95
4.9 Thermodynamic variables of $Nank2-7^\star$, at 12, 20 and 28°C, depending on urea concentration. 98

4.10 Thermodynamic variables of $Nank3-7^{*}$, at 12, 20 and 28°C, depending on urea concentration. 100
4.11 Thermodynamic variables of $Nank2-7^{*}$, at 12°C, depending on urea concentration. 104
4.12 Thermodynamic variables of $Nank2-7^{*}$, at 20°C, depending on urea concentration. 104
4.13 Thermodynamic variables of $Nank2-7^{*}$, at 28°C, depending on urea concentration. 104
4.14 Thermodynamic variables of $Nank3-7^{*}$, at 12°C, depending on urea concentration. 106
4.15 Thermodynamic variables of $Nank3-7^{*}$, at 20°C, depending on urea concentration. 106
4.16 Thermodynamic variables of $Nank3-7^{*}$, at 28°C, depending on urea concentration. 106

List of Figures

1.1 Energetic point of view of protein folding. 5
1.2 Structural evolution for folding 5
1.3 p, T-stability diagram 7
1.4 Schematic view of the stability phase transition diagram 8
1.5 Temperature-dependence of the ΔG^u_{eq}, ΔH and ΔS 9
1.6 Dependance of ΔC_p upon unfolding on ΔASA. 11
1.7 Volume composition of a protein 13
1.8 ΔV^u_{eq} evolution with distinct parameters 15
1.9 Expansivity and volume of the folded and unfolded state of proteins. . . 15
1.10 Expansivity of water and amino acids acids depending on temperature. 17

2.1 Comparison between globular protein and repeat proteins 22
2.2 Sequence alignment of Notch ankyrin domains from different taxa. . . . 25
2.3 Structure of ankyrin motifs and ankyrin repeat protein 26
2.4 Urea-induced unfolding of ankyrin repeat proteins. 27
2.5 Structural transitions of $Nank4-7^*$ 29
2.6 Thermal unfolding of the ankyrin repeats of Drosophila Notch. 31
2.7 Experimental and calculated hydrodynamic properties of Notch ankyrin repeat proteins. 32
2.8 Comparison of CD- and fluorescence-monitored urea unfolding transitions of glycine variants of the Drosophila Notch ankyrin domain. . . . 33
2.9 Fluorescence-detected unfolding and refolding of the Notch ankyrin domain. 34
2.10 A nearest-neighbor thermodynamic description of repeat-protein stability. 36
2.11 Equilibrium energy landscape for the Notch ankyrin domain. 38
2.12 Evolution of tryptophan emission spectra with pressure. 43
2.13 Unfolding curves by area under the curve and mass center of spectra. . 43
2.14 Narrowing of the amide I' band of α-lactalbumin 47

3.1 Evolution of the center of the spectral mass of tryptophan emission spectra with urea and temperature 58
3.2 Evolution of tryptophan emission spectra with pressure. 59
3.3 Evolution of the ΔG^u_{eq} with urea at all temperatures. 60

3.4 Stability diagram of $Nank1-7^*$. 61
3.5 Evolution of ΔV_{eq}^f with temperature for $Nank1-7^*$ 62
3.6 Pressure Perturbation Calorimetry experiment on $Nank1-7^*$ 62
3.7 Evolution of ΔV_{eq}^f and volumes with temperature 64
3.8 Calculated p,T transition diagram of $Nank1-7^*$ 64
3.9 SAXS under pressure experiments on $Nank1-7^*$ at 20°C 66
3.10 Comparison of the pair-distance functions at 24°C 67
3.11 FTIR under pressure measurements of $Nank1-7^*$ at 20°C 69
3.12 Natural logarithm of the p-jump fluorescence detected relaxation time . 70
3.13 Evolution of the ΔV_{eq}^f, $\Delta V_f^{\#}$ and $\Delta V_u^{\#}$ for all the urea concentrations at 20°C. 72
3.14 Evolution of activation and equilibrium volume changes with temperature. 73
3.15 Schema of the evolution of volume changes with temperature for $Nank1-7^*$ 74
3.16 Evolution of $ln(K_f)$ and $ln(K_u)$ over $\Delta G/RT$, depending on temperature. 75
3.17 Evolution of $ln(k_u)$ (= - $ln(k_f)$) with pressure when $\Delta G = 0$ for $Nank1-7^*$ at 20°C. 76
3.18 3-dimensional representation of $ln(k_f)$ as a function of pressure and temperature 77
3.19 Energetic diagram of the protein folding. 78

4.1 Analysis of tryptophan intrinsic fluorescence emission spectra of $Nank4-7^*$. 84
4.2 Center of the spectral mass of $Nank2-7^*$, $Nank3-7^*$ and $Nank4-7^*$ at 20°C, depending on urea and pressure. 85
4.3 Linear regression of ΔG_{eq}^f of $Nank2-7^*$, $Nank3-7^*$ and $Nank4-7^*$ at 20°C 86
4.4 Thermodynamic parameters for folding of deletion constructs of the Notch ankyrin domain 86
4.5 Evolutions of ΔV_{eq}^f at 20°C depending on the number of repeats 87
4.6 Evolution of the mass center of tryptophan emission spectra of $Nank1-6^*$ at 20°C 88
4.7 Linear regression of ΔG_{eq}^f for $Nank1-5^*$ and $Nank1-6^*$ 89
4.8 Conserved alanines allong the $Nank1-7^*$ structure 89
4.9 Center of spectral mass for $NankNAG2$ and $NAG5$ 90
4.10 Evolution of ΔV_{eq}^f depending on the number of repeats for proteins with the first repeat 91
4.11 Sequences of all repeats used 91
4.12 Evolution of the mass center of the spectra for NW and NWC 92
4.13 Evolution of ΔV_{eq}^f at 20°C depending on the number of repeats. 92
4.14 Natural logarithm of τ of the mutants at 20°C 93

4.15 Evolution of the relaxation times 93
4.16 Volume changes obtained at 20°C for the Nank variants and the NW and NWC consensus constructs. 96
4.17 Center of the spectral mass of $Nank2-7^{\star}$ depending on temperature, urea and pressure. 97
4.18 Evolution of $Nank2-7^{\star}$'s ΔV_{eq}^{f} with temperature. 99
4.19 Evolution of the mass center and $ln(\tau)$ of $Nank3-7^{\star}$ at 12, 20 and 28°C. 99
4.20 Evolution of $Nank3-7^{\star}$'s ΔV_{eq}^{f} with temperature. 100
4.21 DSC PPC of NAG5. 101
4.22 DSC PPC of $Nank1-6^{\star}$, and associated expansivity calculation. 102
4.23 DSC PPC of $Nank1-5^{\star}$ and associated expansivity calculation. 103
4.24 Natural logarithm of τ of $Nank2-7^{\star}$ 104
4.25 Evolution of the volume changes of $Nank2-7^{\star}$ depending on temperature 105
4.26 Evolution of the $ln(\tau)$ of $Nank3-7^{\star}$ at 12, 20 and 28°C. 105
4.27 Evolution of the volume changes of $Nank3-7^{\star}$ at 12, 20 and 28°C. . . 107
4.28 Overview of cavities in the $Nank1-7^{\star}$ crystallographic structure. . . . 108
4.29 Cutting of $Nank1-7^{\star}$ structure. 109
4.30 Snapshots of the movements of the first motifs in MD simulations 110
4.31 Evolution of the RMSD of each motif during the simulation on $Nank1-7^{\star}$. 110
4.32 Overview of cavities in the 1-2-3 construct. 111
4.33 Cavity probes and cavity hydration in MD simulations for 3-5 and 5-7 constructs. 113
4.34 Cavities presented and attended respectively in $Nank1-7^{\star}$ and NAG5 115
4.35 MD simulations of the whole $Nank1-7^{\star}$ protein and calculations of voids and hydration. 115
4.36 Pressure perturbation calorimetry plots of $Nank1-7^{\star}$, $Nank1-6^{\star}$, $Nank1-5^{\star}$ and NAG5 117
4.37 PPC of BPTI proteins 118
4.38 PPC of BPTI mutants depending on pH and associated ΔV_{eq}^{u} 119

5.1 Sequence analysis of repeats 124

A.1 Screenshot of home-made Matlab program for DSC-PPC treatments. . 140

Part I

Introduction

Chapter 1

General introduction to protein folding

In 2005, Science magazine placed the understanding of the protein folding in the 125 major still unsolved problems [1]. Indeed, despite fifty years of intensive work, this problem remains incompletely understood.

1.1 Protein Folding and Human health

Proteins are complex macromolecules consisting of a linear chain of amino acids. This linear chain undergoes a folding process to adopt a 3-dimensional native structure, generally required for the activity of the protein. Protein folding is implicated in numerous diseases and biological processes [2, 3, 4], for example, the prion diseases and the associated amyloid fibrils [5, 6, 7] are clearly linked to defects in protein folding [8]. Hence, a profound examination of the mechanisms of folding/unfolding is needed to better understand these diseases, and to discover how to inhibit such processes.

Prions are proteins that increase their number autocatalytically by inducing normal proteins to aggregate into amyloïd fibrils by simple contact [9], and are known to provide transmissible spongiform encephalopathies [7, 5], such as Bovine Spongiform Encephalopathy [10] and Creutzfeldt-Jakob disease (CJD). Other non transmissible folding diseases include Alzheimer, Lou Gehrig and Gerstmann Sträussler Scheinker diseases [11]. The complete understanding of protein folding would help us understanding this kind of diseases, and may eventually provide information that will be useful in developing new therapeutic strategies.

Another point of interest for studying protein folding is the existence of natively unfolded proteins. Many human and viral proteins are unfolded and fold themselves transciently only at the moment of their activity, making the characterization of their folded structures very difficult. Structural approaches for drug design require that the proteins be quite stable in order to obtain a Nuclear Magnetic Resonance (NMR) or X-ray crystallographic structure. Currently, accurate modeling of these structures in-silico based on structural homology is hindered because these proteins are generally highly flexible in their unfolded regions which have poor homology with known protein

structures. Unfortunately, such natively unfolded proteins could in principle be targets for new drugs : one example is the Vif protein, known to be an excellent target for drug-design against HIV-1 [12, 13, 14, 15]. The complete knowledge of the protein folding process may help us find better models or in-silico structures which can be used as a basis for the stabilization of such proteins or the creation of pharmacophores. Finally, a fundamental understanding of protein folding is a real challenge that is interesting enough to merit all our attention from a physical chemistry point of view. The history of science has proved over time that we never know what a discovery can bring in terms of new opportunities, biotechnologies or new fields of research before finding it. But it is likely that the perspectives will be equal to the challenge.

1.2 A short overview of protein folding

Over the last several decades, the characterization of the thermodynamic and physical rules that control the structure and the dynamics of biomolecules, such as proteins [16], has been an intensive and productive field of research [17]. Understanding the forces that govern the stability and folding mechanisms is one of the central problems in the protein biophysics [18, 19, 20, 21, 22, 23, 24, 25, 26, 27]. Thanks to thermal and chemical denaturation, using denaturants like urea or guanidine hydrochloride (GuHCl), a great deal has been learnt about how the intramolecular interactions, such as charge-charge interactions, hydrogen bonds, van der Waals interactions or hydrophobic effects [28, 29, 30, 31] contribute to the proteins' stability [31]. However, the understanding of the mechanism of protein folding still remains incomplete, in particular the destabilizing effects of hydrostatic pressure are not understood.

To address this unanswered thermodynamic question, we will focus in this work on a simple protein model, where the folding is reversible, two-state, without any visible intermediates [32]. From the equilibrium energetic point of view, as represented in figure 1.1 A, the unfolded population has a higher energy than the folded protein, and the difference of energy between the two populations is the Gibbs free energy difference for unfolding (ΔG^u_{eq}). Hence, the larger the ΔG^u_{eq} is in absolute value, the greater the stability of the folded protein is.

Proteins are marginally stable within a range of ΔG^u_{eq} between 20 and 40 $KJ.mol^{-1}$. The reason for this poor stability is generally attributed to the fact that proteins need a high conformational flexibility to permit their function [35], which results from entropy-enthalpy compensation [36, 37, 38]. The highly dynamic character of proteins comes too from the fact that the statistical fluctuations of the thermodynamic variables can be important, highly depending on their hydration [39, 40, 41]. According to the two-state model, a barrier exists between those two energetically different populations, which is only transiently populated during the transition between the two states.

From the kinetic point of view, proteins undergo thermal motions, leading to en-

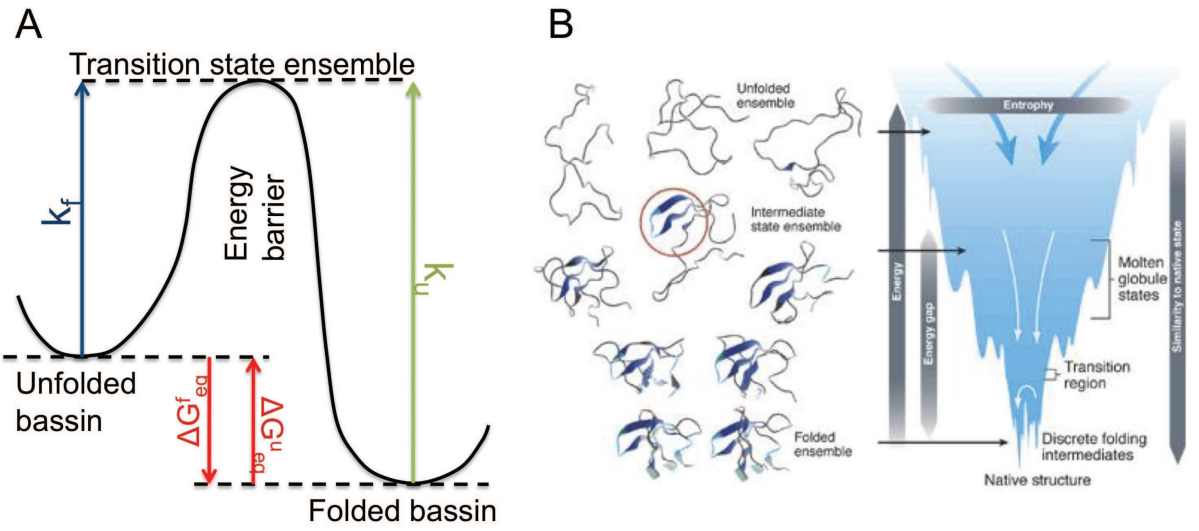

Figure 1.1: Energetic point of view of protein folding. A : simple two state protein energetic diagram. B : Folding funnel of a protein [33, 34]

.

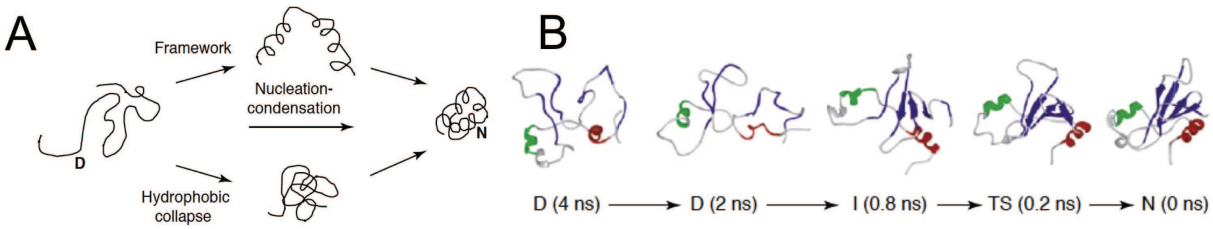

Figure 1.2: Structural evolution for folding. A : Different pathways during protein folding. B : Snapshots of molecular dynamics simulations of Barnase during its folding. Based on [27].

ergetic changes. These energetic motions imply strucural changes. These changes of the unfolded protein can be energetically small, so the protein remains in the unfolded basin while exploring the energetic landscape, as shown in the upper part of figure 1.1 B. From these thermal fluctuations, some fraction of α-helices or β-sheets can appear. Depending upon their position, the protein can fall into energetic traps, but if they are correct the folding process continues, leading to a complexification of the overall structure and to a decrease of internal energy and entropy. All these different structures are transient, as the formation of helices and sheets are thought to occur on very short time scales [42, 43], as shown in table 1.1 and mainly via a mechanism of nucleation [42, 44, 25, 45, 26] and/or hydrophobic collapse [27], as shown in figure 1.2 A. The lower the energy of the protein structure, the more the number of statistical structure possibilities decreases, until the transition state ensemble is reached. At this point the folding process goes on and the energy minimization results on the folded structure, as represented in figures 1.1 B and 1.2 B for Barnase [27]. Considering that all these rearrangements are on different time scales but occur globally together, folding can be a very fast phenomenon [46, 47]. Even if the folded structure is unique, it is highly dynamic, undergoing a wide variety of internal motions on multiple time

Structure	Model system	Method	Timescale
$\alpha - helix$	Ala based peptide (21 aa)	TJ	10-20 ns
			10 ns
	Ala based peptide (21 aa)	TJ	≈160 ns
	Gln based peptide (≈135 aa)	US	≈ 30 ns
			200 ps
$\beta - hairpin$	16 aa from protein G B1	TJ	$6\mu s$
$\beta - turn$	Gly-Ser based peptide (3-4 aa)	TT	≈ 20 ns
loops	Gly-Ser based peptide (6-10 aa)	TT	50-100 ns
	GdmCl-unfolded cytochrome C (62 aa)	PR	40 μs
			≈ 1 μs
collapse	apomyoglobin	TJ	4 μs
	cytochrome C	CF	50 μs

Table 1.1: Rate of elementary steps in protein folding, from [42]. TT = triplet-triplet energy transfer; PR = photodissociation/rebinding; TJ = Temperature jump; US = ultrasonic absorption; CF = continuous flow.

scales. Hence, proteins are complex vibrating systems which "breath" in space during time, with sometimes intermediates or conformers more or less opened or folded. So we have not to take the protein folding process as a static system, but rather as a dynamic system [48]. Nonetheless, contrary to the equilibrium point of view, and in contrast to single molecule experiments or molecular dynamic simulations, in order to observe in an ensemble the folding/unfolding process, we must perturb the system enough to significantly change the populations and then monitor the relaxation to equilibrium.

1.3 Protein folding in the pressure-temperature plane

The earth is predominantly a high-pressure environment, and 62% of the total biosphere is characterized by pressures greater than 100 bar [49]. Recent advances in analytical techniques for studying proteins under pressure have facilitated the investigation of molecular interactions within proteins and protein complexes, and have provided new insights into the structure and function of biomolecules at high pressures [50]. Thanks to the seminal work of Hawley [51], Brandts [52] and Kauzmann [53], we have since the early 1970's a good idea of the behavior of two-state protein folding in the temperature-pressure plane. Assuming that protein folding is of sufficient cooperativity to be modeled as a two-state transition, the native state of the protein is populated only in a defined region of temperature and pressure. This is represented by pointing the ellipse of $\Delta G^u_{eq} = 0$ depending on temperature and pressure, as shown in figure 1.3 A; indeed, the protein is natively folded when ΔG^u_{eq} is positive, and unfolded when ΔG^u_{eq} is negative. This p,T stability of proteins can be generally approximated by an ellipse for all proteins, but each protein's ellipse has it's own shape (as shown in figure 1.3 B), depending on thermodynamic variables. Calculation of the thermodynamic pa-

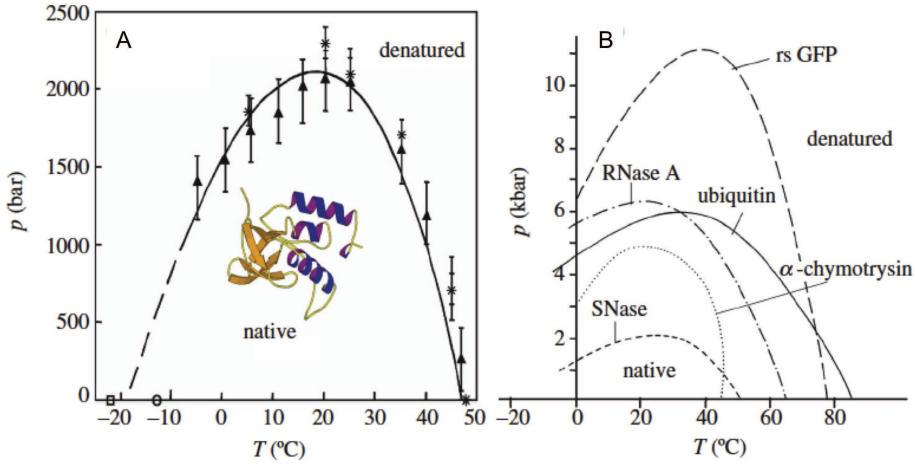

Figure 1.3: p,T-stability diagram of (A) SNase at pH 5.5 as obatined by SAXS (star), FTIR (triangle) and DSC measurements (circle and square) and (B) of several other monomeric proteins. Figure taken from [54].

rameters responsible for this pressure-temperature stability diagram is carried out via the determination of the ΔG^u_{eq} of the protein at each temperature and pressure, and fit to the following equation [54]:

$$\Delta G(p,T) = \Delta G_0 - \Delta S_0(T - T_0) - \Delta C_p[T(ln\frac{T}{T_0} - 1) + T_0]$$

$$+\Delta V_0(p - p_0) + (\Delta\frac{1}{2}\beta)(p - p_0)^2 + \Delta\alpha(T - T_0)(p - p_0) \quad (1.1)$$

which is a second order Taylor series expansion of the equation $\delta\Delta G = \Delta V\delta p - \Delta S\delta T$ around a chosen reference point (T_0, p_0) to (T,p) [55, 56]. In equation 1.1, the first order terms (entropy difference upon unfolding between the folded and unfolded state (ΔS)), the second order term (heat capacity difference upon unfolding between the two states (ΔC_p)) account for the temperature dependance. The pressure stability depends on the first order and the second order terms, the volume difference upon unfolding between the unfolded and the folded state (ΔV), the difference in isothermal compressibility upon unfolding between the two states ($\Delta\beta$), which is generally assumed to be equal to 0, and, finally, the difference of thermal expansivity upon unfolding between the two states of the protein ($\Delta\alpha$)).

Careful inspection of this diagram, as represented in figure 1.4 , reveals interesting aspects. First, because of experimental difficulties in the exploration of negative pressures, only a part of this diagram (at positive pressure) can be obtained experimentally. Despite these difficulties, work from Uden et al [59] shed to light the possible shape of the ellipse at negative pressures, as shown as a gray zone in figure 1.4. Second, we can see that ΔG^u_{eq} changes both with temperature at a constant pressure, and

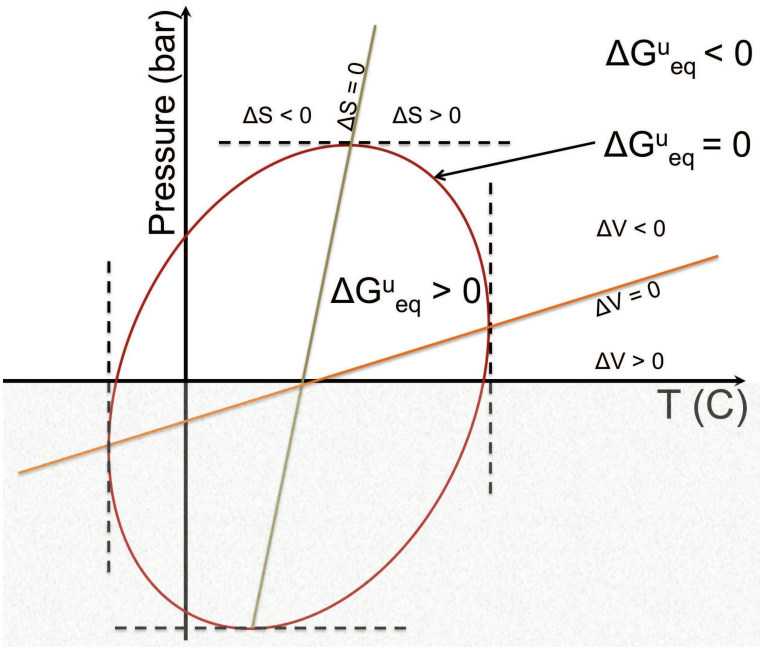

Figure 1.4: Schematic view of the stability phase transition diagram of a protein on the theoretical temperature and pressure ensemble (positives and negatives). In red, the evolution of the $\Delta G^u_{eq} = 0$, in orange the line of the null ΔV^u_{eq} and in green the line of the null ΔS. This schema is based on [51, 57, 58].

with pressure at a constant temperature. Thanks to equation 1.1, we know that this change in ΔG^u_{eq} with pressure at constant temperature comes from the ΔV^u_{eq}, as ΔV^u_{eq} is the only parameter defined in respect to pressure only which does not be equal to 0. Third, the tangent at the largest and lowest pressures where the protein is still folded correspond to the point where ΔS equals 0, so we can draw the null ΔS line. On the left of this line, ΔS is negative and favors folding, and on the right ΔS is positive and favors unfolding. Fourth, we can observe that at the largest and lowest temperatures where the protein is still folded, the tangent correspond to a place where $\delta\Delta G/\delta\Delta P$ tends to 0. As $\delta\Delta G/\delta\Delta P = \Delta V$, at the exact position of the tangent, ΔV equals 0. So we can represent the line of the null ΔV in the whole temperature and pressure range between those two points. In the upper area, ΔV^u_{eq} is negative, and under this line ΔV^u_{eq} is positive. On this figure, we can see that the $\Delta V = 0$ line is not exactly on the X-axis, so there may exist a positive ΔV^f_{eq} zone at high temperature and low pressures. We forced here this representation for a better understanding, but experimentally, this line seems really close to the x axis, and it's not obviously evident to observe this positive ΔV zone, as described by Hawley [51]. The amplitude of the effect depends most importantly on the amplitude of the difference in thermal expansivity between the folded and unfolded states.

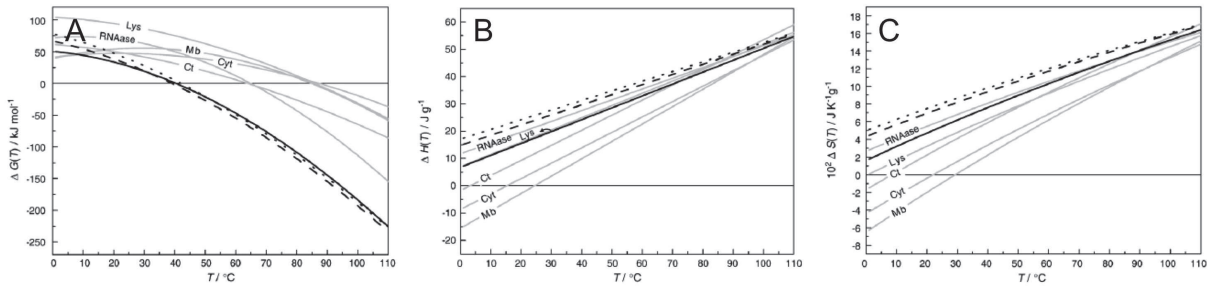

Figure 1.5: Temperature-dependence of A : ΔG^u_{eq}, B : ΔH and C : ΔS, for several proteins : RNase, ribonuclease; Lys, lysozyme; Ct, α-chymotrypsin; Cyt, cytochrome c, and Mb, myoglobin [60]

1.3.1 Temperature effects on protein stability

Protein stability is defined by the Gibbs free energy difference upon unfolding between the folded and the unfolded states (ΔG^u_{eq}), and can be determined experimentally by several methods. Temperature affects protein stability, which is reflected by positive or negative changes in ΔG^u_{eq} depending upon whether a particular change in temperature respectively stabilizes or destabilizes the protein, as shown in figure 1.5 A for several proteins. We know that ΔG^u_{eq} decreases from positive to negative values while increasing temperature because the protein is unfolding (figure 1.5 A). This decrease is non-linear, and comes from the temperature dependence of enthalpy difference upon unfolding ($\frac{\delta \Delta H}{\delta T} = \Delta C_p$, figure 1.5 B) and of entropy difference upon unfolding (ΔS, figure 1.5 C). At constant pressure, equation 1.1 becomes :

$$\Delta G^u_{eq}(T) = \Delta G_0 - \Delta S_0(T - T_0) - \Delta C_p[T(ln\frac{T}{T_0} - 1) + T_0] \tag{1.2}$$

A separation of the enthalpic and entropic contributions provides a better understanding of the origin of the magnitude of ΔG^u_{eq} as well as its temperature dependance.

Enthalpy, H, is a part of the internal energy of the protein that can be given to or taken from the solution and is defined as the capacity to do non-mechanical work and to release heat. The enthalpy of each of the two states can't be measured experimentally, but ΔH can by calorimetric methods. If we look at the plot of ΔH per gram of several protein with temperature, as shown in figure 1.5 B, we can observe that ΔH increases linearly with temperature and is generally small and negative at low and room temperature and highly positive at high temperatures [61]. The slope, $\delta \Delta H / \delta T$, corresponds as noted above to the difference in heat capacity upon unfolding (ΔCp), so ΔH changes with temperature because of ΔCp. We can observe too that ΔH of distinct proteins converges at a temperature of about 110°C [62]. The physical contributions to ΔH has been intensively studied over the years [63, 64, 65, 66, 67, 68, 69, 29, 70, 61, 71] and we have now a good understanding of its properties with temperature. As Privalov

showed in 1979 [62] using experiments with variations of unfolding temperatures as a function of pH or denaturant (GuHCl) [72, 73], ΔH depends on the temperature at which the unfolding occurs. Two effects are mainly responsible for the observed enthalpy difference in protein folding : the disruption of internal interactions within the protein molecule (Van Der Waals, hydrogen bonds, etc ...) and the hydration of the groups that are burried in the native state and exposed in the unfolded state. The enthalpy contributions of hydration and disruption of the internal interactions have opposite signs and cancel each other almost completely. Hilser, Gomez and Freire [63] showed a simple parametric equation to define the ΔH at any temperature with certain structural parameters :

$$\Delta H_u(T) = (a(T).\Delta ASA_{apolar} + b(T).\Delta ASA_{polar}) + \Delta H_{protonation} + \Delta H_{other} \qquad (1.3)$$

The equation 1.3 expresses the ΔH_u as a linear combination of changes in the solvent accessible polar and apolar areas (respectively ΔASA_{polar} and ΔASA_{apolar}) [74] and changes due to protonation and some other effects. Accessible surface area is related to hydrophobicity, and has a major importance [74] in this equation, so ΔH increases with temperature because hydrophobicity of each state is different [63].

If we look at the plot of ΔS per gram of protein versus temperature, as shown in figure 1.5 C p 9, we can easily see that ΔS increases with temperature in a non-linear fashion, from negative to positive values. The $\Delta S(T)$ lines for several proteins seem to converge at a temperature around 110-130°C, depending on data used (110°C for Privalov [62], 114.6°C from Sturtevant's relation [75], 112.8°C for Baldwin [76]). Entropy is a macroscopic property of a system that is a measure of the microscopic disorder within the system of microscopic objects carrying energy, eg. atoms for proteins, as defined by the second law of thermodynamics. Entropy is a measure of the amount of energy which does no work during isothermal energy conversions. The second law of thermodynamics states that spontaneous natural processes increase entropy overall, and an increase of entropy corresponds to an irreversible change. The third law says that for a perfect crystal, as the temperature approaches absolute zero, entropy approaches 0.

Entropy is a parameter of principle importance for understanding protein structure, and it is clear that the entropy of protein unfolding in aqueous media is a sum of a parameter associated with the increase of configurational freedom in the polypeptide chain and another with the hydration of groups that become exposed to solvent during unfolding [77, 78, 79, 80, 72, 31]. These parameters were studied in detail and determined as both negative at room temperature, and proportional to ASA. So again, hydrophobic interactions seem highly important to explain protein stability, and the dependence of entropy on temperature. If we consider the folded protein, while increasing temperature, the disorder increases, and so does the entropy but less than for the unfolded protein. One contribution to the increase in ΔS_u with temperature is

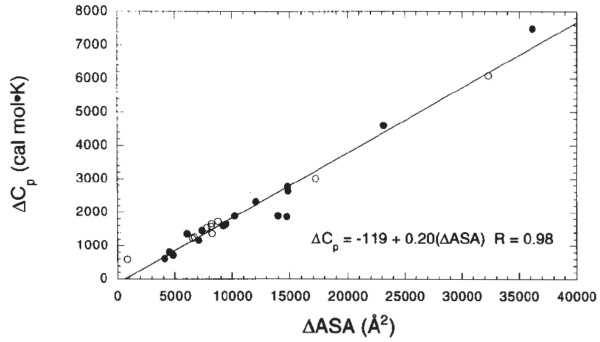

Figure 1.6: Dependance of heat capacity change upon unfolding on ΔASA for 45 proteins. Taken from [86].

the increasing chain entropy of the unfolded state relative to the folded state. Moreover, the hydration term becomes important at high temperatures. It is known that ΔS converges at about 110°C like ΔH [81], and Baldwin et al. [76] proposed that this behavior for entropy might be related to hydrophobic effects, this theory was confirmed later [82]. The reason of the intersect at 110°C for enthalpy and entropy is too suggested by Privalov [62] from the properties of the hydrophobic interactions. The non-linear temperature dependence of the entropy change upon unfolding is like the enthalpy change, due to the difference in heat capacity between the folded and unfolded states.

The temperature dependence of both ΔH_u and ΔS_u is due to $\Delta C_{p,u}$. $\Delta C_{p,u}$ is related as well to the accessible surface area (ASA), as shown in figure 1.6 with the strong correlation factor R=0.98. Hence, the heat capacity difference between the folded and unfolded states represents a key parameter defining the temperature stability of proteins.

The $\Delta C_{p,u}$ per gram of protein is linearly related to the fraction of hydrophobic residues and is independent of the temperature in the temperature range usually studied in experiments, from 5 to 80°C. The heat capacity of the unfolded state is larger than for the folded one, as commonly observed [83, 84, 85].

An equation of ΔC_p has been given [87]:

$$\Delta C_{p,u} = (0.45 \pm 0.02)(\Delta ASA_{apolar}) - (0.26 \pm 0.03)(\Delta ASA_{polar}). \tag{1.4}$$

So it's obvious than hydrophobicity is a major parameter for the ΔC_p. Moreover, another parameter, the m-value, is highly correlated with the ΔC_p, and we know that the amplitude of the linear dependence of stability on denaturant (m-value) is correlated as well with hydrophobicity through the accessible surface area.

Equation 1.4 is empiric, and from Gomez, Hilser, Xie and Freire [88], three major terms account for the absolute heat capacity of a protein : first, one term depends only on the primary or covalent structure of a protein and containing contributions from vibrational frequencies arising from the stretching and bending modes of each valence

bond and internal rotations; a second term contains the contributions of noncovalent interactions arising from secondary and tertiary structure and a third term contains the contributions of hydration. The hydration term contributes respectively about 15% and 40% to the total heat capacity of the native and unfolded states. The contribution of non-covalent structure to the total heat capacity of the native state is positive but very small. The change in heat capacity upon unfolding is primarily given by the increase of the hydration term (about 95%) and to a much lesser extent by the loss of noncovalent interactions [88, 89].

As we have just seen before, the large and positive $\Delta C_{p,u}$ value is commonly attributed to the hydrophobic hydration, although other factors may contribute [75, 90]. This is confirmed by Nemethy and Scheraga [30] who found that hydrophobic interactions become stronger while increasing temperature, so other factors may counteract.

The $\Delta C_{p,u}$ imparts a temperature dependence to entropy and enthalpy that may change their signs and which of them dominate, and the $\Delta C_{p,u}$ of a protein describes the amount of curvature in the plot of ΔG as a function of temperature. That's why protein unfolding usually has a positive $\Delta C_{p,u}$, producing a maximum in stability and sometimes observable cold denaturation.

From this discussion, we can clearly see that the interactions implicated in protein stability with temperature are intimately linked to hydration. Hence, we must clearly understand hydration changes associated with folding to understand the folding process, and pressure perturbations can provide useful information about it. As we have seen earlier, ΔG, ΔC_p, ΔH and ΔS are quite well known, but the other parameters of the equation of p,T stability dagram (ΔV, $\Delta\alpha$ and $\Delta\beta$) are still not well understood. In particular, hydrophobic hydration cannot explain the effect of pressure on protein stability. It is evident that without a complete understanding of this phase transition diagram including pressure effects, we cannot claim to have solved the question of protein folding [91, 92, 93, 54, 94, 95].

1.3.2 Pressure effects on protein stability

At constant temperature

Examination of the p,T stability transition diagram reveals that at a constant temperature, ΔG^u_{eq} changes with pressure. When we look at the equation defining the ΔG^u_{eq} versus p and T, with a constant T, we can observe that the temperature dependent terms ΔS, ΔC_p and $\Delta\alpha$ disappear, as we take $T = T_0$. So the only parameters which can explain this change in stability are the ones depending on pressure only : ΔV and $\Delta\beta$. We know from prior work that the compressibility difference upon unfolding between the folded and unfolded states, $\Delta\beta$, is not large because the two states are compressed with the approximately same magnitude, so that $\Delta\beta$ is very small [93].

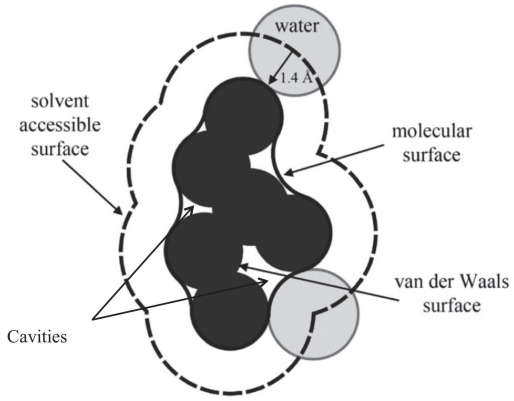

Figure 1.7: Volume composition of a protein

Hence, the only parameter which explains the effects of pressure at constant temperature is the volume difference between the folded and unfolded state upon unfolding : ΔV^u_{eq}. To understand better what happens to the protein's stability upon increasing pressure at constant temperature, we need to clearly define what is the protein molar volume and the ΔV^u_{eq}.

The protein molar volume (V_{pr}) and the volume change upon unfolding (ΔV^u_{eq}) V_{pr} depends on the protein, and is generally in a range of 0.7-0.8 $mL.g^{-1}$ [96, 97, 98, 99]. However, volume changes upon unfolding are small, typically less than 1% of the molar volume (or $<$0.008 $mL.g^{-1}$). To determine a protein molar volume (V_{pr}) is experimentally difficult. Typically, densitometry [97, 99, 100, 85] is used, althought it requires high concentration and accuracy depends on correctly knowing the exact concentration. Molecular dynamic simulations can provide some information as well, but are highly dependent on the type of force fields used and volume calculation methods chosen.

The molar volume of the protein water system is constituted of the volume of the protein and the volume of surrounding water, as described here :

$$V_{pr} = V_{protein} + V_{water} \tag{1.5}$$

where $V_{protein}$ is composed of the volume of the atoms ($V_{prot,atoms}$) and the free space between the atoms inside the structure ($V_{prot,void}$). In the protein, this space comes from the cavities and the mispacking of the structure. The volume of the surrounding water V_{water} is composed by that of the bulk water ($V_{w,bulk}$) and that of water of hydration ($V_{w,hyd}$). The density of water interacting with protein moieties can be significantly different from its density in bulk. So we can rewrite this equation as :

$$V_{pr} = (V_{prot,atoms} + V_{prot,void}) + V_{w,bulk} + V_{w,hyd} \tag{1.6}$$

Now that we defined the volume of a protein system, this definition is operative for

both the unfolded and the folded states of the protein. The volume of the folded state is different than that of the unfolded state, so the difference between those two states, the ΔV^u_{eq}, comes from differences in some of these terms. In order to understand the physical properties and basis of ΔV^u_{eq}, we need to understand from where this difference can come. Since the volume of the atoms can't change upon unfolding regardless pressure or temperature, this equation shows us that the ΔV^u_{eq} must arise from either changes in void volume or changes in the amount of hydration. So we can finally express the ΔV^u_{eq} as the resultant of these terms (based on [101]) :

$$\Delta V^u_{eq} = \Delta V_{voids} + \Delta V_{hydration} = \Delta V_{voids} + \Delta V_{polar} + \Delta V_{apolar} \tag{1.7}$$

where ΔV_{voids} is defined as the difference of volume of the two state of the protein induced by the water filling of the cavities and free spaces present in the folded state upon unfolding, and $\Delta V_{hydration}$ is the difference of the hydration volume upon unfolding, defined as as the change of the volume of the solvent due to hydration of the newly exposed accessible surface area with respect to its volume in bulk solvent. This term contains contributions from hydrophobic and polar hydration. We note that this term should be proportional to ΔASA. The change in void volume is a specific property of the folded state structure, whereas the change in volume due to differential hydration is a property of the unfolded state. The numerical value of ΔV^u_{eq} is still not understood because the parametrization of each factor is not known. However, we know that ΔV^u_{eq}, contrary to ΔC_p, is not strongly correlated with the number of residues (figure 1.8 A), or changes in the accessible surface area of the protein upon unfolding (figure 1.8 B), or the hydration. Nor is ΔV^u_{eq} strongly correlated with the largest cavity (figure 1.8 C). All this can be seen with the poor related correlations presented in figure 1.8 D. This shows us that the problem of the parametrization of the ΔV^u_{eq} remains a challenge.

With varying temperature

ΔV^u_{eq} changes with temperature It is known that the absolute value of ΔV^u_{eq} decreases with increasing temperature. This is due to the difference of thermal expansivity (α) between the two states : $\Delta\alpha$, the second order cross term in equation 1.1. The consequence of this is that pressure effects on protein stability are much more significant at room temperature than above. In fact, as noted above, at higher temperature, pressure can favor the folded state.

The thermal expansivity α of the two states If we look at the plot of the thermal expansivity α of a protein with temperature, as shown in figure 1.9 A obtained from pressure perturbation calorimetry (PPC), we can observe that the two states of the protein have their own expansivity and that these are rather strongly dependent on temperature. We can follow the expansivity of the folded state in conditions of temperature where the protein is totally folded (here from 0 to 50°C), and that of the

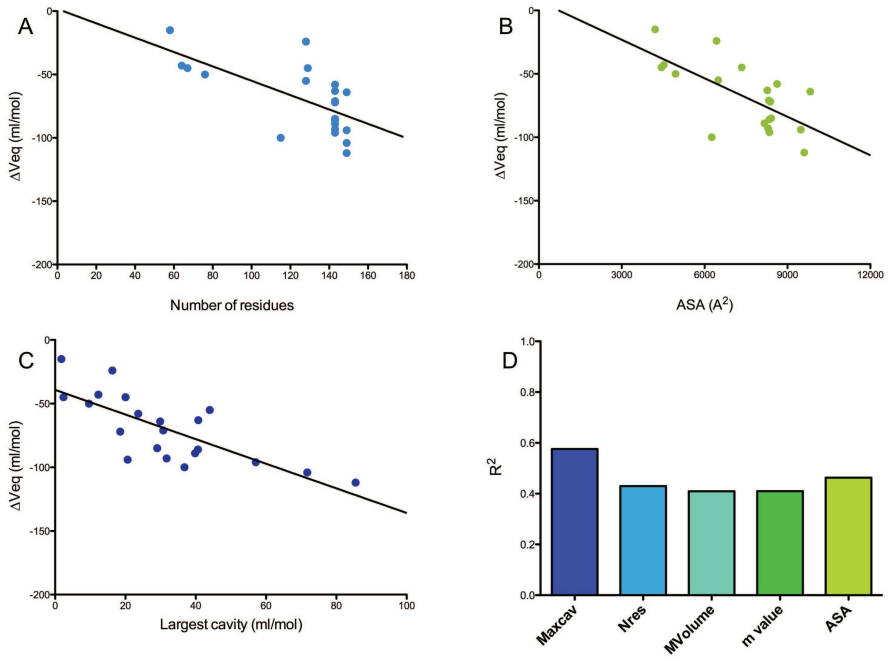

Figure 1.8: ΔV^u_{eq} evolution with distinct parameters. A : ΔV^u_{eq} depending on the number of residues. B : ΔV^u_{eq} depending on the accessible surface area. C : ΔV^u_{eq} depending on the largest cavity in the structure. D : Correlation factors of linearization depending on several parameters. Datas presented here come from CI2 [102], CPSB [103], Ubiquitin [104], P13 [105], Lysozyme [106], SNase [97, 107, 108] and unpublished ones. Some of these datas were recalculated from the 3D structures.

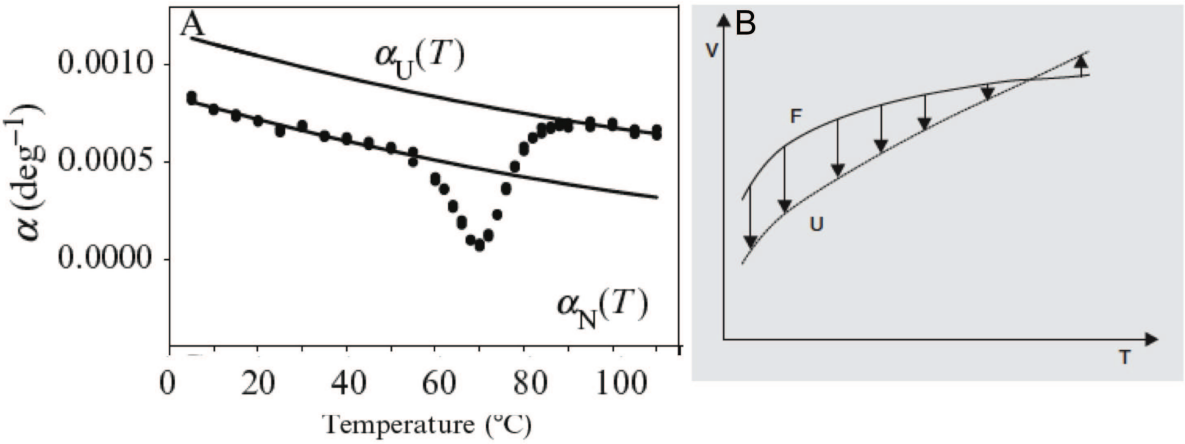

Figure 1.9: Expansivity and volume of the folded and unfolded state of proteins. A : expansivity of the two states of a protein, defined via PPC. B : Volumes of the folded and unfolded states of a protein. Based on [109, 100].

unfolded state when the protein is totally unfolded (here from 85 to 110°C). We extrapolate the value of α of the unfolded state at low temperature. Recently, Tsamaloukas, Pyzocha and Makhatadze described a parametrization equation of α of the unfolded state depending on the expansivity of each amino acid of the chain [110]. In general, α_u (expansivity of the unfolded state) is larger than α_f (expansivity of the folded state), and both decrease non-linearly with temperature in a parallel way. However, the linear temperature dependence of ΔV^u_{eq} for several proteins is consistent with a temperature independent $\Delta\alpha$ [111, 85, 112, 92, 97, 113].

The thermal expansivity of proteins, α, can be defined as the aptitude for the protein to expand when we increase temperature. The expansivity is physically described at constant pressure in function of the temperature by this equation :

$$\alpha = \frac{1}{V}(\frac{\delta V}{\delta T})_p \tag{1.8}$$

where V is the protein state's molar volume (for example the folded state molar volume if we look at the α of the folded state) at a defined temperature T_0, V the volume of the protein state, and δT the temperature difference between T and T_0. Upon increasing the temperature, the protein dehydrates because of the thermal motions and the entropy increase. The interactions between the hydrating water molecules and the protein surface decrease with temperature, leading to an overall decrease in the degree of hydration and net transfer of water molecules to the bulk. In addition, the protein itself expands in space. The more we increase temperature, the more the folded protein expands and the less the protein's expansion can increase without disrupting. In addition, as fewer water molecules interact, the effect of dehydration decreases with temperature. Thus overall α_f diminishes. The hydration term affects the α_u in a similar manner (see figure 1.9 A and B).

The thermal expansivity coefficient of a protein is found to be generally between $1.10^{-3}K^{-1}$ at low temperature and $0.5 - 0.6.10^{-3}K^{-1}$ after the thermal unfolding, at high temperatures [92, 96, 114]. Otherwise, the thermal expansivity coefficient can be quite different depending on the protein, between these limits. We still don't understand clearly the expansivity of the two states, the mechanisms involved, nor their exact parametrization. It is assumed that the major contribution to the expansivities of both the folded and the unfolded states arises from hydration effects, but as noted above, the structural expansion of the folded state may play a role as well particularly for the folded state.

Expansivity of water As shown in figure 1.10, at 3°C, water exhibits a negative expansivity [96]. At this low temperature, water is highly structured, much like ice. And ice floats, i.e. it is less dense than bulk water. Upon heating, transfer of water molecules

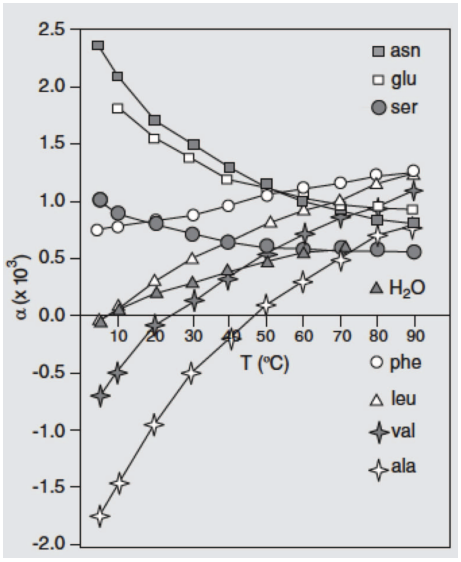

Figure 1.10: Expansivity of water and amino acids in solution depending on temperature [100].

to the bulk upon melting leads to an increase in density, and a decrease in volume, and hence a negative expansivity. This effect is large enough to overcome the temperature-induced increases in bond length. Above 4°C, the expansivity of water becomes positive because the low density structure is melting out and the effects of thermal motion come to dominate. Addition of proteins to water will alter the expansivity in a manner that depends upon whether the solutes favor the denser or the more structured form of their hydrating water.

Expansivity of amino acids and proteins Brandts et al. measured the thermal expansivity of hydrophobic, polar and aromatic amino acids [100, 96]. It appears from this study that the values of the thermal expansivities of proteins result from the combination of relatively large and opposite effects [115, 112]. It can be seen that hydrophobic moieties exhibit large, negative expansivity at low temperature that increases to positive values at high temperatures. The opposite is observed for polar hydration where electrostriction at low temperature leads to increased density (and so smaller volume) of hydrating water molecules. The expansivity of folded proteins appears to be dominated by the contributions of polar amino acids since the values are positive (near $1.10^{-3}K^{-1}$ at 4°C) and decrease in a manner similar to that of serine (figure 1.10), prior to the temperature-induced unfolding. Since unfolded proteins expose a larger fraction of hydrophobic surface area to the solvent than folded proteins, we can imagine that the expansivity of unfolded proteins should be lower than that of folded proteins at lower temperature and higher at high temperature, but this has not been demonstrated. Indeed, the expansivity of the unfolded state is assumed to be higher than that of the folded state, based on the extrapolation and the temperature independence of $\Delta\alpha$ discussed above.

So we can think that the contributions of hydrophobic interactions, polar hydration

and cavities contribute in opposite manner to expansivity.

The difference of thermal expansivity of the two states : $\Delta\alpha$ At low temperature, the hydration of hydrophobic residues exposed upon unfolding should lead to an increase in volume because the density of the water around the hydrophobic residues is smaller than the bulk. But, at the same time, exposure of polar residues (including peptide bonds) and their subsequent hydration, as well as the elimination of cavities should lead to a decrease in volume. At higher temperature, because the expansivity of the folded state has a stronger negative temperature dependence than the unfolded state, the volume of the folded state increases significantly at low temperature, but less as temperature is raised. So one hypothesis is that the evolution of ΔV^u_{eq} with temperature is due to the inability of the folded state structure to expand significantly, and the decreasing effects of hydration at higher temperature on the unfolded state. From equation 1.1, depending upon the magnitude of ΔV^u_{eq} and $\Delta\alpha$, it is possible that at high temperatures the application of pressure can actually lead to the folding of proteins [51, 52, 53, 57, 58, 59, 116, 91, 92, 93, 54]. The fact that the ΔV^u_{eq} can change it's sign was termed by Chalikian and Breslauer the "Volume change paradox" [101]. While recent work has provided some insight, much work remains to understand and parametrize the volumetric properties of proteins.

1.4 ΔV^u_{eq} : an incomplete understanding of pressure effects on protein stability

As clearly explained by Kauzmann [117] : *"A fundamental understanding of the behavior and properties of matter requires a fundamental characterization and comprehension of the response of the matter under consideration to the two fundamental thermodynamic variables, temperature and pressure"*. Pressure is an important thermodynamic parameter, commonly used to study a wide variety of chemical and physical processes [118, 119]. The importance of ΔV for protein folding studies was brought to light in 1914, when Bridgman unfolded ovalbumin with pressure [120]. High pressure has proven to be useful for protein unfolding, intermediate stabilization as for apo-myoglobin [121, 122, 123, 124], refolding of recombinant, oligomeric, and aggregated proteins. But we still don't understand exactly how pressure acts. Differences in molar volume between the protein conformational states constitute the basis for the effects of pressure on the equilibrium between these states. A quantitative parameterization of the volume changes, at the basis of pressure effects, accompanying protein conformational transitions will be required to better understand at the same time the protein folding and the effects of pressure [100]. A relatively large percentage of the proteins that have been studied using pressure perturbations can be unfolded completely in a pressure range

accessible in liquid cells (<10 kbar) [91, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134]. A large number of studies have been aimed at understanding the mechanism of pressure-temperature unfolding, with various proteins, like trypsin and chimotrypsin [135], bovine pancreatic trypsin inhibitor (BPTI) [136], ANS-apomyoglobin complex [122], ribonuclease A [52], Staphyloccocal nuclease [85, 137] and many others. All these proteins have one point in common : they are globular. It is difficult in such globular systems to test the importance of sequence, size and shape in determining volume changes because any significant change therein leads to massive structural disruption. So in order to understand better the effects of pressure on protein stability and the parametrization of the ΔV, we intend to use another model : a modular repeat protein.
The protein stability dependance on pressure doesn't seem to behave like the components of the hydrophobic model [31]. Moreover, the liquid hydrocarbon model [76] fails systematically if we extend it to the effects of pressure on protein folding [117], even if it helped a lot for the comprehension of the relative contributions of the intramolecular interactions, as the hydrophobic effect, for the chemical and thermal denaturation of proteins [80, 76, 138, 61, 139, 140]. Pressure holds the promise to inform on the contributing forces to protein stability.

Although the action of pressure on proteins was never clearly explained, the major hypothesis to explain the unfolding under pressure is that pressure acts by pushing the water molecule from the solvent inside the protein, favoring the smaller volume [141, 49]. Water, entering into the hydrophobic cores of the protein may not be acting directly on hydrophobic interactions, but would fill the cavities, making them disappear, leading to unfolding. This should lead to a decrease in molar volume of the protein, and disruption of interactions is accompanied by the unfolding of the protein. Hydration, the way water molecules interact with protein surface and in the protein interior is a really important part of the protein folding problem [142]. Hydration plays a major role in the native state [125] and transition state ensemble [143, 144, 145]. During the dynamics of folding, water is released from the protein surface and replaced mostly by hydrophobic interactions.

To better understand the ΔV^{u}_{eq}, we can try to dissect its contributions : the molar volume of the protein decreases with the exposition of the polar and charged groups to the solvent, because these amino acids will tend to form interactions with surrounding water. Volume can increase when we expose hydrophobic residues to solvent, indeed, the hydrophobic amino acids buried in the protein structure naturally form compact hydrophobic cores, including voids without any water. When the hydrophobic residues are exposed, the tight cores are broken, leading to smaller voids at each positions, for which the total sum is lower than the hydrophobic core size. Unfolded proteins present overall more surface area to the solvent, and their degree of solvation is larger than for folded proteins. In addition, folded proteins bury a higher percentage of non-

polar as opposed to polar residues in their native structures, and hence, the nature of the hydration of the protein surface is different in the folded state compared to more open, disordered states. Both the degree and type of hydration can contribute to the differences in specific volume between folded and unfolded states [100]. Another factor that increases volume is the exposition to solvent of the solvent excluded volumes, as some internal voids of the protein will be filled with water. Unarguably, proteins are highly densely packed in their folded states. Nonetheless, their interiors present void volumes that exclude solvent and protein moieties due to their size and distribution in the structure. Thus, the opening up of protein structure allowing solvent to occupy an internal volume previously excluded from solvent constitutes a clear contributor to changes in the degree of hydration upon changes in conformation. So the ΔV^u_{eq} should be the sum of all these opposite forces. Thus, in a general sense, whatever water is filling cavities or hydrating exposed surfaces, it leads to differences in volume. Hydration lies at the basis of pressure effects on folding. The challenge remains to evaluate the amplitude of the hydration and void volume terms in the overall pressure effect, and this at all temperatures.

By accessing these terms, we can obtain more insight about hydration dynamics, the thermodynamics of the folding and the energetics of the native, unfolded and transition state ensemble (TSE).

1.5 Objectives

Many questions remain unanswered about the ΔV^u_{eq}, it's physical properties and thermodynamical basis. All the facts exposed above drive us naturally to pose several questions about protein folding.

- What is the contribution of voids and cavities to ΔV^u_{eq}, and how is this contribution distributed in the protein structure ?
- Are all the voids important to the ΔV^u_{eq}, or only some particular ones ? And how much ?
- How do size and sequence affect the differential hydration contribution to ΔV^u_{eq} ?
- Does the shape of the protein influence its ΔV^u_{eq} ?
- Can we imagine to determine precisely ΔV^u_{eq} for proteins ab initio ?

By using model systems to try to answer these specific questions we hope to gain significant insight into the physical basis for pressure effects on proteins. Hence, we decided to use a model protein that permits to dissect all the parameters better than globular proteins : the ankyrin repeat protein.

Chapter 2

A model for probing pressure effects on protein folding : the Ankyrin repeat domain of Notch Receptor

2.1 Advantages of repeat proteins for studies on protein folding

As previously described, globular proteins have been used intensively for the studies of protein folding and in particular for studying pressure effects on protein stability. However, a major problem exists that is inherent with these proteins. Indeed, as we can see in the middle part of figure 2.1 A, for globular proteins like RNase A, the 3D structure (left part of figure 2.1 A) leads to interactions between amino acids that are close within the structure but far in terms of sequence. For example, contact order maps (middle of figure 2.1 A) show that residue 100 connects to residue 20 and 75, or residue 80 connects to residues 19, 50 and 120. All these long range interactions make the protein folding problem more complex. The folding energies are hard to dissociate because close contacts between distinct chains promote the folding cooperativity, and the local loss of structure destabilize the whole ensemble of protein regions. Hence, if we want to study precisely the individual contributions to protein folding, modular proteins for which each module only interacts with its close neighbors in sequence are highly advantageous. Hence, as we can see in figure 2.1 B and C for two repeat proteins, the contact orders are easier to analyze, with only close contacts interactions which are symmetric and can easily be modeled.

Moreover, globular proteins are complex, irregular and made of heterogeneous architectures, as reported in the left part of figure 2.1 A. These complex architectures, compared to simple architectures of repeat proteins (left part of figure 2.1 B and C), make protein folding studies more difficult and make the comparison between different globular proteins more complex, as different 2D structures are mixed within the final

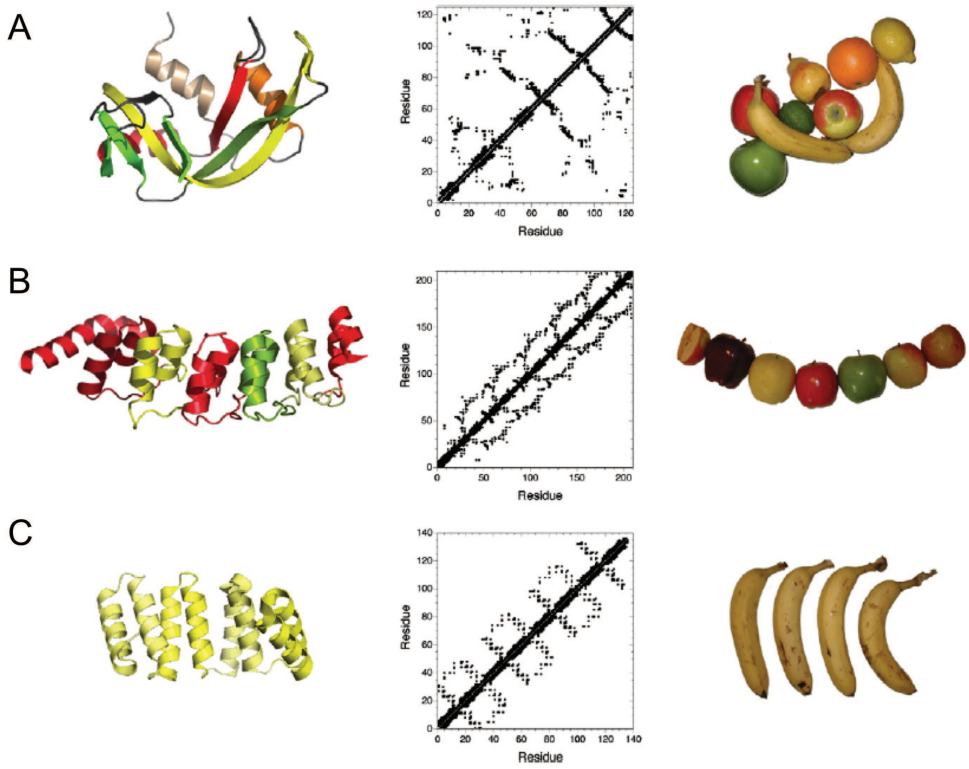

Figure 2.1: Comparison between structures of globular protein and repeat proteins. (A) RNase A, a globular protein (7RSA.pdb), is compared to repeat proteins (B) Ankyrin 1-7 $\star$, our model protein (1OT8.pdb) and (C) 2FO7.pdb a consensus TPR protein. On the left, the ribbon diagram showing the secondary and tertiary structures. In the middle, the contact order map, highlighting the lack of long-range interactions for repeat proteins and the regular pattern of tertiary structure. On the right, a "fruit-mix" vision of the tertiary and secondary structure of the proteins, to make a better visual comparison of the dissection problem. This figure was taken from [146].

3D structures. It is like trying to compare apples and oranges (right part of figure 2.1 A) instead of apples only (right part of figure 2.1 B). This heterogeneity brings up another problem : the *comparison* problem between proteins. Nowadays, with the increasing number of proteins used and for which we understand quite well the folding mechanism, the challenge is to understand the general principle of protein folding. As noted by Baker et al. [24], the folding rules should be simpler than we thought because of this apparent complexity in globular proteins. It's generally difficult to find a general principle from several datasets and proteins, and if we want to compare the thermodynamic variables between proteins, we need to use proteins with similar architectures.

Proteins containing motif repetitions present major advantages for protein folding studies because of the abolition of the *dissection and comparison* problems, as they present repetitions of structurally identical motifs, in a linear, modular and simple packing. The *comparison* problem disappear because the contributions of each motif can be easily evaluated at all the levels. As there are no long-range interactions

between separated parts of these proteins, the *dissection* problem disappears as well, and the study of the folding parameters become largely simplified.

A last advantage of this kind of repeat proteins is that we can insert easily mutations within the sequence that will only have local effects. This leads to an accurate analysis of the integrality of the folding process, depending on the parameters we change. For example, we can easily change the number of repetitions [147] or insert a unique mutation in one only motif, or to the contrary insert similar mutations in all the motifs at the exact same conserved positions.

Hence, we can expect to answer our questions and to deduce convincing general elements about the protein folding mechanism and in the instance, the effects of pressure. Our choice as model protein was the ankyrin domain of the Notch Receptor : Nank. With this model system, we hope to be able to test the additivity of the ΔV^u_{eq}, answer questions concerning the contributions of polar and hydrophobic hydration, test the effects of cavities, and in so doing, lay the foundations of the parametrization of ΔV^u_{eq}.

2.2 Ankyrin repeat proteins

2.2.1 Background of the ankyrin repeat proteins

The ankyrin motif, whose name comes from the cytoskeleton adaptor protein Ankyrin that contains 24 copies of this so-called motif, is found in tandem in a large variety of proteins responsible of protein-interaction signal mediation, like the Notch receptor (the Notch pathway is a ubiquitous transmembrane signaling pathway involved in a variety of cell-fate decisions in higher eukaryotes [148, 149, 150, 151, 152, 153, 154]), the subunit β of the GA-binding protein transcription factor [155, 156], the cyclin-dependent kinases inhibitors p16, p18, p19 [157, 158, 159], etc ... These ankyrin repeated motifs are known to play important biological roles, in particular in the signaling and anti-tumoral activity fields, and it has been shown recently that the folding mechanism of these repeats has a major impact on their functions [160]. Although our major interest is more fundamental than the specific activity of these structures, it is evident that a good understanding of the folding of these repeats allows a better global insight on the biological mechanisms that ensue from their structure and stability.

2.2.2 Sequences of ankyrin motifs

In the Notch receptor of Drosophila, seven ankyrin repeats (six standards and a putative one) have been identified from analysis of the primary sequence [161, 162], based on the Ankyrin consensus sequence (see table 2.1). Weak similarity to the ankyrin repeat consensus exists in the sequence immediately C-terminal to the sixth ankyrin

Consensus[a]	xtx**OULHøAø**xxttxt**øø**tx**LL**txtxtxxxxxx
2° structure[b]	bb aaaaaaaaa aaaaaaaaaa bb
Repeat 1	cgl**TPL**m**IAA**vrggzq**VI**sd**LL**aqgaelnatmdk
Repeat 2	tge**T**s**LHLAA**rfarad**AA**kr**LL**dagadancqdn
Repeat 3	tgr**TPLHAAV**aadamg**VF**qi**LL**rnratnlnarmh
Repeat 4	dgt**TPL**i**LAA**rlaieg**MV**ed**L**itadadinaadn
Repeat 5	sgk**TALHWAA**avnnte**AV**ni**LL**mhhanrdaqdd
Repeat 6	kde**TPL**f**LAA**regsye**AC**ka**LL**dnfanreitdh
Repeat 7	mdrl**P**rd**VA**serl.hhd**I**vr**LL**dehvprspqml

Table 2.1: Sequence of the 6 identified and the seventh putative Ankyrin Repeat form Drosophila melanogaster Notch. a) Repeat boundaries are from Bork [163] except for repeat 7, which is the 32 residues immediately following the sixth repeat identified in [163]. x = any residue, t= turn, O : Ser or Thr, U = Pro or Ala, ϕ = hydrophobic, z = insertion of 16 residues (gldtgediennedsta), . means gap. Residues that match the consensus are in upper-case, bold. b) Secondary structure asignments (a = α-helix, b = β-strand) are based on comparisons of sequences of ankyrin repeats whose structures have been determined to high resolution [164, 156, 165, 166, 167, 168].

repeat of Drosophila Notch, although this C-terminal region is not identified as a standard ankyrin repeat in the multiple sequence alignment of Bork [163], but is taken into account as a putative seventh repeat. It was suggested that the low sequence similarity of this C-terminal putative seventh repeat may be a result of its role as a "capping" ankyrin repeat which interacts with solvent on one face, rather than with another ankyrin repeat. The six motifs and the putative seventh one are present in our construct $Nank1-7^{\star}$.

The sequence conservation among the seven ankyrin repeat sequences of the Notch ankyrin domain is modest (17% average pairwise identity [148], table 2.1); however the sequence conservation of analogous repeats from different taxa is quite high (approximately 70% identity [170, 148, 152], as shown in figure 2.2). Because of this low conservation in the sequence of all the 7 motifs, we figured that we will be able to test the importance of the sequence on the ΔV_{eq}^{f} in an unbiased approach, by taking in consideration the natural sequence diversity and evolutions. However, identification of the precise beginning and ending of ankyrin repeats from primary sequence can be difficult [163].

2.2.3 Structure of the ankyrin motif and protein

Although the sequence variability of ankyrin repeats is high in our constructs, the three-dimensional structure of ankyrin repeat domains is quite similar [164]. The ankyrin motif is made of 33 amino acids, and is folded in two antiparallel α-helices, connected to each other by a short loop, as shown in figure 2.3 A. These motifs are repeated several times in tandem, as shown in figure 2.3 B and C, and each motif is linked to the adjacent one by a short β-hairpin. So ankyrin repeat proteins have a regular topology, consisting of regularly spaced modules where close contacts are

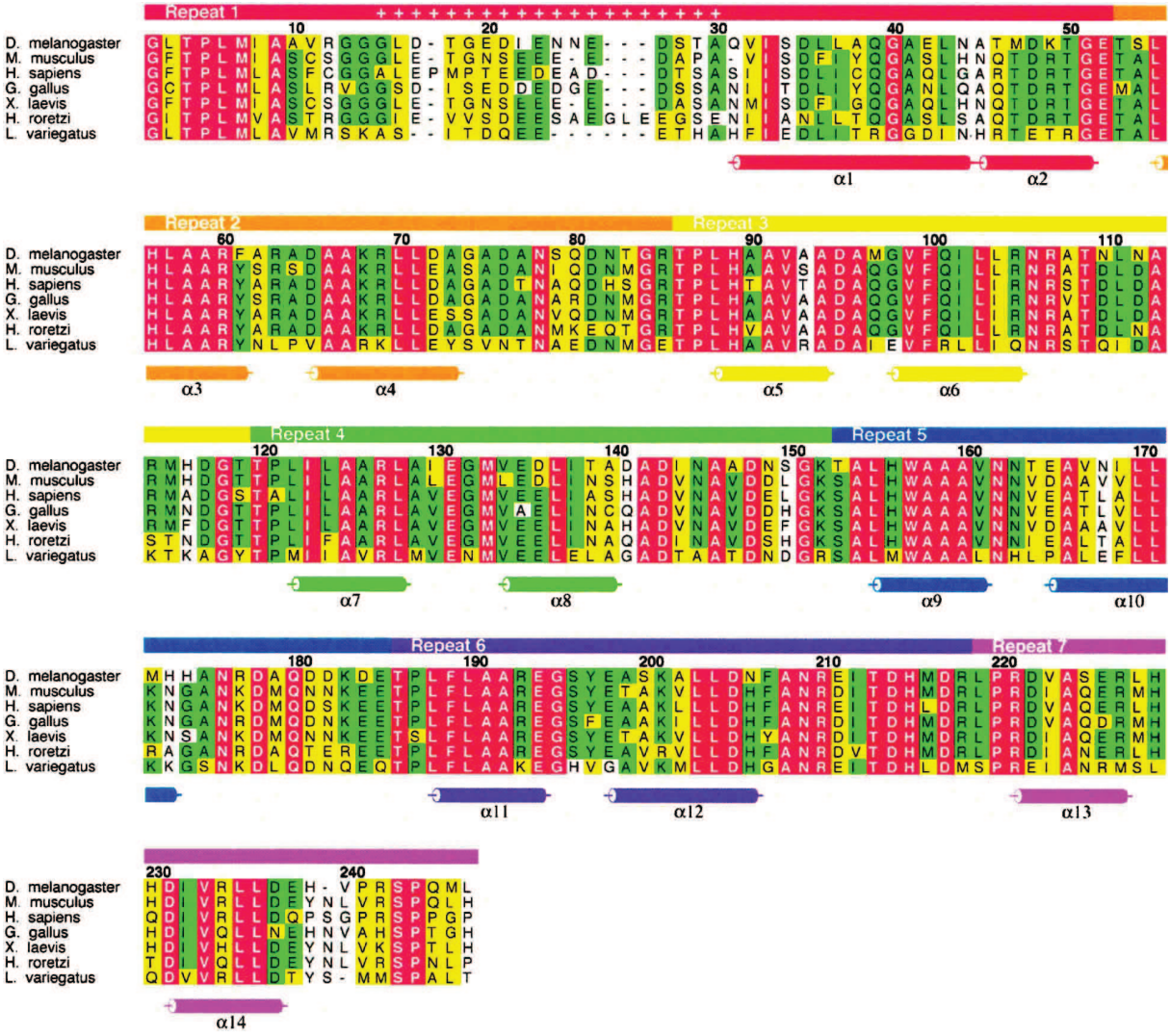

Figure 2.2: Sequence alignment of Notch ankyrin domains from different taxa. The bar above the alignment indicates the boundaries of the seven ankyrin sequence repeats. Insertions in the first repeat are indicated by plus symbols in the bar above the alignment. α-helices are shown below the sequences as determined using DSSP. Absolutely conserved residues are colored red, positions with greater than 50% identity are colored green, and positions that contain conservative amino acid substitutions are colored yellow. The numbering refers to the Drosophila ankyrin construct studied here; residue 1 corresponds to residue 1902 of the full length Notch receptor. This figure was made using the program ALSCRIPT. Figure from [169]

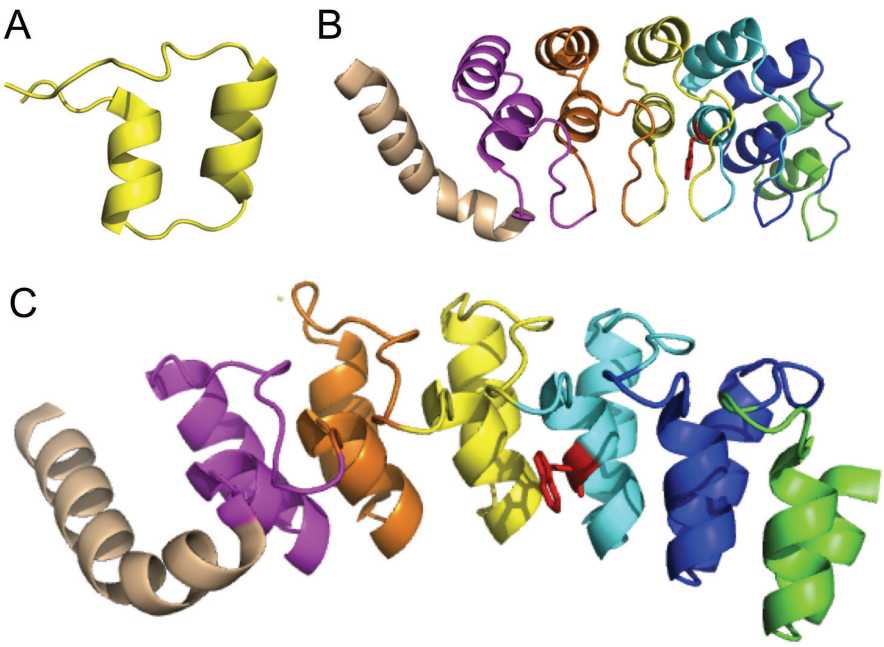

Figure 2.3: Structure of ankyrin motifs and ankyrin repeat protein. A : a single ankyrin motif, composed of two $\alpha-helices$, joined together by a short loop. B : Crystallographic structure of our ankyrin repeat protein, composed of 7 motif repetitions, seen from the top. C : Lateral view of the protein structure. The single tryptophan, residing in the fifth repeat (cyan), is shown in red. The motifs are colored as follows : the first partially folded N-terminal motif in pink, and the second to seventh motif from purple to green.

made by residues that are close in primary sequence, as previously explained [148, 152]. The packing of these motifs creates a tight hydrophobic core at the interface between two motifs, that we call hydrophobic interfaces. Actually, we know the 3D structure of $Nank1-7^{\star}$ [169] (PDB : 1OT8) , and of one of it's mutant, $Nank4-7^{\star}$ [171] (PDB : 1YMP) which conserves the same architecture. $Nank1-7^{\star}$ contains seven ankyrin motifs. The first repeat (N-terminal one, pink in figure 2.3 C) is known to adopt a partially unfolded non regular ankyrin fold from the crystallographic structure even though its sequence is recognized as a standard ankyrin sequence [172, 169]. Surprisingly, the putative seventh ankyrin repeat (C-terminal one, in green in figure 2.3 C) appears to be well folded despite its low sequence identity with the ankyrin consensus. In the middle of the fifth motif (cyan in figure 2.3 C) is found a unique tryptophan, highly important for our fluorescence studies. Despite it's position, where it's side chain points out of the protein, this tryptophan is dehydrated in the native state, because of the packing of the side chains of its neighbors which blocks the water entry into the structure.

2.3 Following the unfolding by fluorescence spectroscopy

The presence of a unique tryptophan in the core of the fifth motif is very important for our intrinsic tryptophan fluorescence experiments under pressure, as described

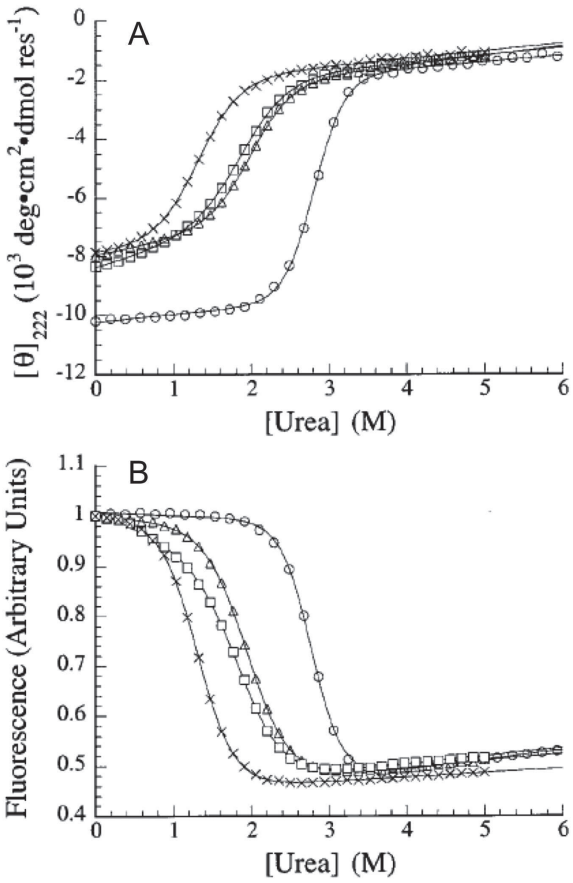

Figure 2.4: Urea-induced unfolding of ankyrin repeat proteins, monitored by CD and fluorescence. Unfolding transitions of Nank 1-7 * (circle), Nank 1-6 * (triangle), Nank 1-6 (cross) and Nank 1-5* (square), monitored by (A) circular dichroism and (B) Fluorescence spectroscopy. Results of the analysis of these datas are in table 2.2.

in materials and methods (2.5 p 42). Urea-induced unfolding transitions of Drosophila Notch ankyrin repeat polypeptides have cooperative, sigmoidal shapes when monitored by both Circular Dichroïsm (CD) spectroscopy regarding the α-helix content (figure 2.4 A) and tryptophan intrinsic fluorescence, which measures the polarity of the tryptophan in repeat 5, as shown in figure 2.4 B. Both CD and fluorescence unfolding curves can be fitted to estimate stability as a function of urea concentration. Furthermore, the team of Doug Barrick found that for a given ankyrin repeat polypeptide, similar thermodynamic parameters ($\Delta G^0_{H_2O}$, m-value) are obtained from CD and fluorescence-monitored transitions, as provided in table 2.2 [152]. Thus, the tryptophan faithfully reports on the chemical denaturation of $Nank1-7^*$.

$Nank1-7^*$ does not appear to self-associate below 200 μM, such that studies carried out at concentrations lower than that should be reversible [148].

	$\Delta G°_{H_2O}$ (kcal·mol^{-1})		m (kcal·mol^{-1}·M^{-1})	
	CD	fluorescence	CD	fluorescence
Nank1−7*	8.03 ± 0.09 (n = 6)	7.99 ± 0.04 (n = 4)	2.79 ± 0.02 (n = 6)	2.80 ± 0.01 (n = 4)
Nank1−6*	4.14 ± 0.05 (n = 4)	4.04 ± 0.01 (n = 3)	2.06 ± 0.02 (n = 4)	2.05 ± 0.01 (n = 3)
Nank1−6	2.74 ± 0.04 (n = 3)	3.00 ± 0.07 (n = 3)	2.05 ± 0.02 (n = 3)	2.29 ± 0.05 (n = 3)
Nank1−5*	3.55 ± 0.25 (n = 3)	3.70 ± 0.16 (n = 3)	1.85 ± 0.10 (n = 3)	1.97 ± 0.04 (n = 3)

[a] Parameters estimated by fitting the linear free energy equation (eq 1) to the unfolding transitions. Reported values are the mean of the number of independent unfolding transitions analyzed (n, as given in parentheses). Uncertainties are standard errors on the mean, that is, the standard deviation divided by $n^{1/2}$. Conditions: 25 mM Tris·HCl, 150 mM NaCl, pH 8, 4 °C.

Table 2.2: Urea-induced unfolding parameters for Ankyrin repeat polypeptides from the Drosophila Notch Protein Monitored by CD and Fluorescence

2.3.1 Existing mutants used in this work

During the last decade, a large number of mutants and variants were created by the team of Doug Barrick, based on the original $Nank1-7$ [172, 173, 174, 175]. These variants were designed to elucidate the thermodynamic and kinetic folding mechanisms. Early in the course of the studies, cysteines were replaced in the sequence to improve stability [148], and when these cycteines are replaced, a * is added to the name of the mutant. As you can see for example in figure 2.4 B, the urea-induced unfolding of $Nank1-6^{\star}$ (triangles), the one with mutated cysteines, is clearly shifted to higher urea concentrations than for $Nank1-6$ (cross), proving its increased stability.

In order to make the nomenclature of all the mutants easier, we will use a notation dependent on the motifs which are present in the structure, as follows. For example, the main original protein is the N-terminal ankyrin part of the Notch receptor, containing motifs 1 to 7, so we will call it Nank1-7. As we wanted to use stable proteins during long time periods as required by our experiments, we preferred using the replaced cysteines variants. Hence, our base for this work was the largest construct we have with replaced cysteines : $Nank1-7^{\star}$.

$Nank1-7^{\star}$ as a reference for pressure-induced unfolding studies

$Nank1-7^{\star}$ is the most stable ankyrin repeat protein of this set of proteins, as we can observe in figure 2.4 B that the middle of its denaturation curve with urea is at about 2.8M of urea. This is confirmed by the large value of its $\Delta G^0_{H_2O}$ at 4°C (8.03 (CD) or 7.99 (fluorescence) $kcal.mol^{-1}$) and at 20°C (6.7 $kcal.mol^{-1}$). We can observe too that its folded state plateau is fairly parallel to the x-axis, so that there is no change in the structure of $Nank1-7^{\star}$ at low urea concentrations, and that its unfolding appears to be two-state both by fluorescence and CD experiments.

N-terminal truncated constructs

$Nank1-6^{\star}$ is a construct when the seventh putative repeat was deleted. It is stable in solution with a $\Delta G^0_{H_2O}$ of 4.1 $kcal.mol^{-1}$ at 4°C and 2.8 $kcal.mol^{-1}$ at 20°C (see tables

	Circular dichroism		Fluorescence	
Construct	ΔG°_{H2O}, kcal·mol^{-1}	*m* value, kcal·mol^{-1}·M^{-1}	ΔG°_{H2O}, kcal·mol^{-1}	*m* value, kcal·mol^{-1}·M^{-1}
Nank1–7*	−6.65 ± 0.04 (5)†	−2.85 ± 0.02	−6.78 ± 0.6 (5)	−2.91 ± 0.03
Nank1–6*	−2.85 ± 0.03 (4)	−1.76 ± 0.01	−3.01 ± 0.03 (4)	−1.85 ± 0.02
Nank1–5*	−2.69 ± 0.04 (5)	−1.73 ± 0.03	−2.84 ± 0.16 (3)	−1.67 ± 0.08
Nank1–4*	+0.37	−0.85	ND	ND
Nank4–7*	+0.089	−2.12	ND	ND
Nank3–7*	−1.81 ± 0.09 (3)	−1.58 ± 0.06	−1.70 ± 0.23 (3)	−1.67 ± 0.08
Nank2–7*	−4.96 ± 0.04 (8)	−2.36 ± 0.02	−5.05 ± 0.04 (5)	−2.44 ± 0.03
Nank2–6*	−1.96 ± 0.002 (3)	−1.58 ± 0.02	−2.03 ± 0.02 (3)	−1.67 ± 0.02
Nank2–5*	−1.73 ± 0.03 (6)	−1.39 ± 0.02	−1.54 ± 0.03 (3)	−1.34 ± 0.01

Numbers in parentheses denote number (*n*) of independent folding curves. ND, not determined.
†ΔG°_{H2O} for Nank1–7* differs from that reported in ref. 25, which was determined in the presence of 150 mM NaCl.

Table 2.3: Thermodynamic parameters for folding of deletion constructs of the Notch ankyrin domain, at 20°C.

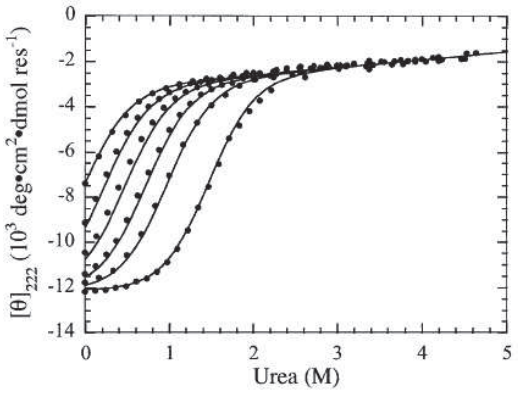

Figure 2.5: Structural transitions of $Nank4-7^{\star}$ in a mixed urea/TMAO solvent system monitored by circular dichroism at 222 nm. Urea-induced denaturation in increasing TMAO concentrations (from left to right at 0, 0.125, 0.25, 0.375, 0.5, 0.75 M TMAO).

2.2 and 2.3), and the middle of its urea-induced unfolding transition is at about 2M of urea, as shown in figure 2.4 B.

$Nank1-5^{\star}$ lacks the two C-terminal repeats, and is stable with a $\Delta G^{0}_{H_2O}$ of 3.5 $kcal.mol^{-1}$ at 4°C and 2.7 $kcal.mol^{-1}$ at 20°C (see table 2.2 and 2.3).

C-terminal truncated constructs

$Nank2-7^{\star}$ has the six N-terminal repeats, lacking the first partly folded repeat. As we can see in table 2.3, this mutant is stable with a $\Delta G^{0}_{H_2O}$ upon unfolding of 4.9 $kcal.mol^{-1}$ at 20°C [176].

$Nank3-7^{\star}$ is a short construct lacking the first two repeats and is stable in solution with a low $\Delta G^{0}_{H_2O}$ of 1.8 $kcal.mol^{-1}$ at 20°C (see table 2.3).

$Nank4-7^{\star}$ is a particular construct. Urea-induced unfolding experiments at ambient temperature produce only a partial unfolding transition that lacks a native baseline (leftmost transition in figure 2.5) [175], so this mutants is not stable. Increasing TMAO concentrations shifted the urea-induced unfolding transitions of this protein to higher urea concentrations, without affecting the apparent steepness of the transitions. The

	$\Delta G°_{u,H_2O}$ (kcal mol^{-1})		m-Value (kcal mol^{-1} M^{-1})	
	CD	Fluorescence	CD	Fluorescence
WT	5.86 ± 0.06	5.85 ± 0.05	2.68 ± 0.03	2.66 ± 0.02
NAG1	6.07 ± 0.06	6.04 ± 0.03	2.82 ± 0.01	2.77 ± 0.02
NAG2	2.35 ± 0.04	2.35 ± 0.02	2.57 ± 0.02	2.65 ± 0.03
NAG3	2.22 ± 0.04	2.36 ± 0.03	2.65 ± 0.05	2.83 ± 0.05
NAG4	3.54 ± 0.04	3.76 ± 0.05	2.97 ± 0.03	3.13 ± 0.04
NAG5	3.77 ± 0.04	3.95 ± 0.06	2.93 ± 0.04	3.05 ± 0.05
NAG6	3.10 ± 0.06	3.79 ± 0.09	1.84 ± 0.01	2.15 ± 0.02
NAG7	3.4 ± 0.2	3.39 ± 0.03	1.8 ± 0.1	1.90 ± .0.01

Parameters were estimated by fitting equation (8) to urea unfolding transitions; reported parameters are mean values from three or more independent unfolding transitions. Uncertainties are standard errors on the mean. Conditions: 25 mM Tris, 150 mM NaCl (pH 8.0), 20 °C.

Table 2.4: Urea-induced unfolding parameters for Ankyrin repeat polypeptides from the Drosophila Notch Protein Monitored by CD and Fluorescence

value of $\Delta G^0_{H_2O}$ determined from the global fit for $Nank4-7^\star$ is 0.09 $kcal.mol^{-1}$, so that without urea, the folded and unfolded states are equiprobable. Adding 750 mM of TMAO leads to complete folding of $Nank4-7^\star$, as we can see in figure 2.5 (rightmost part) [175]. Trimethylamine N-oxide (TMAO) is a naturally occurring osmolyte that is found in several marine organisms containing elevated intracellular urea concentrations [177]. Numerous studies have investigated the effect of TMAO on proteins and described its stabilizing effects [178, 179], and defined that TMAO promotes folding to specific, biologically relevant native states [180] because interactions between the TMAO and the peptide backbone destabilize the denatured state, shifting equilibrium toward the native state [181].

Alanine to glycine mutants based on the complete protein $Nank1-7\Delta$: NAGs

In this work we will also use two alanine to glycine mutants, NAG2 and NAG5. This destabilizing mutation is placed respectively in the second and fifth repeats, but they are stable with $\Delta G^0_{H_2O}$ of respectively 2.3 and 3.7 $kcal.mol^{-1}$ at 20°C, as reported in table 2.4. These mutants were originally created to determine the limits of cooperativity, and the substitutions were of analogous alanine residues in the first helix of each repeat, allowing the same perturbation to be examined at different positions in the protein [172].

What is known about these variants

$Nank1-7^\star$ was found to have a free energy of unfolding nearly twice that of $Nank1-6^\star$, lacking the C-terminal repeat [152], suggesting that this putative seventh ankyrin repeat is an integral part of the Notch ankyrin domain.

The observation that inclusion of the putative seventh ankyrin repeat shifts the unfolding transition to higher denaturant concentration and temperature (figures 2.4 B and 2.6) provides evidence for strong cooperative interactions in the Notch ankyrin domain [152]. The entire unfolding transition is shifted to strongly destabilizing conditions and is sharpened. Thus, the entire ankyrin domain (including regions near the

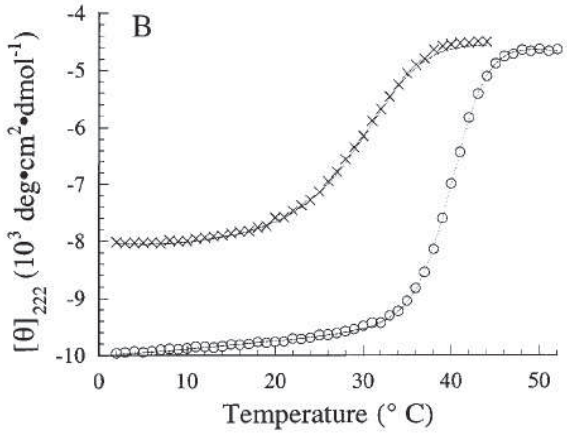

Figure 2.6: Thermal unfolding of the ankyrin repeats of Drosophila Notch. Comparison of thermal unfolding transitions for $Nank1-7^{\star}$ (circles) and $Nank1-6^{\star}$ (cross) in 1 M urea. Conditions: 150 mM NaCl, 25 mM TrisHCl, pH 8.1M

N-terminus) is stabilized by the putative seventh, C-terminal repeat.

The small change in stability between $Nank1-6^{\star}$ and $Nank1-5^{\star}$ that results from inclusion of the sixth repeat in the $Nank1-5^{\star}$ sequence, compared with the large change resulting from inclusion of the putative seventh repeat, may suggest that different ankyrin repeats contribute different amounts to the overall stability of the domain. All of the structural and thermodynamic data are consistent with a model in which the C-terminal region folds as a seventh ankyrin repeat, and that the sixth repeat is only folded if the seventh repeat is included [152].

Sedimentation velocity data are also consistent with partial unfolding of the sixth repeat in the absence of the putative seventh repeat. The variation of measured Notch ankyrin shape factors (and computed axial ratios) with repeat number is inconsistent with a simple model in which deletion of ankyrin repeats simply produces a shorter array of folded repeats, but is consistent with the a model in which the sixth repeat is unfolded in the absence of the seventh repeat (i.e., $Nank1-6^{\star}$) [148]. In addition to an increase in $\Delta G^0_{H_2O}$, the steepness of the unfolding transition is greater for $Nank1-7^{\star}$ than for $Nank1-6^{\star}$. This increase in steepness is reflected in the fitted m-values: $Nank1-7^{\star}$ has an m-value of 2.79 $kcal.mol^{-1}$, compared to 2.05 $kcal.mol^{-1}$ for $Nank1-6^{\star}$ (table 2.2). As m-values are often considered to be measures of the change in the exposed surface area on unfolding [182, 86], these results suggest a significantly larger change in the exposed surface area on unfolding of $Nank1-7^{\star}$ than $Nank1-6^{\star}$ [152], and an equal exposed surface for $Nank1-5^{\star}$ and $Nank1-6^{\star}$. All these results converge to identify that deletion of repeat seven from the Notch ankyrin domain results in partial unfolding of repeat six, and in decreases in free energy, enthalpy, entropy and heat capacity of unfolding of around 50% each [183, 149]. However, further deletion of repeat six does not significantly affect the structure of the five remaining N-terminal ankyrin repeats [172, 148, 152].

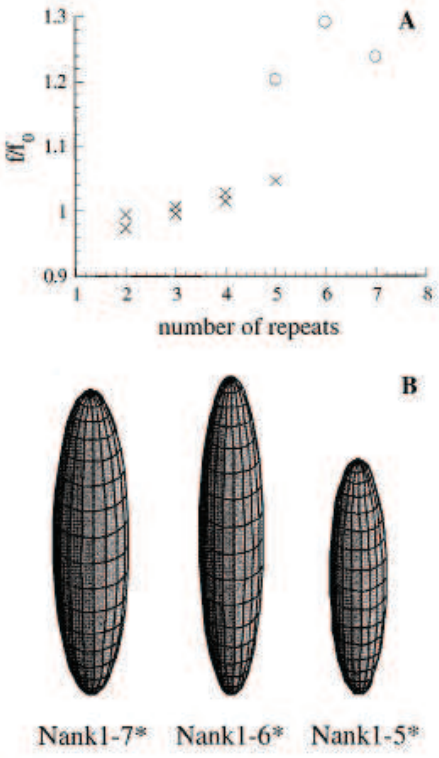

Figure 2.7: Experimental and calculated hydrodynamic properties of Notch ankyrin repeat proteins. A : Comparison of experimentally determined shape factors (f/f_0) for Notch ankyrin polypeptides (circles) with values calculated from bead models based on ankyrin repeat proteins of known structures (crosses). B : Prolate ellipsoids of revolution with frictional properties matching experimentally determined values for the Notch ankyrin constructs.

Frictional studies confirms this point of view. Indeed, frictional properties of $Nank1-7^*$, $Nank1-6^*$, and $Nank1-5^*$ differ substantially from one another, and suggest that the molecular dimensions of these polypeptides do not simply vary in proportion to the number of repeats. Instead, the determined shape factor (f/f0) is larger for $Nank1-6^*$ than for $Nank1-7^*$, indicating that removal of the putative seventh ankyrin repeat increases the frictional coefficient of the remaining polypeptide relative to the corresponding hydrated sphere as shown in figure 2.7. To match the experimentally determined hydrodynamic data, the prolate ellipsoid that is frictionally equivalent to $Nank1-6^*$ would have to be more elongated than the ellipsoid matching the frictional properties of $Nank1-7^*$ (figure 2.7 B). So a possible explanation for the low observed ellipticity is that one of the repeats on the N-terminus of $Nank1-6^*$ is disordered.

For $Nank1-7^*$, the wavelength of maximum emission in the tryptophan fluorescence spectrum is 333 nm [148], indicating that the environment surrounding the tryptophan is significantly nonpolar. For the shorter Notch ankyrin constructs ($Nank1-6^*$, $Nank1-5^*$), there is a small shift to longer wavelength, with a maximum emission wavelength of 335.5 nm for both $Nank1-6^*$ and $Nank1-5^*$. These results indicate a slight perturbation around the repeat 5 tryptophan of $Nank1-6^*$ and $Nank1-5^*$, and quenching experiments [148] indicated that for $Nank1-6^*$ and $Nank1-5^*$, the trypto-

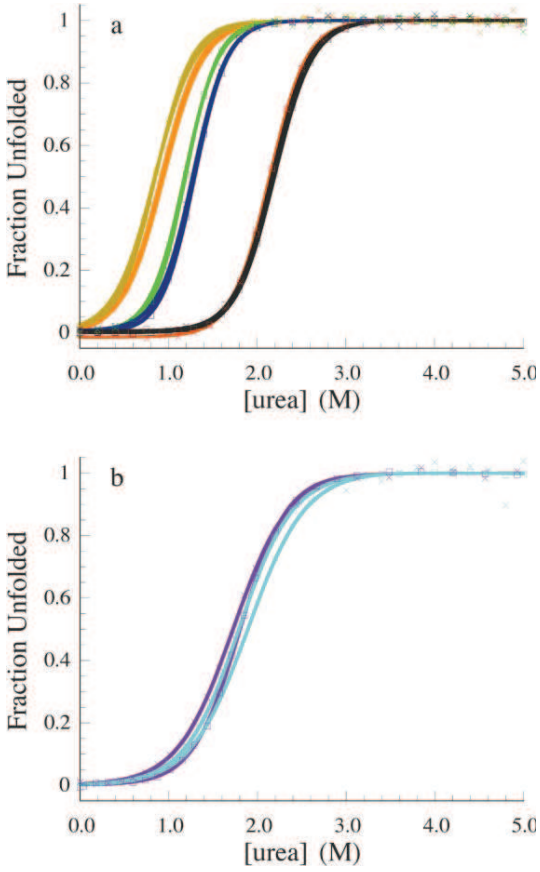

Figure 2.8: Comparison of CD- and fluorescence-monitored urea unfolding transitions of glycine variants of the Drosophila Notch ankyrin domain. Urea-induced unfolding transitions monitored by CD (cross) and fluorescence (squares) spectroscopy of (a) NAG1 (red), NAG2 (orange), NAG3 (gold), NAG4 (green), NAG5 (blue), and $Nank1-7\Delta$ (black); and (b) NAG6 (purple), and NAG7 (aqua). Solid lines show the best-fitted curves to the unfolding data using a modified form of equation (8) in which the native and denatured baseline slopes were set to zero, and the native and denatured baseline intercepts were set to zero and one, respectively. Conditions: 25 mM Tris–HCl, 150 mM NaCl, 2.0–5.0 μM protein (pH 8), 20°C.

phan is in a similar environment (sequestered from solvent), and that this environment differs somewhat from that of $Nank1-7^{*}$ for which tryptophan is more sequestered [148]. So $Nank1-6^{*}$ and $Nank1-5^{*}$ seem identical in terms of tryptophan environment.

The visible portion of the N-terminal first repeat of $Nank1-7^{*}$ adopts a different conformation in each of the three polypeptides present in the asymmetric unit the crystal used to determine its 3D structure. This heterogeneity is consistent with the observation that a core mutation in the first Notch ankyrin sequence repeat has very little effect on stability, as observed in figure 2.8, because the destabilizing alanine to glycine mutation (NAG1) unfolded curve is highly similar to that of the wild type protein ($Nank1-7\Delta$) [172], confirming that the first repeat is not fully-folded, as shown in its 3D structure and by comparing unfolding curves in figure 2.8 [172].

In the case of $Nank2-7^{*}$ and $Nank3-7^{*}$, all the repeats appear well folded and stables. And the lack of the first repeat, which caps the protein $Nank1-7^{*}$, induces a lower stability for $Nank2-7^{*}$.

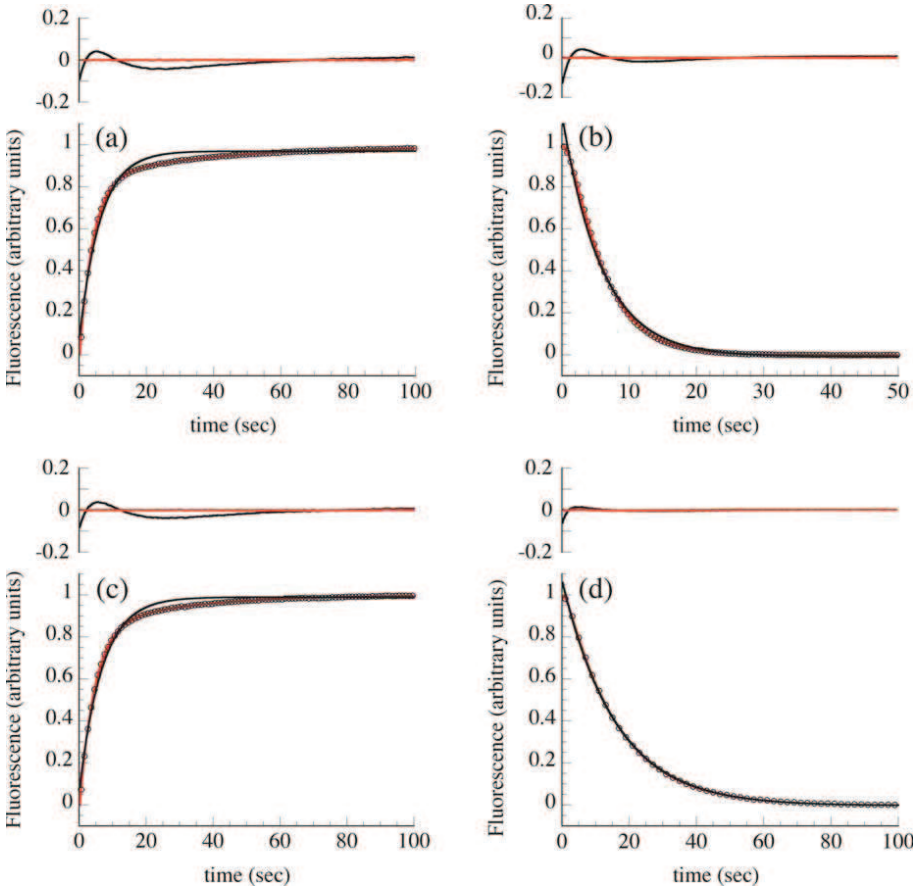

Figure 2.9: Fluorescence-detected unfolding and refolding of the Notch ankyrin domain. $Nank1-7\Delta$ ((a) and (b)) and $Nank1-7^*$ ((c) and (d)). Refolding ((a) and (c)) is initiated by rapid dilution from 3.4 M to 0.8 M urea. Unfolding ((b) and (d)) is initiated by rapid addition of urea from 0 M to 4.0 M. Continuous black and red lines show single and double exponential fits (and residuals, upper), respectively. For both refolding and unfolding, data and residuals are scaled, based on fitted amplitudes, such that the observed fluorescence change varies from zero to one.

As we can see in figure 2.9, while monitoring urea-induced unfolding for both $Nank1-7\Delta$ and $Nank1-7^{\star}$ with fluorescence spectroscopy, respectively b and d, they seem to have a simple exponential kinetics, meaning a simple two state unfolding, even if the first seconds don't fit well. But for refolding (respectively a and c), the fluorescence kinetics are not single exponentials [184], meaning that the folding process is not so simple. The slow, minor refolding phase was shown to be limited by prolyl isomerization in the denatured state (D), and the minor unfolding phase, which appears as a lag during fluorescence-detected unfolding, is consistent with an on-pathway intermediate (I). The structure in this intermediate, along with the preceding rate-limiting transition state ensemble, is restricted to the central repeats [153]. Proline residues often lead to complexity in protein folding because prolyl peptide bonds can populate both cis and trans isomers in the unfolded state [185, 186, 187]. The Notch ankyrin domain has five consensus proline residues in repeats 1, 3, 4, 6, and 7. Although the proline in repeat 1 is disordered in the crystal structure, those in repeats 3–7 are all trans in the crystal structure. Two additional proline residues are located in the seventh repeat in the loop region C-terminal to the second helix of $Nank1-7^{\star}$; this region has been omitted from $Nank1-7\Delta$. Thus the $Nank1-7^{\star}$ construct contains seven proline residues, whereas $Nank1-7\Delta$ contains five. The effects of prolyl isomerization appear to be restricted to the slowest phase of refolding [184, 173]. Given the large number of proline residues in the Notch ankyrin domain, it is possible that the intermediate may result from a prolyl isomerization reaction within the native state ensemble. The intermediate appears to be midway between N and D in folding free energy and denaturant sensitivity, but has Tryptophan fluorescence properties close to N [184]. Complete elimination of heterogeneity in $Nank1-7^{\star}$ requires substitution of all seven proline residues.

We expect the kinetics of folding/unfolding of Nank variants to be simplified under pressure, and that proline isomerization won't perturb them.

2.3.2 A well known folding mechanism

Thermodynamic studies of different ankyrin repeat proteins demonstrated that they generally possess a highly cooperative two-state folding mechanism [152, 148]. However, some experiments seem to show an intermediate extremely close to the folded form of the protein, and almost undetectable by usual methods [184, 188, 174]. Owing to their repetitive structures, linear repeat-proteins have two aspects that make them ideal subjects for simple nearest-neighbor models. First, they are constructed of repeating units, at the structure and sequence level, and second, direct contacts are limited to adjacent repeats. The folding mechanisms of these proteins were studied in details by the group of Doug Barrick, notably by using the Ising model [189].

The Ising model is a well known model that treats of the behavior of linear arrays of repetitive interacting subunits. Given this structural simplicity, the free energy of

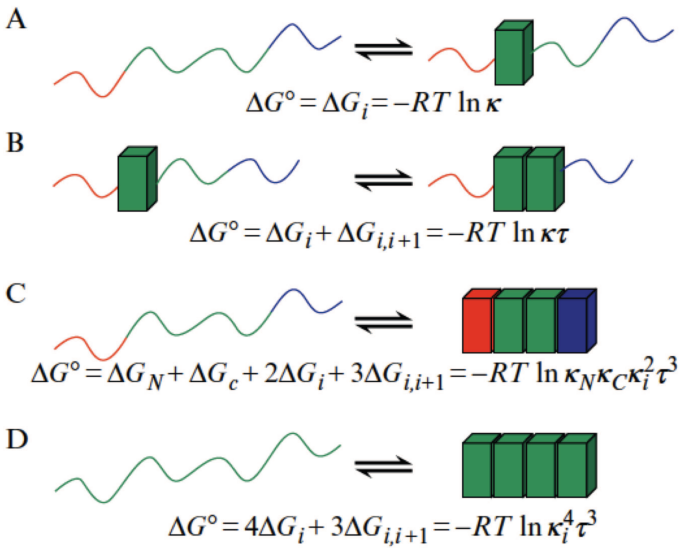

Figure 2.10: A nearest-neighbor thermodynamic description of repeat-protein stability. The first two lines (A, B) show single-repeat steps in folding (individual folded repeats are shown as blocks), whereas the last two lines (C, D) show overall folding reactions the fully denatured to the fully native state. Green repeats depict identical sequences, such as consensus repeats, and the red and blue repeats represent N- and C-terminal caps.

repeat protein folding can be expected to have two dominant contributions : the ΔG of intrinsic folding of individual units, ΔG_i (figure 2.10 A), and the ΔG of interaction at the interface between a repeat (i) and its neighbor (i+1) : $\Delta G_{i,i+1}$ (figure 2.10 B). So as seen in figure 2.10 B,C and D, when the protein folds the free energy corresponds to the sum of each of the ΔG of folding of repeats that have two interactions with neighbors, the ΔG of folding of the N and C termini as they have different ΔG of folding due to their position, plus the ΔG of interactions for each of the interfaces. For $Nank1-7^\star$ for example, $\Delta G^0 = \Delta G_N + \Delta G_C + 5\Delta G_i + 6\Delta G_{i,i+1}$. Values obtained from Ising model for Nank proteins are reported in table 2.5 where we can remark that interaction free energies are the main forces involved in the protein stability.

From the Ising model , it has been shown that there is a high degree of energetic additivity for motifs during the folding process. What happens structurally in the motifs spreads by nucleation to the other motifs and all the motifs seem to react in the same manner. All these data combine with each other to provide us an interesting environment for studying protein folding. And we were interested in determining whether a protein with this high degree of energetic additivity would exhibit additivity in its volume change of unfolding.

Because of the high energy between each motif and interfaces, there is a high cooperativity of folding presented by the ankyrin repeat protein. Hence, the modifications in the environment of the fifth repeat tryptophan should be relevant of the global folding of all the motifs, as confirmed for several ankyrin repeat constructs by comparison between Circular Dichroïsm (CD) and tryptophan fluorescence experiments for

	Consensus TPR (Kajander *et al.*, 2005)[a]	Consensus Ankyrin (Wetzel *et al.*, 2008)[b]	Consensus Ankyrin[d]
ΔG_N	n.d.	10.6 ± 0.6[c]	5.2 ± 0.1
		9.2 ± 1.1	5.4 ± 0.1
ΔG_R	2.30 ± 0.04	3.3 ± 0.2	4.4 ± 0.1
	2.26 ± 0.07	1.9 ± 1.2	4.4 ± 0.1
ΔG_C	n.d.	10.6 ± 0.6[c]	6.8 ± 0.1
		9.3 ± 1.9	6.8 ± 0.1
$\Delta G_{i,i+1}$	-4.52 ± 0.04	-14.2 ± 0.7	-11.2 ± 0.2
	-4.13 ± 0.10	-11.8 ± 1.5	-11.2 ± 0.2
m_i	0.57 ± 0.01	1.1 ± 0.1	0.75 ± 0.01
	0.57 ± 0.01	1.1 ± 0.2	0.75 ± 0.02
m_{cap}	n.d.	0.83 ± 0.04	n.d.
		0.65 ± 0.09	

Notes: Energies are in kcal·mol^{-1}; *m*-values are in kcal·mol^{-1}·M^{-1}. *n.d.*, not determined in the model used.

[a] Parameters for CTPR folding are for single helices. Parameters have been converted from the original formulation to ΔG_i and $\Delta G_{i,i+1}$, to allow comparison with other data. Errors were propagated assuming the published *H* and *J* values to be uncorrelated.

[a,b] The top line for each parameter gives estimated parameter values and uncertainties given by the authors. To facilitate comparison, the bottom line (*italics*) gives parameter values based on our fits, with uncertainties determined by bootstrap analysis as described above.

[c] For the consensus ankyrin repeats of Wetzel *et al.*, ΔG_N and ΔG_C are assumed identical, and are fitted as a single parameter.

[d] For the consensus ankyrin repeats from our laboratory, parameters and errors in the top line come from resampling guanidine titrations as described; errors in the bottom line (*italics*) come from bootstrap analysis as described.

Table 2.5: Free energies calculated with Ising model.

denaturant induced unfolding [152, 172] (figure 2.4 and table 2.2 and 2.4), if we make the assumption that folding is here a two-state mechanism, which should be well adapted.

Based on this Ising model, it is possible to make it more complex and use experimental ΔG values from equilibrium denaturation studies of each variants in order to determine the free energy of folding of each individual repeat and the interface interaction energies, as it was done in [176]. Using variants with different number of repeats and complex Ising model, they constructed a free energy landscape, shown in figure 2.11.

The application of energy landscape theory to the protein folding problem provides a step toward depicting the complexity of the protein folding process by showing the interplay between structural coordinates and representing the degree of conformational heterogeneity (and thus entropy) associated with various stages of folding [176]. The free energies can be represented as a surface or free energy landscape, in which different partly folded conformations are arranged by the number of folded repeats on one axis, and the localization of folded structure (toward the N terminus or C terminus) on the other axis (figure 2.11 A, gray-shaded surface). The vertical displacement and color of each partly folded conformation, each of which is depicted as a horizontal square, represents the folding free energy relative to the fully unfolded ensemble (at zero free energy). This experimentally determined energy landscape differs from most depictions of folding funnels in that a substantial early energy barrier separates the denatured ensemble from the rest of our landscape (figure 2.11 B). Atop this barrier, i.e., at the tiers containing one and two folded repeats, there are significant

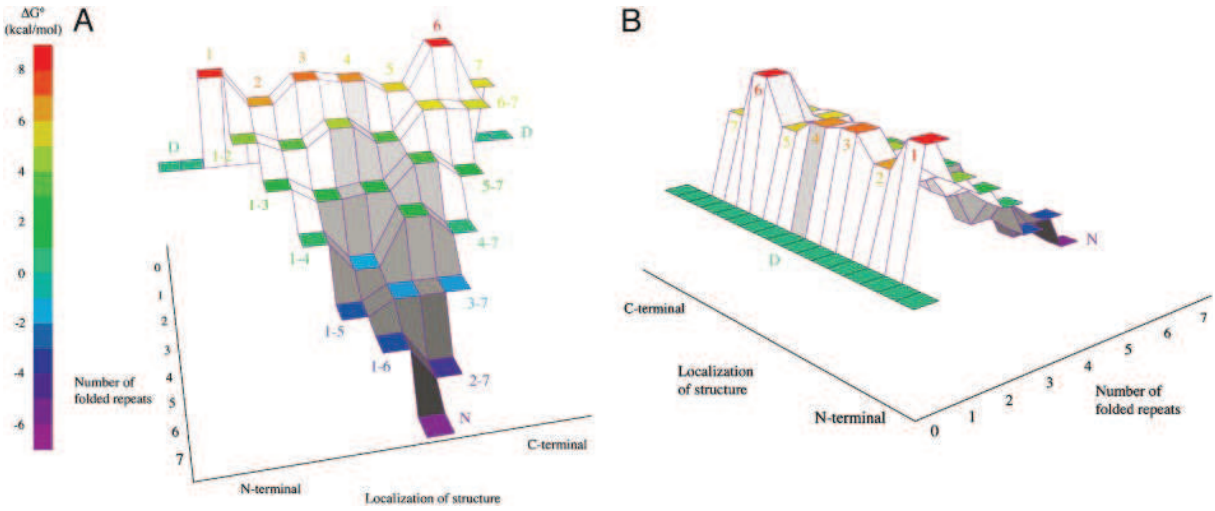

Figure 2.11: Equilibrium energy landscape for the Notch ankyrin domain. Free energies were calculated from the linear heterogeneous model with experimentally determined coefficients. Free energies of each partly folded conformation are represented by the height along the vertical axis and by color. The denatured ensemble is presented as a flat tier (aqua) at a free energy of 0. Shown are the views from the native side (A) and the opposite side (B) showing the denatured ensemble and the early barrier. The diagonal surfaces connecting neighboring conformations are meant to aid the eye but do not represent the energetics of partial folding of single repeats. Diagonal surfaces connecting partly folded conformations that contain a folded fourth repeat are represented with gray shading; those connecting conformations lacking a folded fourth repeat are unshaded. This figure was prepared with MATHEMATICA (Wolfram Research, Champaign, IL). Figure taken from [176].

differences in stability between different conformations. In particular, the two single-repeat structures composed of folded repeats 2 and 5 have low energies, especially compared with structures with single-repeat structures involving repeats 1 and 6, suggesting that these low-energy points may initiate preferred routes for folding.[176] For the repeat-4-centered landscape, the low-energy channel that connects the denatured state ensemble to the native conformation involves the sequence 4-45-345-2345-12345-1234563 native. The team of Doug Barrick showed that for the mutants, the folding is generally initiated in the central repeats, and that the TSE is folded-like only in the central motifs, even if the structure of the TSE can sometimes shift to the extremities.

Our model protein, the Nank 1-7* protein, and all its ensemble of already existing variants and mutants have been studied in detail by the team of Doug Barrick at the John Hopkins University, Baltimore. We have a wealth of information concerning their behavior at 1 atmosphere. The Ankyrin repeat proteins appears to be a model of choice for the study of the additivity, the importance of sequence and size in determining the value of ΔV^u_{eq}. Since these studies are the beginning of a complete physical characterization of the ΔV^u_{eq}, we will use it to study the temperature-pressure effects on protein stability.

Part II

Materials and methods

2.4 Production and purification of Ankyrin proteins and mutants

BL21(DE3) bacterias were transformed by the heat-shock method. Plasmids pET15b (Novagen, Madison, WI) were used to contain genes of various mutant proteins. The protocol is detailed as follows : 1 μL of DNA is aded to 200 μL of competent cells. These cells are conserved on ice during 20 minutes, then placed at 42°C during 90 seconds. The heat shock occurs by placing the cells again on ice during 3 minutes. Cells are then grown at 37°C in 0.1 $mg.mL^{-1}$ Ampicillin Luria-Bertani (LB) medium, under shaking. The induction of production by the lactose operon is provided by adjunction of IPTG (0.2 mM final concentration).

After 4 hours of induction, cells were centrifuged at 8000 rpm during 20 minutes, and bacterial pellets were resuspended in Tris 50 mM NaCl 300 mM pH 7.5. After lysis with French Press (Thermo Electron Corporation, Waltham, MA) and centrifugation at 18000 rpm, fractions of interest were taken apart (Nank 1-7*, 2-7*, NAG5 are in the supernatant, Nank 3-7*, 4-7*, 1-6*, 1-5*, 2-5* are in the pellet).

For soluble proteins in the supernatant, the solution was filtered at 0.45 μm before purification. In the case of insoluble proteins into the pellet, pellets were resuspended in 50 mM Tris 300 mM NaCl 6M Urea pH 7.5 until complete solubilization, and then filtered at 0.45 μ m before purification.

All the proteins have a His 6 tag, which allows us to use Nickel affinity chromatography, with a Hitrap 5mL column (GE Healthcare), eluted with a concentration gradient of Imidazole. For the proteins in the pellets, this step was done in 6M urea buffer. Fractions of interest are collected and pooled together. An extra dialysis into Tris 50 mM NaCl 300 mM pH 7.5 is needed for proteins in urea.

Fractions are then purified by size exclusion chromatography on HiLoad S75 26/60 or HiLoad S75 16/60 (GE Healthcare), depending on the average protein concentration, against a 50 mM Tris 300 mM NaCl pH 7.5 buffer.

Fractions of interest are pooled after analysis by SDS-Page gels, and then concentrated. The final concentration, around 10 $mg.mL^{-1}$, is checked by 280 nm absorbance spectroscopy, using a NanoDrop 3300 (Thermo Scientific, Wilmington, DE), by assuming a molar extinction coefficient of 3550 $L.mol^{-1}.cm^{-1}$.

In the unique case of Nank 3-7*, as the construction is a MBP fusion, we used TEV protease to cleave the MBP part of the protein, between the Nickel affinity and size-exclusion chromatography.

Proteins are then aliquoted and stored at -80°C.

2.5 High-pressure Fluorescence Spectroscopy

The high-pressure fluorescence spectroscopy measurements were done in a high-pressure cell coupled with a fluorescence detection system (ISS, Champaign, IL, USA). The excitation light is provided by a xenon lamp at 290 nm in order to excite the intrinsic tryptophan, and delivered to the sample by an optic fiber. The emission light is followed from 320 to 450 nm. Temperature of the sample and of the high pressure cell is maintained constant by a external thermostat.

We follow the evolution of the maximum of the tryptophan fluorescence emission peak at different pressures, to probe the immediate environment of the tryptophan. Indeed, it is known that when the tryptophan is totally buried within the structure, and so inaccessible to the solvent, it's maximum of intensity is around 330 nm, and when it is completely exposed to solvent, it's maximum is around 355 nm [190, 191, 192].

These data are effective in an ideal case, but we have to keep in mind that depending on the different amino acids that are in the close neighborhood of the tryptophan, it can be seen a shift of this maximum to the blue wavelengths, that can highly change the obtained results [191, 193, 194].

The visible change in the tryptophan emission fluorescence spectrum is seen at two different levels. First, the tryptophan, changing more and more from a position where it is buried in the structure to a position where it is exposed to solvent, during the unfolding process, shows a decrease in signal intensity. Secondly, this decrease in intensity (figure 2.13 A) is correlated with the shift of the maximum emission peak, as shown in figure 3.2.

We can represent the unfolding curves depending on pressure as a plot of the area under the curve or of the mass center of the spectra, depending on the pressure applied to samples. The area under the curve presents two problems : the errors are large depending on the protein concentration, and the maximum of the emission peak is sometimes hard to define precisely by a gaussian fitting, leading to larger errors.

In order to avoid these problems and to have more coherent results, we preferred to use the mass center of each spectrum to treat the raw datas.

The experiments were performed at different pressures, for different temperature and urea concentrations. For each tryptophan pressure-dependent emission spectrum, the center of mass of the spectrum was calculated with the ISS Vinci software, using :

$$< \lambda >_j = \sum_j F_j \lambda_j / \sum_j F_j \tag{2.1}$$

where j = 320, 321, 322, ... , 444, 450 nm.

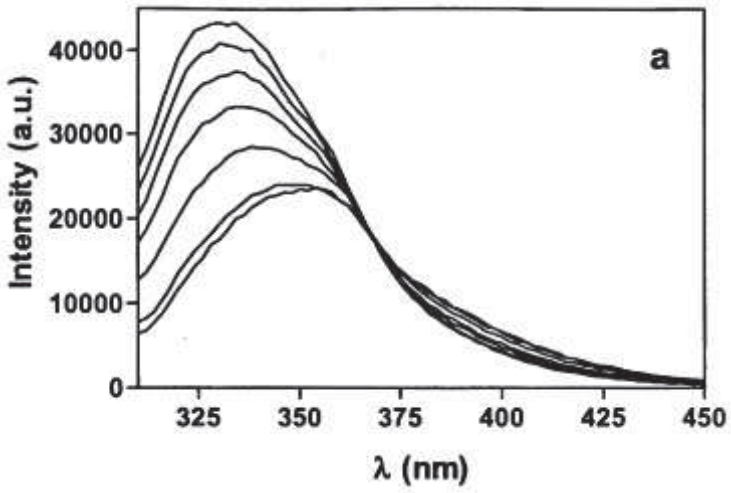

Figure 2.12: Evolution of tryptophan emission spectra with pressure. Each spectrum was obtained under 290 nm excitation light, emission was recorded from 310 to 450 nm. From atmospheric pressure (the higher spectrum) to 3000 bars (spectrum at the bottom), protein unfolding is continuously observed, here for the protein P13MTCP1 [125]. We can observe the decrease of intensity and the shift of the maximum of each peak, depending on the unfolding of the protein.

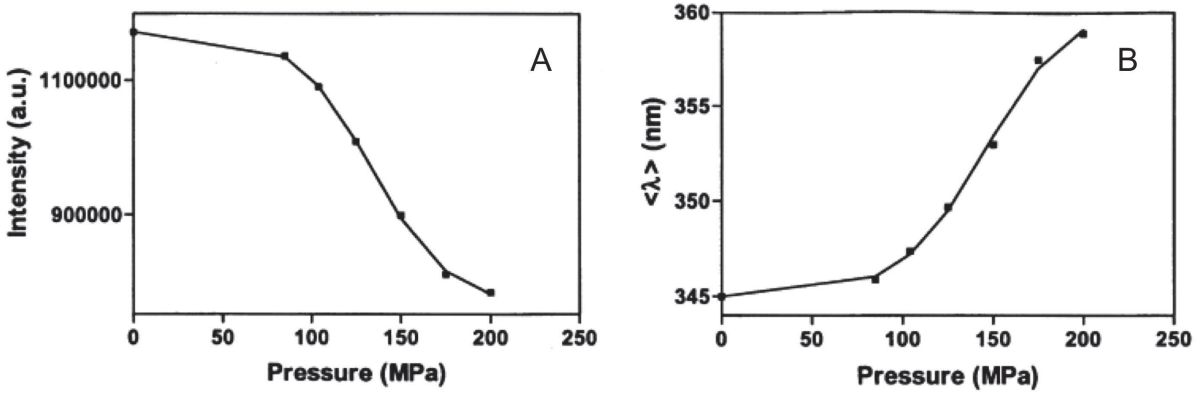

Figure 2.13: Representation of protein P13MTCP1 unfolding curves depending on the area under the curve (intensity) or the mass center of the spectra. A : Tryptohan emisson intensity decrease with pressure induced unfolding. B : Change in mass center of the spectra during pressure-induced unfolding

For the experiments of pressure-induced unfolding, the ΔG_u is assumed to evolve linearily with the pressure p :

$$\Delta G_{ui} = \Delta G_u^0 + p_i.\Delta V_u^0 \tag{2.2}$$

where :

$$\Delta G_{ui} = -RT.lnK_{ui} \tag{2.3}$$

and :

$$K_{ui} = \frac{(< \lambda >_f - < \lambda >_i)}{(< \lambda >_i - < \lambda >_u)} \tag{2.4}$$

where $< \lambda >_f$, $< \lambda >_u$) and $< \lambda >_i$ are respectively the center of spectral mass of folded, unfolded state and at the mixed-population state at pressure i. For each measurement, these calculations were made by the BioEQS [195, 196] program by taking into account the quantum yield [98] of the protein, which is around 0.5 in our case. BioEQS is a highly used and accurate free program for these calculations [91, 197, 198, 199, 200, 201, 107].

Between each tryptophan emission spectrum at equilibrium, we performed pressure jumps to reach the next equilibrium pressure. Those jumps are less than 150 bars, in order to stay in the limits of the Eyring transition state theory, regarding the ln(τ), described in equation 2.6. The tryptophan peak will shift from $\approx$ 333 nm to $\approx$ 355 nm during the complete unfolding. So the average $\approx$ 340 nm wavelength is the best one to be able to completely see the intensity decay. So during these jumps, we can follow in real time the relaxation of the proteins, by exciting at 290 nm, and following the intensity decay at 340 nm. The pressure jumps are monitored until complete equilibrium.

The relaxation time of the protein (τ) is determined by these pressure jumps kinetics by :

$$I(t) = I_0.e^{-t/\tau} + C \tag{2.5}$$

where I(t) is the maximum intensity of the fluorescence peak at time t, I_0 is the intensity at time 0, and C is the asymptotic value at each final pressure of a pressure jump.

In agree with the Eyring transition state theory, the logarithm of τ can be expressed like :

$$ln(\tau)_{(p)} = ln(\frac{1}{k_{f(p)} + k_{u(p)}}) \tag{2.6}$$

where $k_{f(p)}$ and $k_{u(p)}$ are respectively the rate constant of folding and unfolding, at pressure p. Also, we know that :

$$K_{ui} = k_{f(p)}/k_{u(p)} \tag{2.7}$$

where K_{ui} is the apparent equilibrium constant obtained from equation 2.3. Indeed, the variation of folding and unfolding rate constants are expressed relative to the rate constants depending on pressure : k_f^0 and k_u^0, as described above :

$$k_{f(p)} = k_f^0.e^{\frac{-p.\Delta V_f^\#}{RT}} \tag{2.8}$$

$$k_{u(p)} = k_u^0.e^{\frac{-p.\Delta V_u^\#}{RT}} \tag{2.9}$$

where $\Delta V_f^\#$ and $\Delta V_u^\#$ are respectively ΔV of activation of folding and unfolding. Those ΔV of activation define the folding volume difference ΔV_{eq}^f between the unfolded and the unfolded state of the protein :

$$\Delta V_{eq}^f = \Delta V_f^\# - \Delta V_u^\# \tag{2.10}$$

After treatment of the equilibrium pressure-induced unfolding curves, the analysis of the plots of $ln(\tau)$ as a function of the final pressure of pressure jumps is performed using equation 2.6, substituting $k_f(p)$ and $k_u(p)$ by equations 2.8-2.9. During these calculations, $\Delta V_f^\#$ and $\Delta V_u^\#$ are constrained using equation 2.10.

2.6 Small Angle X-ray Scattering (SAXS) under pressure

Small Angle X-ray Scattering (SAXS) is a powerful tool for investigating tertiary structures of proteins. By using this technique, biological macromolecules can be studied in their native conditions in aqueous solution [58, 202].

In the case of monodisperse and highly diluted proteins, the SAXS signal I(q) is the average of the squared modulus of the fourier transform of the electronic density of proteins, $\rho(\overrightarrow{r})$.

$$I(q) = n.P(q) = n.\langle[\int_{Vol} \rho(\overrightarrow{r}).e^{i\overrightarrow{q}\,\overrightarrow{r}}\,d^3r]^2\rangle_\Omega \tag{2.11}$$

where n denotes the protein's number density, P(q) the protein's form factor and $< . > \Omega$ the spherical average. The wavevector transfer $q = (4\pi/\lambda)\sin\theta$ is a function of the scattering angle 2θ and the incident wavelength λ. For small q values, the scattered intensity can be expressed by the Guinier approximation :

$$I(q) = I(0).e^{-R_g^2.q^2/3} \tag{2.12}$$

The radius of gyration R_g, describing the characteristic size of a scattering object, is given as :

$$R_g^2 = \int_{Vol} \rho(\overrightarrow{r}).r^2 d^3r / \int_{Vol} \rho(\overrightarrow{r}) d^3r = \int_0^{D_{max}} \rho(\overrightarrow{r}) dr \tag{2.13}$$

where D_{max} is the maximum dimension of the protein and $p(\overrightarrow{r})$ is the so-called pair-distance distribution function, characterizing the shape and size of the particle.

The Guinier approximation is valid for q-values up to q_{max} less than or equal to 1.3/Rg. For the SAXS experiments, protein solutions were prepared immediately before the measurements were performed.

To avoid pressure-induced changes of the pH value, 50 mM bis-Tris and 25 mM Tris buffer were used at pH 7.5. The concentrations of the solutions were 10.0 $mg.mL^{-1}$. At these conditions, the solution is dilute enough so that the single scattering approximation is valid. However, the scattering signal is still of sufficient intensity to obtain an adequate signal-to-noise ratio.

The SAXS experiments were performed at beamline BL9 of DELTA (TU Dortmund, Germany). High pressure SAXS up to 3 kbar was carried out in a special sample cell employing two flat diamond windows. After azimuthally averaging of the SAXS spectra, the scattering signals were normalized and corrected for solvent and background scattering by subtracting the scattering curve of the corresponding pure buffer solution. In order to obtain the radii of gyration, the low q-range showing a linear curve shape in the q^2 vs. log(I(q)) representation was fitted according to Guinier's approximation. The pair-distance distribution functions were calculated by fitting the scattering data using the program GNOM. To reveal if the proteins are folded or denatured, SAXS curves are shown in the so-called Kratky plot, i.e., q vs. q^2 I(q).

Vibration	*Diagnostic specification*	*Wavenumber (cm-1)*
Amide I band	Secondary structure	1610-1690
Amide II/II' bands	H/D exchange, protein compactness	1515-1560
Tyrosine ring	Tertiary structure compactness	1515
Antisymmetrical stretching of -COO- groups	Type of chelation of Ca2+ in Ca-binding	1560-1590
C=0 stretching in CO- bound Cyt. P-450$_{cam}$	Heme pocket environment	1900-1980

Table 2.6: Vibrational bands of diagnostic value for high-pressure FTIR studies

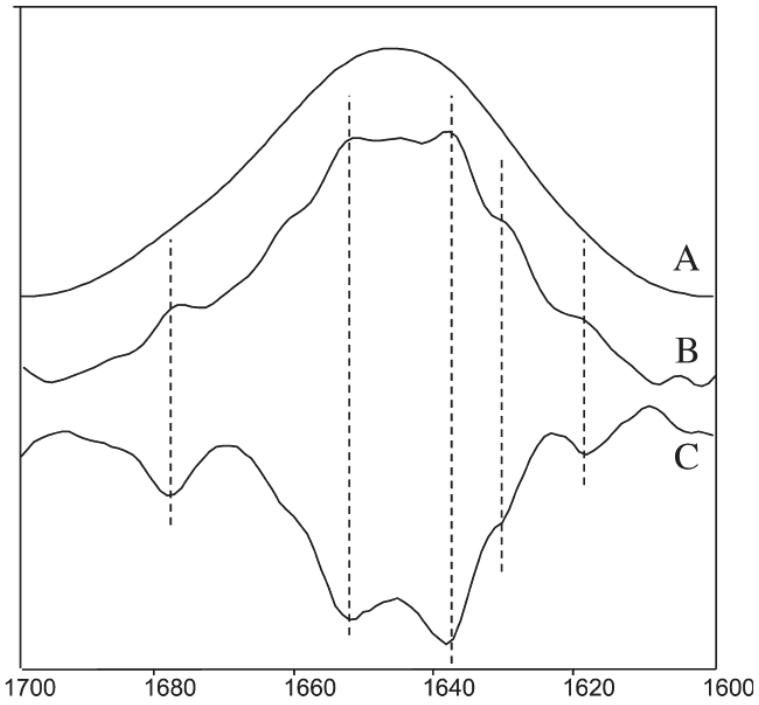

Figure 2.14: Narrowing of the amide I' band of α -lactalbumin. A : the original absorption spectrum, B : spectrum after self-deconvolution, C : the second derivative spectrum. This picture comes from [203]

2.7 Fourier Transform Infra Red Spectroscopy (FTIR) under pressure

The variety of chemical groups that are present in the amino acids in proteins induces quantity of vibrational bands, which can be sensitive to structural changes, for example those induced by pressure. Thus, the FTIR technique [203] can provide an overview of the structural modifications under pressure. Those vibrational bands and their respective wavelengths are shown in table 2.6.

In the major part, only the vibration of the amide I band is used under pressure.

By adding pressure to the system, the effect of compression on the bonds is characterized by a shift to the blue wavelengths of the infrared bands; the systems maintained by hydrogen bonds, like proteins, show a red shift. As the vibrational energy highly

Structures	*Deuterated form (cm-1)*	*H/D-caused red-shift (cm-1)*
α-Helix	$\approx$ 1654	$\approx$ 5
β-Sheet	1615 and 1685	
Random coil	$\approx$ 1645	$\approx$10
3_{10}-Helix	$\approx$ 1639	$\approx$ 15
Turns	$\approx$ 1682-1662	5

Table 2.7: Typical assignments of the elements of protein secondary structure after complete H/D-exchange

depends on the presence and geometry of hydrogen bonds, the frequence of the band of the amide I is affected by the secondary structure, where are located the amide bonds [203].

Thus, by deconvoluting the absorbance spectra, we obtain peaks specific of a wavelength, as shown in figure 2.14. By auto-deconvoluting with the Fourier method, we can recreate the derivative of each spectrum, which helps us increasing the resolution. The peaks appearing by this method, reported in table 2.7, will evolve by shifting, depending on the pressure we add.

As urea has bonds that are FTIR active, we preferred to use ^{13}C labeled urea, in order to make it's pics disappear from the spectra [204]. Samples were prepared at 2% weight-volume of the protein, in D_20, in 50 mM Tris 150 mM NaCl pH8, with labeled urea.

Pressure was applied to the samples by a diamond anvil. For each pressure step, 256 scans were summed. The analysis and deconvolution of the spectra was performed using the program GRAMS/AI 8.0. For comparison reasons, each pics were normalized.

2.8 Densitometry

To perform the densitometry measurment, we used a DMA 5000 density meter (Anton Paar, Courtaboeuf, France).

For performing highly accurate temperature-dependent density measurements, the DMA 5000 contains an oscillating tube built of glass. The temperature range for operation of the apparatus is from 10 to 70 °C. Temperature was controlled with a built in precision temperature sensor with an accuracy of 0.01°C. About 1.2 cm3 of the sample solution were needed to fill the measuring cell. The accuracy of density measurements as a function of temperature is very high, about $5.10^{-5} g.cm^3$. The density of protein solutions was measured at ambient pressure from 10 to 95 °C.

The apparent specific volume Vp was calculated from the protein solution (d) and

the solvent (d_0) densities using:

$$Vp = \frac{m_0}{m_{prot}}\{\frac{1}{d} - \frac{1}{d_0}\} + \frac{1}{d} \tag{2.14}$$

where m_{prot} is the mass of the solute (protein) and m_0 is the mass of the solvent. The absolute error of Vp for measurements of different sample preparations is about $\approx 1.5\%$.

2.9 Differential Scanning Calorimetry (DSC) and Pressure Perturbation Calorimetry (PPC)

2.9.1 DSC

The DSC is a powerful tool to understand and analyze the evolution of the protein state with temperature, that help discover thermodynamical values, like ΔCp. The cells of the DSC (GE Healthcare), filled with the protein solution in one cell and reference in the second cell, are heated at the same time. The difference of heat absorbed or produced by the protein solution, compared to the reference cell, is recorded and compensated by the DSC. In this way, we can obtain the δQ depending on the temperature and on the heating gradient : $\delta T/\delta t$, where T is the temperature and t the time. It's known that :

$$Cp_{pr} = (\frac{\delta Q}{\delta T})_p = (\frac{\delta H}{\delta T})_p \tag{2.15}$$

By assuming our system is a simple two-state system, we assume that the Cp_{pr} will evolve depending on temperature, as the sum of the Cp of the two states : Cp_N and Cp_U, respectively for the heat capacity of the native and unfolded state. This leads to a baseline composed by the sum of the heat capacity of each population as a function of temperature :

$$Cp_{pr} = F_N(T)Cp_N + F_U(T)Cp_U \tag{2.16}$$

where F_N and F_U are respectively the population of protein in the native and unfolded state. As we have shown before in the High-pressure Fluorescence section, p 42, we know the relation between $\Delta G(T)$ and K_{eq} from equation 2.3, and we know too the link between the entropy, the enthalpy and the gibbs free energy :

$$\Delta G(T) = \Delta H - T\Delta S(T) \tag{2.17}$$

We can define the unfolded population of proteins $F_U(T)$ as :

$$F_U(T) = \left(\frac{e^{\frac{-\Delta G(T)}{RT}}}{1 + e^{\frac{-\Delta G(T)}{RT}}}\right) = \left(\frac{e^{\frac{-\Delta H(T)}{RT}} e^{\frac{\Delta S(T)}{R}}}{1 + e^{\frac{-\Delta H(T)}{RT}} e^{\frac{\Delta S(T)}{R}}}\right) \tag{2.18}$$

So the population of folded proteins can be expressed as :

$$F_N(T) = 1 - F_U(T) \tag{2.19}$$

At the center of the transition, we can observe in the data an excess of Cp, which comes from the denaturation of the protein. This excess can be described as :

$$Cp_{excess} = \frac{\delta}{\delta T} F_U(T) \tag{2.20}$$

We can then use the equation 2.18 to derive it :

$$\frac{\delta}{\delta T}(F_U(T)) = \left(\frac{e^{\frac{-\Delta H(T)}{RT}} e^{\frac{\Delta S(T)}{R}} \frac{\Delta H(T)}{RT^2}}{(1 + e^{\frac{-\Delta H(T)}{RT}} e^{\frac{\Delta S(T)}{R}})^2}\right) = \frac{K_{eq}(T)}{(1 + K_{eq}(T))^2} \frac{\Delta H(T)^2}{RT^2} \tag{2.21}$$

So we can finally obtain the equation of the evolution of the Cp of the total protein depending on temperature, with :

$$Cp_{exp}^{total}(T) = F_N(T) Cp_N(T) + F_U(T) Cp_U(T) + \frac{K_{eq}(T)}{(1 + K_{eq}(T))^2} \frac{\Delta H(T)^2}{RT^2} \tag{2.22}$$

which reflects the whole temperature unfolding process on the explored 164-6 formappl temperature scale. As we can see, the difference between Cp_U and Cp_N defines the ΔCp of the protein, and the equations will give us the enthalpy of this protein.

2.9.2 PPC

The PPC is an add-on module for the DSC, owning the ability of pushing pressure quickly in the cells and to release it with the same speed. After equilibration of the DSC cells, a pressure change of 5 bars ($\approx$ 60 psi) is made. This change in pressure will lead to a difference of temperature (ΔT) between the two cells, which is recorded. Then an energy ΔQ is applied on one of the cells, to re-equilibrate the system temperature. Once the equilibrium is reached, pressure is released, leading to an inverse temperature change between the cells. the equilibrium is re-established by the PPC machine. Then, a new temperature of equilibrium of the two cells is reached, and this process repeats.

At the end of the experiment, we obtain a ΔQ value for each temperature, which are directly proportional to the coefficient of expansivity of the solution : α.

In order to obtain only the α of the protein, the solvents and salts have to be taken into account. So we are obliged to do water-water, water-buffer, buffer-buffer, buffer-protein solution experiments, each time for identical temperatures and parameters. We then substract all the previous experiments from the buffer-protein solution experiment.

The PPC measures very small changes in the absorbed or produced heat when the pressure is provided and released. We know from the second thermodynamic law that for reversible changes :

$$\delta Q = T\delta S \tag{2.23}$$

where δS is the entropy change at temperature T, for a reversible temperature change δQ.

Adding a pressure change at constant temperature changes this equation into :

$$(\frac{\delta Q}{\delta P})_T = (\frac{T\delta S}{\delta P})_T \tag{2.24}$$

The Maxwell relationship :

$$(\frac{\delta S}{\delta P})_T = -(\frac{\delta V}{\delta T})_P \tag{2.25}$$

can be used with 2.23 to derive this new equation :

$$(\frac{\delta Q}{\delta P})_T = -T(\frac{\delta V}{\delta T})_P = -TV\alpha \tag{2.26}$$

where α is the expansivity, linked with the heat as defined in equation 2.23. By integrating equation 2.25, one can find :

$$Q_{reference} = -TV_{cell}\alpha_s\Delta P \tag{2.27}$$

As our solution contains m_{pr} grams of protein, and m_s grams of solvent, the total volume of the cell, V_{cell} can be seen as :

$$(\frac{\delta Q_{sample}}{\delta P})_T = -T(m_{pr}V_{pr}\alpha_{pr} + m_sV_s\alpha_s) \tag{2.28}$$

By integrating 2.28 for a small pressure jump of ΔP, we can obtain the expansivity of the protein at temperature T in terms of expansivity of the buffer α_s and heat change,

as described in equation 2.29 :

$$\Delta Q_{pr/s} = Q_{pr} - Q_s \tag{2.29}$$

$$\alpha_{pr} = \alpha_s - \frac{\Delta Q_{pr/s}}{T\Delta P m_{pr} V_{pr}} \tag{2.30}$$

$$\alpha_s = \alpha_{H_2O} - \frac{\Delta Q_{s/H_2O}}{T\Delta P V_{cell}} \tag{2.31}$$

So, we are finally able to express the expansivity of the protein in a more simple way :

$$\alpha_{pr} = \alpha_{H_2O} - \frac{\Delta Q_{s/H_2O}}{T\Delta V_{cell}} - \frac{\Delta Q_{pr/s}}{T\Delta P m_{pr} V_{pr}} \tag{2.32}$$

In this equation, all the parameters are obtained from the PPC, but not V_{pr}, which can be obtained by densitometry (see 2.8 . The concepts of PPC and DSC (see Material and Methods p. 2.9.1) are close, and the relationship between $\alpha(T)$ and $\Delta V_{pr}/V_{pr}$ is of the same order than the relation between Cp(T) and ΔH. So we can try to make them linked in a global analysis.

By taking into account a two-state system, the expansivity obtained experimentally, α_{exp} is constituted of two terms : the expansivity of the folded (α_N) and unfolded (α_U) states, compared to their own population F in the solution at each temperature. Then, like in the DSC (2.16), we can express α_{exp} like that :

$$\alpha_{exp} = F_N\alpha_N(T) + F_U\alpha_U(T) \tag{2.33}$$

Using 2.3 and 2.17, we can solve $F_U(T)$, taking into account that $K_{eq} = F_U(T)/F_N(T)$ and $F_U(T) + F_N(T) = 1$:

$$F_U(T) = \left(\frac{e^{\frac{-\Delta G(T)}{RT}}}{1 + e^{\frac{-\Delta G(T)}{RT}}}\right) = \left(\frac{e^{\frac{-\Delta H(T)}{RT}} e^{\frac{\Delta S(T)}{R}}}{1 + e^{\frac{-\Delta H(T)}{RT}} e^{\frac{\Delta S(T)}{R}}}\right) \tag{2.34}$$

We know that during the unfolding, the apparent expansivity is in fact an expansivity excess in regard to the expansivities of the two states, which is relative to $\Delta V_{pr}/V_{pr}$. If we assume that ΔV_{pr} is temperature-independent, we can define α_{excess} as :

$$\alpha_{excess} = \frac{1}{V_{pr}}\frac{\delta}{\delta T}(F_U(T)\Delta V_p) = \frac{\Delta V_{pr}}{V_{pr}}\frac{\delta}{\delta T}F_U(T) \tag{2.35}$$

And it's derivative of 2.35, in agreement with 2.34, is :

$$\frac{\delta}{\delta T}F_U(T) = \left(\frac{e^{\frac{-\Delta H(T)}{RT}}e^{\frac{\Delta S(T)}{R}}\frac{\Delta H(T)}{RT^2}}{(1+e^{\frac{-\Delta H(T)}{RT}}e^{\frac{\Delta S(T)}{R}})^2}\right) = \frac{K_{eq}(T)}{(1+K_{eq}(T))^2}\frac{\Delta H(T)}{RT^2} \tag{2.36}$$

By substituting 2.34 and 2.35 :

$$\alpha_{excess}(T) = \frac{K_{eq}(T)}{(1+K_{eq}(T))^2}\frac{\Delta H(T)}{RT^2}\frac{\Delta V_{pr}}{V_{pr}} \tag{2.37}$$

Finally, the expensivity obtained experimentally by PPC can be expressed by combining of 2.33 and 2.37

$$\alpha_{exp}^{total}(T) = F_N(T)\alpha_N(T) + F_U(T)\alpha_U(T) + \frac{K_{eq}(T)}{(1+K_{eq}(T))^2}\frac{\Delta H(T)}{RT^2}\frac{\Delta V_{pr}}{V_{pr}} \tag{2.38}$$

Tris buffers are known to change their pH with temperature, of $\approx$ 0.01 pH unit per degree, so we used Phosphate buffer (Phosphate 50 mM NaCl 150 mM pH 7.5) that doesn't have any pH variation with temperature. Phosphate buffers are known to change their pH at high pressures, but as the pressure used in PPC is very small (5 bar), the change is so small that it can be ignored. The ankyrin protein has tendency to aggregate at high temperatures, after unfolding, so we included 1M urea to avoid these problems, as described in [172]. Proteins were dialyzed to this urea buffer in two steps, overnight, and 50 mL of this buffer was conserved before using it, to do the water-buffer, buffer-buffer and buffer-protein scans, to avoid pipetting and concentration errors. Such a concentration of denaturant is here sufficient to avoid aggregation and to destabilize a little the protein, but not enough to unfold it. So we still obtain the whole unfolding progression with temperature. The Tm will shift to lower temperature, because of this urea, but as we used always the same protocol, all the datasets permit comparison between each other.

2.10 Molecular Dynamics Simulations

We base our simulation protocol on [205]. Simulations were run in GROMACS 4.0.4 [206]. using the OPLS-aa force field [207] and SPC-E water model [208]. The constructs we used for simulations are based on the original 1OT8 pdb structure ($Nank1-7^{\star}$). The model used for $Nank1-7^{\star}$ contains residues 25 to 239, 25 to 112 for 1-3 construct, 80 to 177 for 3-5 construct, 146 to 239 for 5-7. NW and NWC structures were created by sequence and structure homology with the MODELER software [209]. These conformations were used as starting points for 10 ns long MD simulations at 293 K and 1 atm. Counterions were added to obtain electroneutrality. Temperature was controlled using a Nosé-Hoover thermostat with a 5-ps coupling time [210] and pressure was con-

trolled using a Parinello-Rahman barostat with a 5 ps coupling time [211]. The system compressibility was set to $4.6.10^{-5} bar^{-1}$.

Long-range electrostatic interactions were calculated using particle-mesh Ewald [212] with a grid spacing of 1.2 Angströms and cubic interpolation. Van der Waals energies were cut off at 8 Angströms. A 2 fs time step was used and the non-bonded list was updated every 10 integration steps.

Part III

Results

Chapter 3

$Nank1-7^{\star}$

The modular architecture of ankyrin repeat proteins can help us find the exact positions of the theoretical voids which lead to the ΔV^{u}_{eq} within the structure, and to discriminate between voids which ones are or are not important for the ΔV^{u}_{eq} calculation. With our systematic approach in temperature and pressure, we hope to clarify the contributions and the physical and thermodynamical parameters that are relevant for the ΔV^{u}_{eq}. Here, we report a thorough study of the pressure-induced unfolding of $Nank1-7^{\star}$ [169].

3.1 Pressure-induced equilibrium unfolding at various temperature

The pressure-induced equilibrium unfolding of $Nank1-7^{\star}$ was characterized as a function of temperature by measuring the average wavelength of the fluorescence emission of the single tryptophan residue of the protein (figure 2.3 B and C page 26), located in the 5th ankyrin repeat.

From our preliminary results on the $Nank1-7^{\star}$ protein at 20°C, it was evident that $Nank1-7^{\star}$ was too stable to unfold completely at 3000 bar, the highest pressure we could access for fluorescence studies. Denaturant was used to move the unfolding into our available pressure range (up to 3000 bar). Adding denaturant in concentrations such as the protein is destabilized but still folded or at the early beginning of the unfolding curve, allows pressure to act on it, leading to unfolding. Experiments were first done using GuHCl, an usual denaturant [138, 213, 91, 214, 215, 216, 217], but for manipulation purposes and in order to be able to compare results between our experiments and those of Barrick's group, we preferred using urea for all the following experiments [216, 172, 204, 218], as results were comparable. In the case of $Nank1-7^{\star}$, a single urea concentration was not sufficient to obtain a complete unfolding curve under pressure, as the destabilized protein was still too stable to unfold completely. Hence we performed a systematic analysis at several urea concentrations, in order to obtain data for the whole unfolding process. The pressure profiles were collected at urea concentrations ranging from 1.9 to 2.4 M, depending upon the temperature, to observe

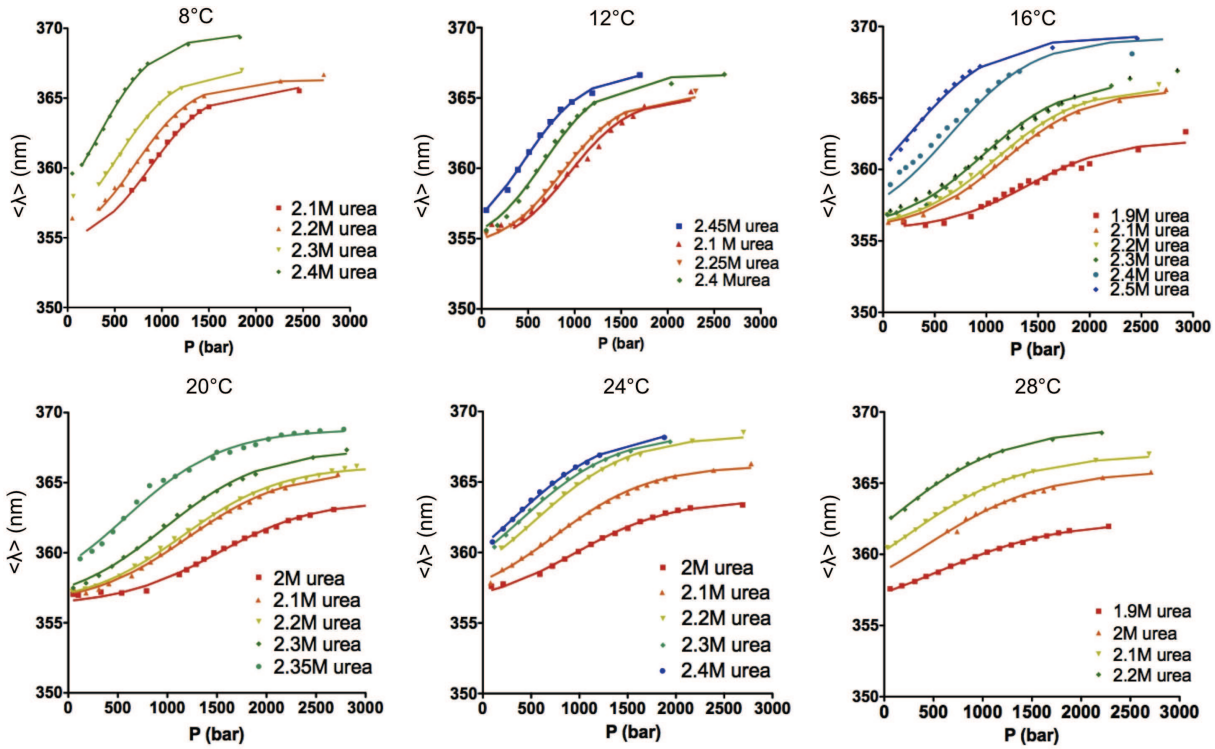

Figure 3.1: Evolution of the center of the spectral mass of tryptophan emission spectra with urea and temperature. *Nank*1 – 7* at 40 μM in 50 mM Tris , 150 mM NaCl , pH 7.5 is unfolded by use of pressure with various urea concentrations, at 8, 12, 16, 20, 24 and 28°C. In each graph, urea concentrations used are represented from the lowest (red squares) to the highest (blue circles).

unfolding over our experimentally accessible pressure range (0.1–3500 bar). We note that at the micromolar protein concentrations used for the fluorescence experiments, these transitions were entirely reversible.

Figure 3.1 represents the evolution of the mass center of the tryptophan emission spectra depending on the pressure applied to the protein solution (*Nank*1 – 7* 40 μM, Tris 50 mM, NaCl 150 mM, pH 7.5), at 8, 12, 16, 20, 24 and 28°C for several urea concentrations. The urea concentrations used to explore the complete denaturation of *Nank*1 – 7* are represented in different colors, from red for the lowest to blue for the highest one, and were defined thanks to the previous characterization work of the unfolding process by Barrick's group via denaturant-induced unfolding [152, 172]. As observed in figure 3.1, from 8 to 28°C, the protein remain folded at low pressure and low urea concentration, defining in each case a stable folded plateau. These plateaus were confirmed by fluorescence experiments at 6M GuHCl. We can note that, depending on the temperature used, the folded plateau of *Nank*1 – 7* changes from $\approx$ 355 nm at 8°C to $\approx$ 357.5 nm at 28°C, coming from a red shift of the tryptophan emission spectra. This is explained by the fact that *Nank*1 – 7* undergoes thermal fluctuations leading to a higher hydration level of its folded state upon increasing temperature. While increasing pressure, we can see that the protein unfolds through the red shift

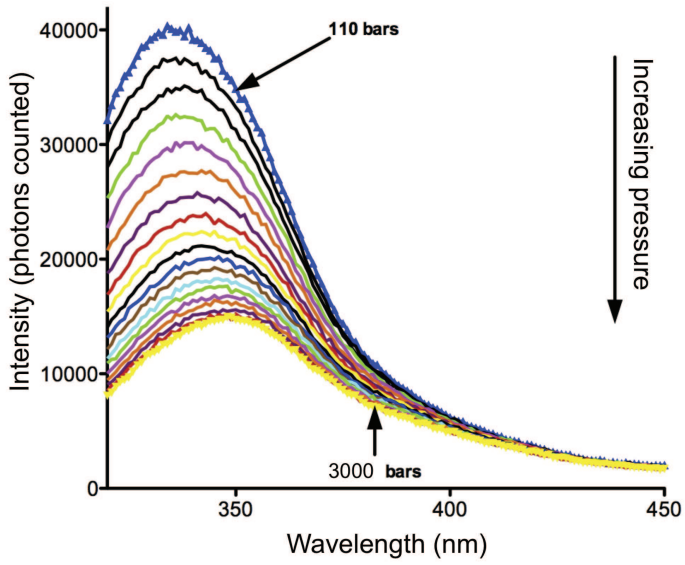

Figure 3.2: Evolution of tryptophan emission spectra with pressure. Each spectrum was obtained under 290 nm excitation light and emission was recorded from 310 to 450 nm. From atmospheric pressure (the higher spectrum, in blue) to 3000 bar (spectrum at the bottom, in yellow), the protein unfolding is continuously observed, here at 2.1M of urea, by a decrease of intensity and the shift of the maximum of the tryptophan spectrum. The consequence of the red shift is the increase of the center of the spectral mass of the tryptophan emission spectrum, as explained in materials and methods.

of the tryptophan emission spectrum as reported by the related increase in center of spectral mass (figure 3.2). At all temperatures, the application of pressure results in a sigmoidal red shift in the fluorescence emission (figure 3.1) that exhibited a distinct high-pressure plateau for the average emission wavelength. However, the magnitude of the pressure-induced red shift was strongly dependent upon the concentration of urea. The higher the urea concentratin in solution, the less pressure needed to induce unfolding and the higher the plateau. In all cases, the volume change upon unfolding was found to be independent of the urea concentration by individual curve analysis, so data were globally fit by the BioEQS program [195, 196] at each temperature, linking the volume change across all urea concentrations, for ΔV_{eq}^{f}. The urea-dependent free-energy change of unfolding (ΔG_{u}^{0}) was not linked across the datasets. The numerical values are reported from the global fits in table 3.1.

The increase in center of spectral mass of the tryptophan emission spectrum for the different plateaus with urea concentration reveals increased solvent relaxation around the excited state of the tryptophan at the high-pressure, indicative of an increased degree of solvent accessibility. At the highest urea concentrations used, at all temper-

[Urea] M	1.9		2.0		2.1		2.2		2.3		2.35		2.4		2.45		2.5	
T (°C)	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol	$\Delta G°_u$ kcal /mol	$\Delta V°_f$ kcal /mol
8					1.6	71.05	1.32	71.05	0.95	71.05			0.54	71.05				
12					1.8	69.5	1.66	69.5					1.28	69.5	0.89	69.5		
16	1.91	52.3			1.59	52.3	1.49	52.3	1.32	52.3			0.89	52.3			0.36	52.3
20			1.7	44.2	1.36	44.2	1.3	44.2	1.1	44.2	0.65	44.2						
24			0.96	42.5	1.03	42.5	0.55	42.5	0.39	42.5			0.33	42.5				
28	0.54	38.9	0.41	38.9	0.19	38.9	0.00	38.9										

Table 3.1: Thermodynamic parameters (ΔG_u^0 and ΔV_{eq}^f) for the pressure-induced unfolding of Nank1-7 as a function of temperature and urea concentration

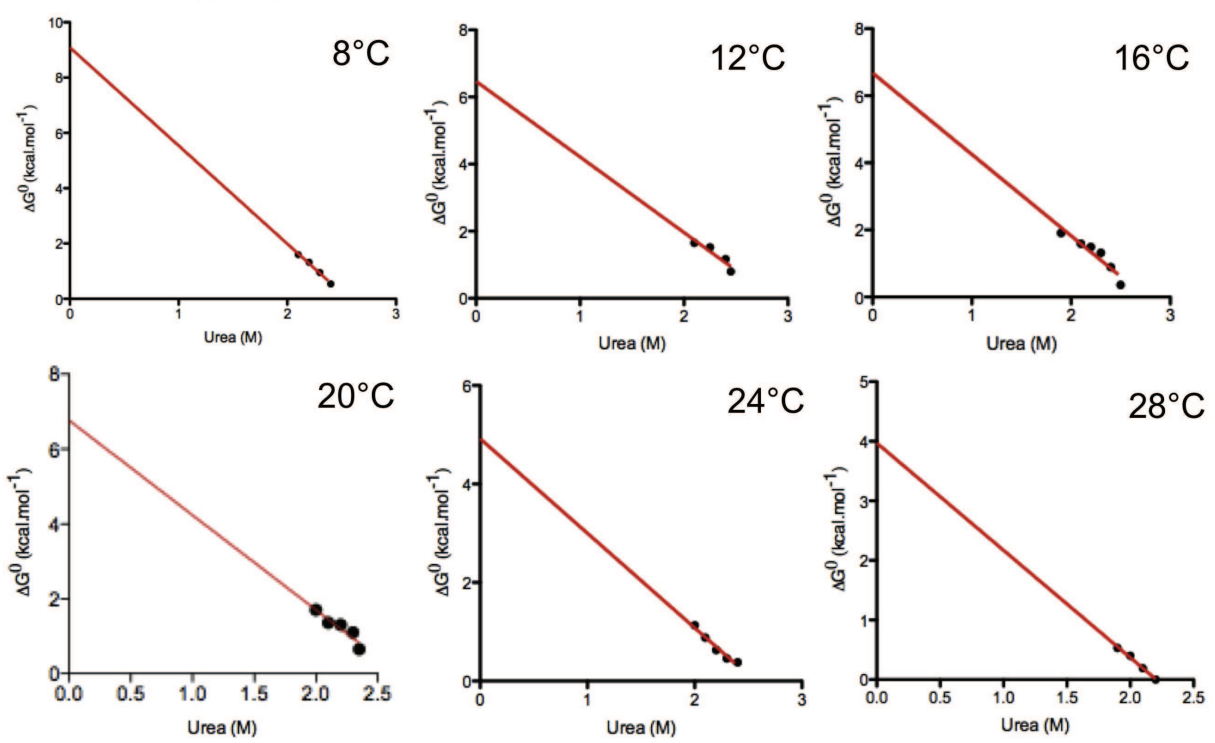

Figure 3.3: Evolution of the ΔG_{eq}^u with urea at all temperatures. All the $\Delta G_{eq,urea}^u$ were fitted by linear regression to obtain the ΔG_u^0.

atures, the final high-pressure plateau corresponds to the average emission wavelength observed for N-acetyL-tryptophanamide in water, i.e., a completely accessible tryptophan moiety. We hypothesize that the unfolding of $Nank1-7^*$ by pressure conforms to a two-state model, without any intermediate, with a different degree of tryptophan exposure in the pressure-unfolded state depending upon the urea concentration, which can come from a high plasticity of the unfolded state.

As shown in figure 3.3, the urea-dependent Gibbs free energy difference upon unfolding ($\Delta G_u^0(i)$) obtained for each urea concentration-dependent pressure-induced unfolding curve were collected and used for linear regression to obtain the Gibbs free difference upon unfolding in absence of urea (ΔG_u^0), as it is known that $\Delta G_u^0(i)$ depends linearly with denaturant concentration, as described in the following equation :

$$\Delta G_u^0(i) = \Delta G_{H_2O}^0 + m[Urea] \tag{3.1}$$

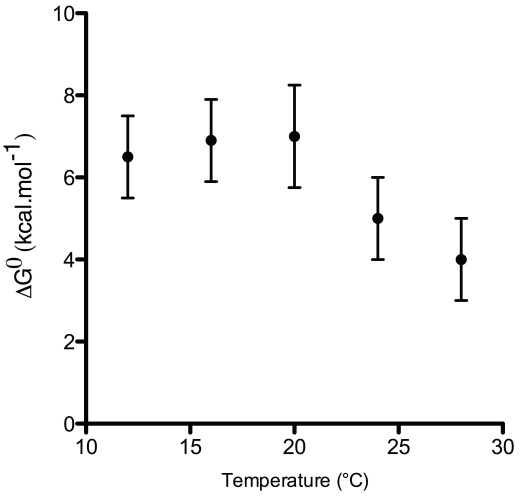

Figure 3.4: Stability diagram of $Nank1-7^{\star}$.

Although the uncertainty of these extrapolated values is large, given the small number of points (four to five urea concentrations) and the long extrapolation (≈2 to 0 M urea), the values calculated for ΔG_u^0 at 16 and 20°C (6.9 ± 1.0 and 7.0 ± 1.25 kcal/-mol, respectively) are within the error of those previously reported from urea unfolding studies at atmospheric pressure at 16 and 20°C (8.09 ± 0.07 and 7.52 ± 0.04 kcal/-mol, respectively). ΔG_u^0 at the other temperatures tested are with a larger error in comparison with known values, but within a $\pm 2kcal.mol^{-1}$ error, even if they are a little underestimated. These errors are close to the standard errors known for ΔG_{eq}^u calculations [219]. The nonlinear temperature dependence of ΔG_u^0 in the absence of urea, as shown in figure 3.4, is consistent with a significant heat capacity change upon unfolding, which is typical of proteins [218]. All our ΔG_u^0 are a little low compared to published values, but within the error [152].

From these results, we were able to define the evolution of ΔV_{eq}^f with temperature for $Nank1-7^{\star}$, as shown in figure 3.5. We can observe that there is a large decrease of the ΔV_{eq}^f from 71.05 $mL.mol^{-1}$ at 8°C to 38.9 $mL.mol^{-1}$ at 28°C. This decrease seems linear in this range of temperature, and we can imagine that if linearity is maintained above 28°C, the ΔV_{eq}^f changes its sign at ≈ 48°C, but when we tried high-pressure fluorescence experiments at 32°C, data were too noisy because of a bad signal to noise ratio, and we were not able to continue this work without making large approximations or putting the setup at risk because of the maximum attainable temperature of the pressure vessel. It appears that $Nank1-7^{\star}$ is a good model to test the ΔV_{eq}^f change of sign at high temperatures by extrapolation of data at low temperatures.

We found a strong dependence of ΔV_{eq}^f upon temperature, $\approx -1.7mL.mol^{-1}.K^{-1}$, which arises from the difference in thermal expansivity, $\Delta\alpha$, between the unfolded and folded states, as explained in introduction. As our setup wasn't able to define ΔV_{eq}^f at more than 28°C, we performed pressure-perturbation calorimetry in order to define it at the melting point, as shown in figure 3.6. As we can see, the expansivity of both states

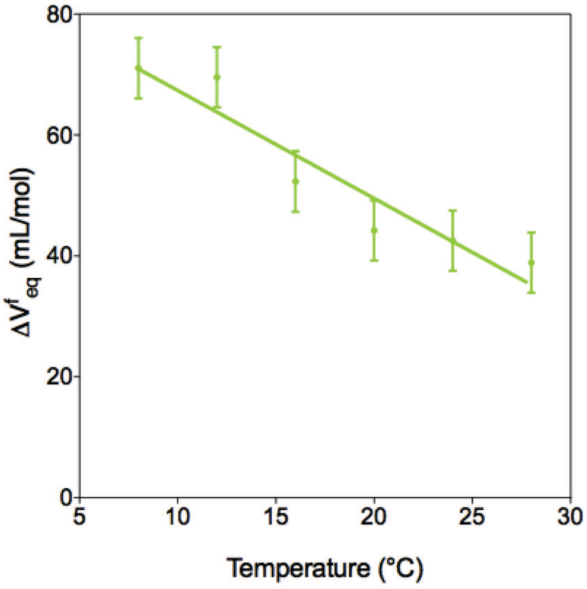

Figure 3.5: Evolution of ΔV^f_{eq} $(mL.mol^{-1})$ with temperature (°C) for $Nank1-7^*$.

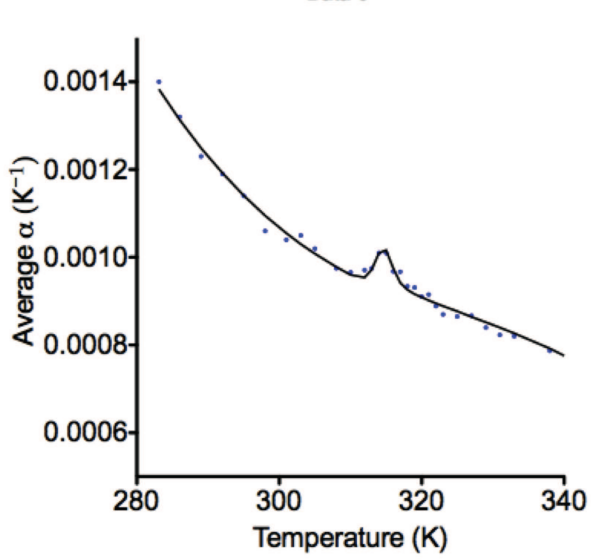

Figure 3.6: Pressure Perturbation Calorimetry experiment on $Nank1-7^*$. Experiment done at 12 mg/mL of $Nank1-7^*$, in phophate buffer 50mM, 150 mM NaCl, 1M urea.

(folded state at low temperature and unfolded state at high temperature) are between $1.4x10^{-3}$ and $0.8x10^{-3}K^{-1}$. These values are higher than usual, and we suspect some aggregation of the protein in solution, as PPC needs at least $10mg.mL^{-1}$ of protein. Because of this apparent aggregation, we can't really hypothesize anything about the expansivities of the two states, but we can be sure of the change appearing at the melting point. In order to avoid aggregation after the Tm, we performed PPC with a buffer containing 1M of urea, so the apparent Tm is less than the normal one. At the Tm, we can observe that there's a positive change in the plot, which area corresponds to the ΔV^u_{eq}. From this experiment, we learnt that at 315 K (42 °C), the ΔV^u_{eq} is + 8.2 $mL.mol^{-1}$, and so ΔV^f_{eq} is - 8.2 $mL.mol^{-1}$. This negative ΔV^f_{eq} was confirmed too by another PPC experiment in Tris buffer (data not shown). This experiment told us that ΔV^f_{eq} changes its sign at high temperature. We saw here the exact pattern known as the volume change paradox and which is simply due to $\Delta\alpha$. By using the ΔV^f_{eq} at 42 °C obtained from PPC, we can observe in figure 3.7 A that the ΔV^f_{eq} becomes null at about 40°C. Even if fluorescence experiments predict the change of sign at 48°C and PPC at 40°C, we can't have an exact value of the temperature at which ΔV^f_{eq} becomes null, but we can be sure it is between 40 and 50°C. In all the cases, we managed to prove with a high confidence the ΔV^f_{eq} change of sign of $Nank1-7^\star$.

To better understand the volumetric properties of $Nank1-7^\star$ upon increasing temperature, we performed densitometry measurements, as shown in figure 3.7 B. As we can observe, the volume of the folded state at low temperature increases with temperature until the the unfolding transition is reached. During the transition, we follow a weighted average of the volume of each state. After complete unfolding, we observe only the unfolded state volume which is higher than the one of the putative folded state, and which continues to increase while increasing temperature. Interestingly, densitometry measurement confirms that at about 32°C, the ΔV^f_{eq} is very low, explaining the poor signal-to noise ration in high-pressure fluorescence study, and confirms that after the transition, the ΔV^u_{eq} turns positive, contrary to the negative one at low temperatures. As we don't know exactly how expansivity evolves with temperature for the folded state at high temperature, we can't know exact ΔV^f_{eq} values at high temperatures after the transition, but the change of sign is evident. Here again, we can obviously see the volume change paradox, although the temperature is not well defined. This experiment confirms the hypothesis that the ΔV^f_{eq} change its sign because of the $\Delta\alpha$, as the evolution of the volume of each state (driven by the expansivity of each state) is different while increasing temperature. Now that we can be sure of it, it is clear that the expansivity of the folded, unfolded and transition states of proteins are an important part of the understanding of the effects of pressure on proteins.

The p-T phase diagram of $Nank1-7^\star$ calculated from the values of ΔG^u and ΔV^f_{eq} obtained at 2.1 M urea, as shown in figure 3.8, presents a shape that is similar to that of previously reported p-T phase diagrams for proteins [111, 51, 220, 221, 57, 58, 222, 202].

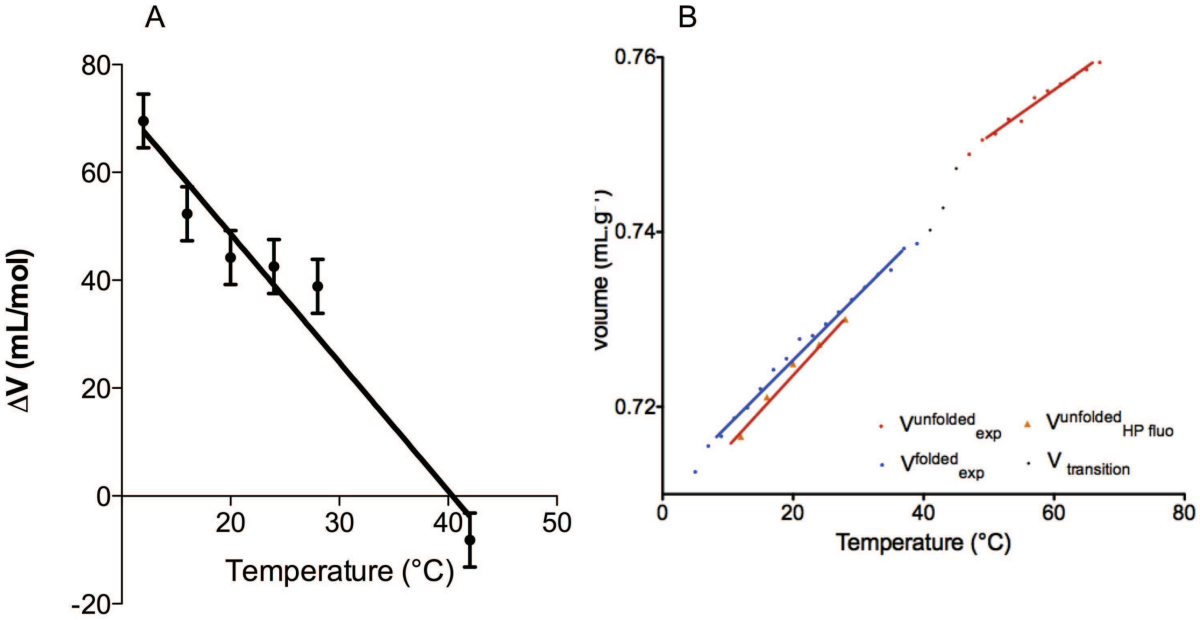

Figure 3.7: Evolution of ΔV_{eq}^{f} and volumes with temperature. A : Evolution of ΔV_{eq}^{f} obtained by high-pressure fluorescence spectroscopy with temperature, from 8 to 45.5 °C. B : Densitometry measurement of *Nank*1 – 7*, in 50 mM tris , 150 mM NaCl, pH 7.5. In blue and red triangles are represented respectively the folded and the unfolded state volume, in black circle are the unfolding transition. In orange triangles are represented the unfolded state volume at low temperature, recalculated from the high-pressure fluorescence obtained ΔV_{eq}^{f}.

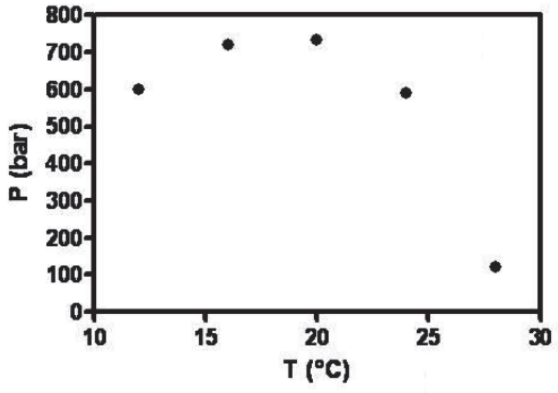

Figure 3.8: Calculated p,T transition diagram of *Nank*1 – 7* at 2.1M urea, from data obtained by high pressure fluorescence equilibrium measurements

We can see here that our technique permits us to clearly define the ΔV_{eq}^{f} value with low approximations on a large scale of temperature from equilibrium data.

3.2 Chain condensation in the pressure unfolded state

To investigate the structural changes of $Nank1{-}7^{\star}$ associated with changes in pressure and urea concentration, we performed pressure-dependent SAXS experiments [214, 145] (as described in Materials and Methods 2.6 p 45) at 24°C. We note that, although no aggregation was observed under pressure, these experiments were irreversible upon release of pressure due to aggregation at the high concentrations required for SAXS measurements. To have consistent data, we need to be sure that at each pressure applied, the quality of the datasets is good enough to allow an accurate study. As shown in the SAXS intensity profiles in figure 3.9 A, D and G, we observe a loss of signal when we apply pressure to the sample at all the urea concentrations used. This loss of signal leads in a lesser defined end of the plots, between 1.5 and 2.5 q/nm^{-1}. This decrease of intensity comes from the loss of contrast between the signal of the protein and the one of water, because water becomes more dense with pressure, as explained earlier. Despite this loss of sensitivity, data are clear enough to be used and interpreted. In the absence of urea, no appreciable change in the scattering profiles was observed between atmospheric pressure and 3 kbar (figure 3.9 A), and no significant change appear from 0 to 3 kbar in the pair distance profile figure (3.9 B), indicating that the protein is in its native state on the range of pressure applied. However, in the presence of 2.0 or 2.2 M urea, both the shape and the intensity of the scattering profiles are strongly pressure-dependent (figure 3.9 D and G respectively), and the pair-distribution functions (figure 3.9 E and H) reveal a significant increase in their average and largest pairwise distance with increasing pressure and urea concentration. We can remark too that at 2 and 2.2M of urea, the pair distribution functions at atmospheric pressure have profiles that remain highly similar to the native protein at atmospheric pressure, even if we can see a modest amount of unfolding because of the little enlargement of the peak, as urea destabilizes the proteins. We know that at these urea concentrations, the protein is just undergoing the beginning of the unfolding process, so this is consistent with both the fluorescence experiments shown in figure 3.1 and the denaturant-induced equilibrium unfolding, as shown in figure 2.4 p 27. While increasing pressure, at 2M urea (E), the profile changes to reach a larger, more expanded profile between 2000 and 3000 bars. This is indicative of an unfolding of the protein. According to the SAXS data, the unfolding seems to stop at 2000 bar, and we can't distinguish any differences between the different plots between 2 and 3 kbar. This is consistent with our pressure fluorescence spectroscopy measurements, as the inflections of the curves leading to the plateaus begin at about 2000 bars. By looking at the experimental data at 2 and 2.2M of urea together, we can see that the

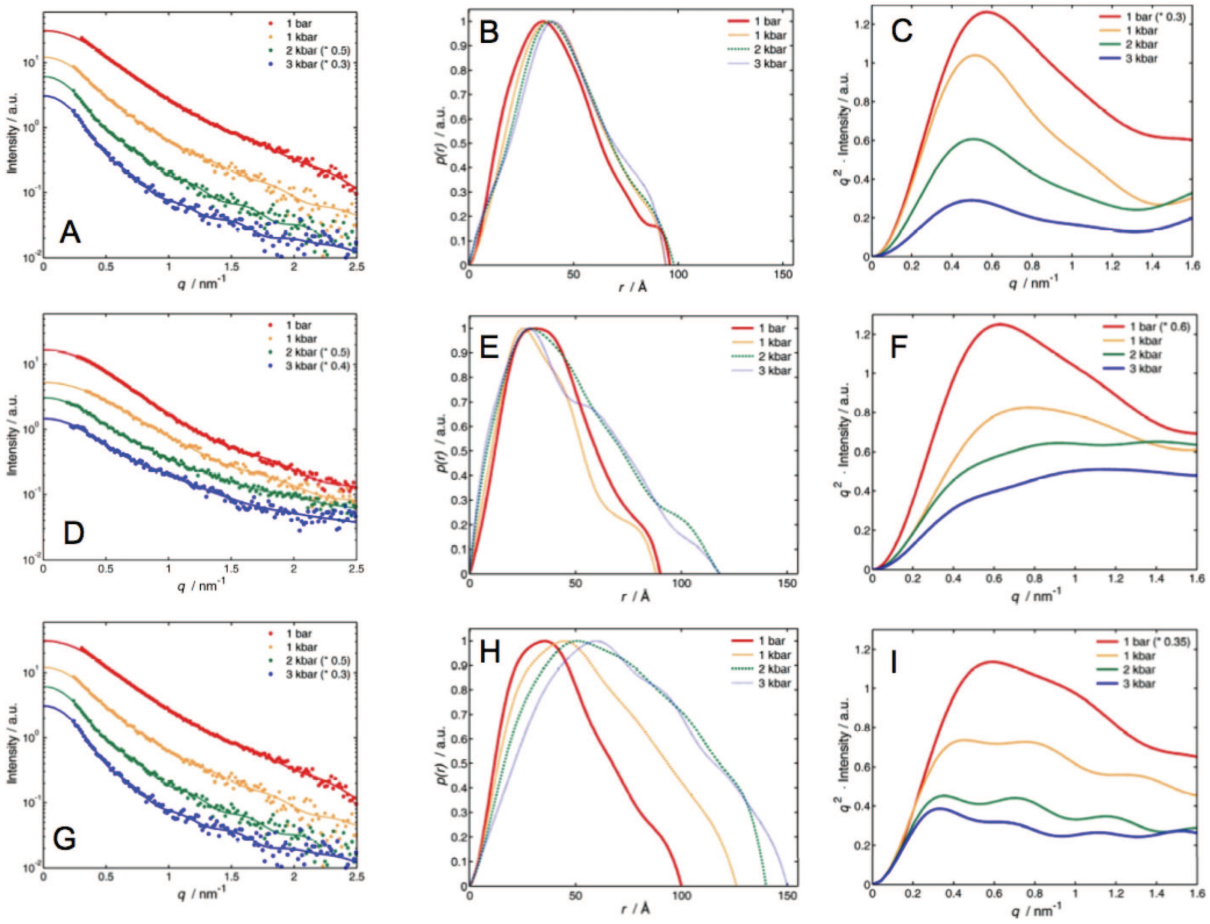

Figure 3.9: SAXS under pressure experiments on *Nank*1 – 7* at 20°C. A, D and G : intensity profiles of the SAXS datas respectively at 0, 2 and 2.2M of urea, from atmospheric pressure to 3 kbars. B, E and H : Pair distance distribution functions at different pressures and at respectively 0, 2 and 2.2M urea, from atmospheric pressure to 3 kbars. C, F and I : Kratky plots for respectively 0, 2 and 2.2M urea, from atmospheric pressure to 3 kbars [113].

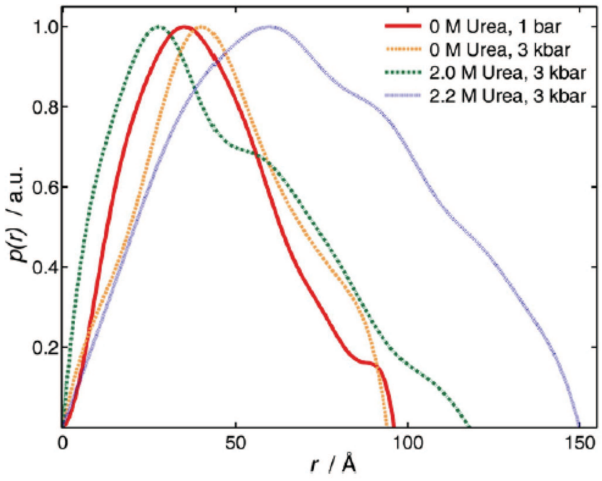

Figure 3.10: Comparison of the pair-distance functions at 24°C, at 1 bar and 0 M urea (red line), and at 3 kbar in presence of 0 M urea (yellow doted line), 2 M urea (green dotted line), and 2.2 M urea (blue line) [113].

profile of 2M urea at 2 kbar and the profile of 2.2M urea at 1 kbar are highly identical, and we can see that 2.2M 1kbar is a little more unfolded. This is similar to what was seen in the fluorescence denaturation curves of figure 3.1 at 24°C, with the same mass centers of the spectra at 2M urea at 2 kbar and 2.2M urea at 700 bars, and 2.2M urea a little more unfolded than 2M urea at 2 kbar. We can be sure that in all those experiments, the unfolding process is the same. Figure 3.10 shows a comparison of the pair-distribution functions for 0, 2, and 2.2 M urea obtained at 3 kbar, corresponding to the high-pressure plateau observed in the fluorescence profiles. The pair-distribution function is much broader at the highest urea concentration, indicating a significantly more expanded chain. Kratky plots at the high pressure plateau for 0, 2, and 2.2 M urea (figure 3.9 C, F and I) indicate that the chain retains an organized structural character at 3 kbar in the absence of urea but behaves as a Gaussian chain at high pressure for both urea concentrations. As the loss of structuration is consistent in the SAXS data with a complete unfolding at high pressure for each urea concentration, we can conclude from those experiments that the different plateaus we observed by fluorescence are not because of differences in the degree of unfolding of the tertiary structure, but because of differences in the close environment of tryptophan, coming from a high plasticity of the unfolded chain. To better understand this process, we used FTIR under pressure to probe the secondary structures with pressure.

3.3 Secondary structure in the pressure-unfolded state

Fourier Transformed Infra Red (FTIR) Spectroscopy is a powerful tool for dissecting the amount of the different secondary structures within the proteins. We used FTIR under pressure, as described in materials and methods 2.7 p 47, to understand the

secondary structure changes with pressure and urea, in order to understand if some resulting secondary structures would explain the changes in the fluorescence plateaus. To assess the evolution of secondary structure accompanying the fluorescence- and SAXS-detected unfolding transitions, high-pressure FTIR experiments were carried out in the absence or presence of varying concentrations of urea at 20°C. So that the urea bands would not obscure the protein amide bands, we used ^{13}C-labeled urea as a denaturant [204, 223]. Analysis of the shift as a function of pressure, as reported in figure 3.11 A, C and E, give after deconvolution the ratios of secondary structure of each type, in each condition.

Like the SAXS experiments, the FTIR experiments were irreversible upon return to atmospheric pressure due to aggregation at low pressure, but again no aggregation occurred under pressure during experiment, because pressure destabilizes protein oligomers and aggregates [224, 225]. In the absence of urea (figure 3.11 B), the pressure-induced loss of $\approx$ 25% α-helical content took place over a very broad pressure range (between $\approx$ 2.5 and 6 kbar), mostly beyond the range of our fluorescence and SAXS high-pressure cells. The lack of any change in secondary structural content below 3 kbar in the absence of urea is consistent with the insignificant changes in fluorescence and SAXS over the same pressure range. In the presence of 2M urea (figure 3.11 D), a $\approx$ 25% loss in α-helical content occurred between 1.5 and 4.5 kbar, with about half of the change occurring below 3 kbar, corresponding to the plateau observed in the fluorescence and SAXS experiments. Due to the irreversible nature of the FTIR transition, calculation of an equilibrium volume change is not possible, although we note that the sensitivity of the FTIR transition to pressure is substantially less (by $\pm$50%) than that observed in the reversible fluorescence-monitored transitions. At 2.35 M urea (Fig. 5 C), we observed a lower helical content already at atmospheric pressure, and the $\pm$35% pressure-induced loss of α-helix occurred between 0 and 4 kbar. At 2.2 kbar and 2.35 M urea, the fluorescence spectrum of *Nank*1 – 7* is as red-shifted as possible (equivalent to N-acetyl L-tryptophanamide in water), and the chain expansion, as observed by SAXS, has also reached a maximum. Therefore, the FTIR results indicate that the *Nank*1 – 7* α-helices resist despite considerable disruption of the tertiary structure by pressure. Indeed, it has been shown previously that pressure has very little effect on isolated α-helices [226, 227]. These data at equilibrium allow us to have a better understanding of things happening at each pressure. Indeed, with fluorescence, we follow the tryptophan environment changes depending on pressure, and this environment seems more and more hydrated (known from fluorescence datasets), while some secondary structures remains (FTIR datas). This indicates that only the tertiary structure is responsible for the tryptophan hydration, what is consistent with its external position within the structure, as seen in figure 2.3 B and C, p 26.

Moreover, despite the differences in helical content, the ΔV_{eq}^{f} calculated from fluorescence experiments remains constant between the different urea concentrations. So

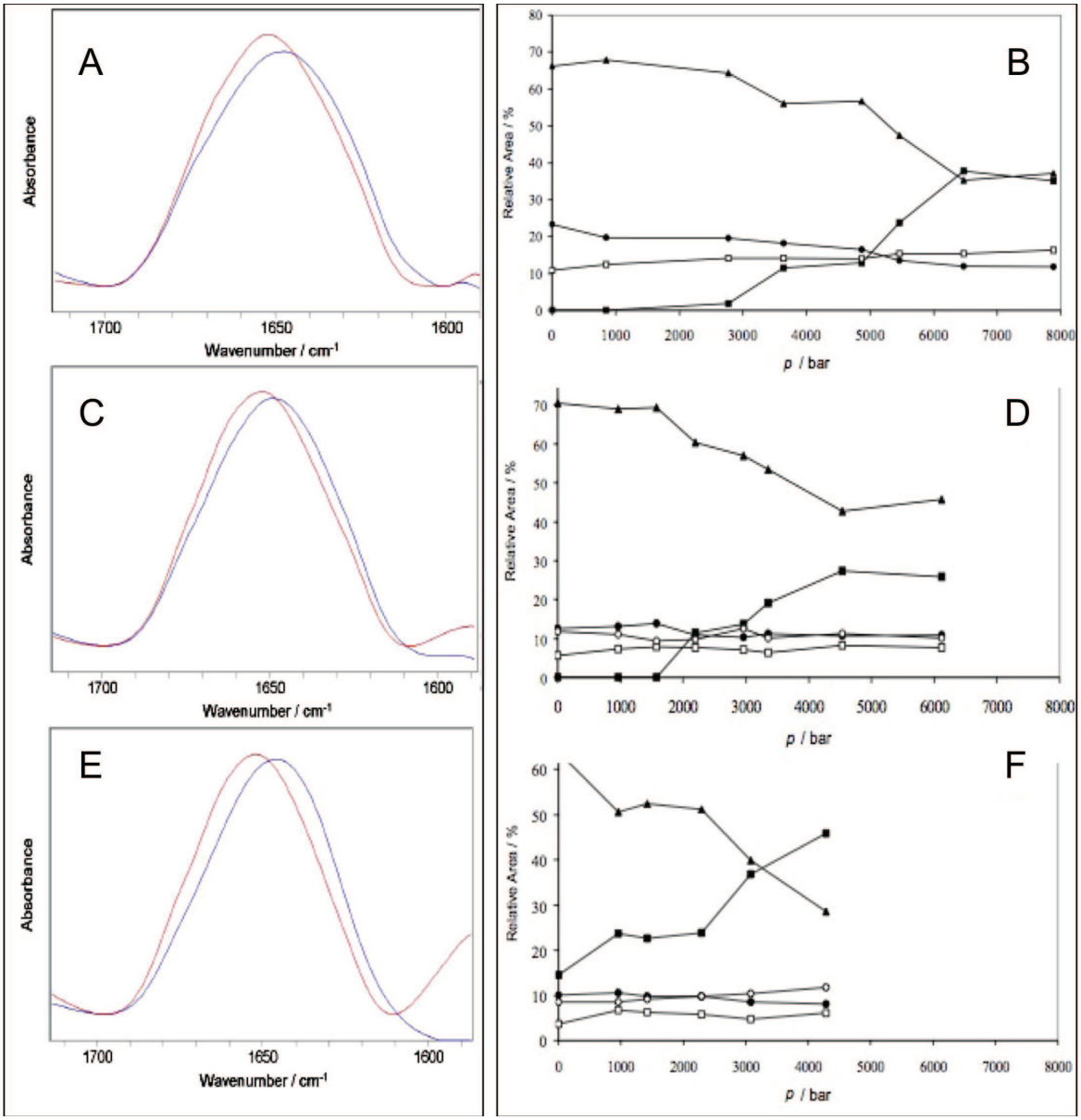

Figure 3.11: FTIR under pressure measurements of $Nank1-7^{\star}$ at 20°C. A, C and E : Absorption spectra of the amide I band, respectively at 0, 2 and 2.35M of urea. B, D, F : Details of secondary structures at 0, 2 and 2.35M of urea respectively. The black and white circles corresponds to loops, black triangles to $\alpha - helices$, black squares to random-coil and white squares to $\beta - sheets$ [113].

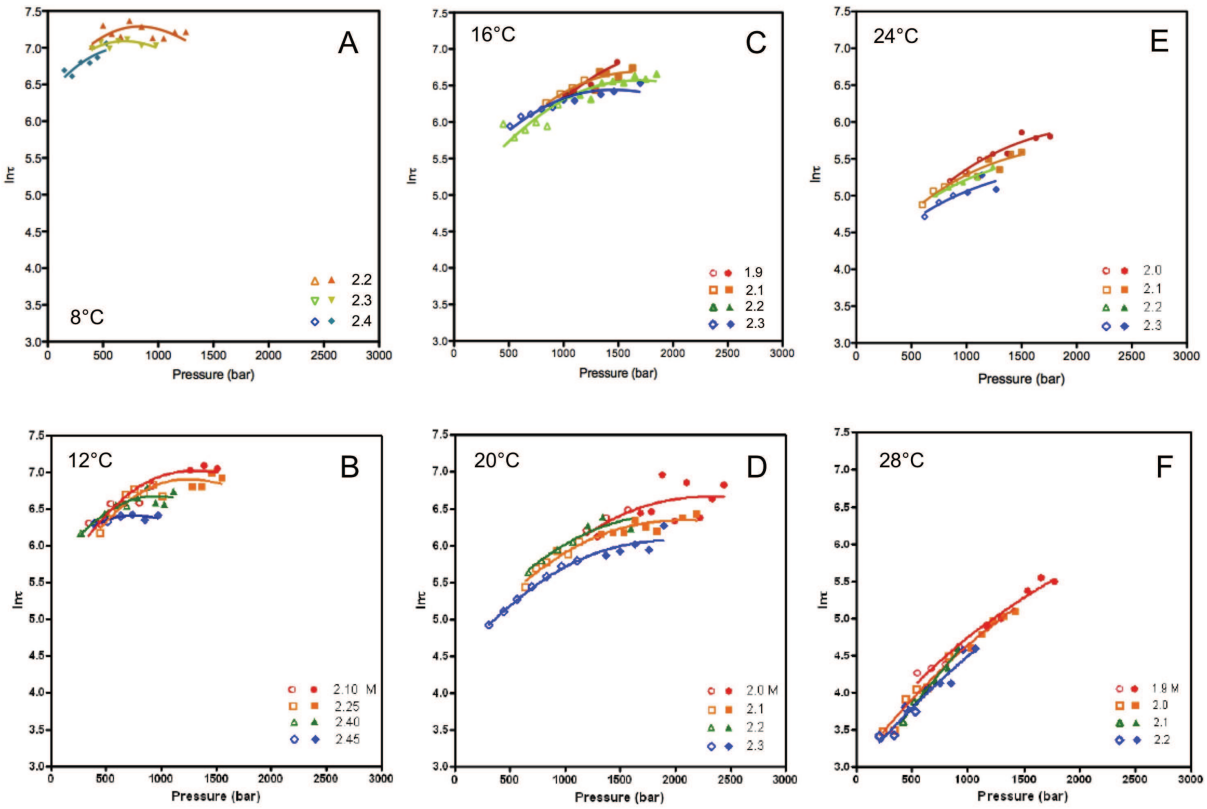

Figure 3.12: Natural logarithm of the p-jump fluorescence detected relaxation time for *Nank*1–7* as a function of pressure for varying urea concentrations (as indicated in the Figures) at A) 8 °C, B) 16 °C, C) 24° C, D) 12°C, E) 20°C and F) 28°C. Open symbols represent data points below the transition midpoint while closed symbols represent data points above the transition midpoint.

we can observe that the ΔV_{eq}^{f} does not depend on these residual helices. This can be a clue to say that the ΔV_{eq}^{f} is highly dependent on the 3D structure of the protein and that changes in secondary structure don't change anything in the structuration of the voids leading to the ΔV_{eq}^{f}. So we can hypothesize that hydration changes occuring with secondary structure formation is not a parameter of the ΔV_{eq}^{f}. As helices do not contain voids, we can also be sure that the ΔV_{eq}^{f} can't change because of mispacking in the secondary structure. The first step in the accurate determination of the provenance of the ΔV_{eq}^{f} on this model protein has just been done.

3.4 Pressure-jump kinetics

From the pressure-jumps kinetics, we have a view of the time dependence of the transition from the folded population to the unfolded one. Hence, we have the possibility to calculate the rate constant for folding and unfolding depending on the final pressure of the jumps. From these data, we are able to obtain the relaxation times of the protein, for each final pressure of the pressure-jumps, and the plot of the $ln(\tau)$ depending on pressure is highly interesting (figure 3.12). Pressure-jump relaxation profiles

[Urea] M	1.9		2.0		2.1		2.2		2.3		2.35		2.4		2.45		2.5	
T (°C)	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol	$k°_f$ (s^{-1})	$\Delta V^{\#}_f$ kcal /mol
8					0.0066	66.19	0.00138	43.61	0.00127	46.68			0.00125	59.94				
12					0.0038	60.28	0.0042	47.58					0.0029	37.15	0.0027	29.76		
16	0.0091	43.64			0.0061	40.8	0.0066	41.98	0.005	39.61			0.0044	44.16			0.0030	44.54
20			0.0098	37.53	0.0087	38.20	0.0076	38.8	0.0098	41.05	0.0135	60.21						
24			0.014	45.03	0.013	39.76	0.011	45.46	0.0114	46.80			0.0069	59.11				
28	0.029	54.6	0.032	58.45	0.037	69.61	0.026	62.46										

Table 3.2: Kinetic parameters from the p-jump fluorescence studies on Nank1-7

were obtained as a function of temperature and urea concentration by monitoring the fluorescence intensity at the maximum of the folded-state fluorescence emission as a function of time after each pressure jump. The relaxation time under all these conditions was consistent with a single exponential decay of the profiles of intensity versus time after the pressure jump. When we look at the plot of the natural logarithm of the relaxation time (ln τ) of $Nank1-7^{\star}$ against pressure (figure 3.12), we can observe for all temperatures that the more we increase the urea, the lower are the $ln(\tau)$ values because the more we increase the urea concentration, the more we destabilize the folded state of the protein, and the more ΔG_{eq} decreases, and hence the folding decreases.

Examination of the plots of the natural logarithm of the relaxation times versus pressure (figure 3.12 A-D) reveals that, as in the case of the equilibrium unfolding, urea has no real effect on the volume change of unfolding, as the $ln(\tau)$ exhibit the same slope for each temperature at all urea concentrations. However, it can be seen that the high-pressure slopes of the profiles are strongly dependent on temperature, with a distinct downward curvature (Chevron-like behavior) observed at low temperature (8 and 12°C, figure 3.12 A and B) and high pressure but fairly linear profiles over the entire unfolding transition at higher temperatures (C to F).

Data obtained at all temperatures were analyzed for the activation volume and atmospheric rate constant for folding, respectively $\Delta V_f^{\#}$ and k_f^0 as described in materials and methods, and results are summarized in table 3.2. The corresponding parameters for unfolding ($\Delta V_u^{\#}$ and k_u^0) were derived from the equilibrium free-energy and volume changes. As shown in figure 3.13, the ΔV_{eq}^f and $\Delta V_f^{\#}$ are roughly the same, within the error at 20°C for all the urea concentrations used, and this pattern is general for all temperatures used, depending on the quality of measurements. We can conclude that denaturant doesn't change anything for all the volumetric properties (ΔV_{eq}^f, $\Delta V_u^{\#}$ and $\Delta V_f^{\#}$) of the protein. It is assumed that urea destabilizes proteins by taking the place of water in the surrounding of the protein, so it seems that changing the hydration shell of the protein does not have a significant effect on ΔV properties of the protein because it changes the volumes of the folded, unfolded and transition states the same way. So it seems that in the parametrization of the ΔVs, the $\Delta V_{hydration}$ term from the surrounding water is close to 0. The major aspect for the treatment of these datasets

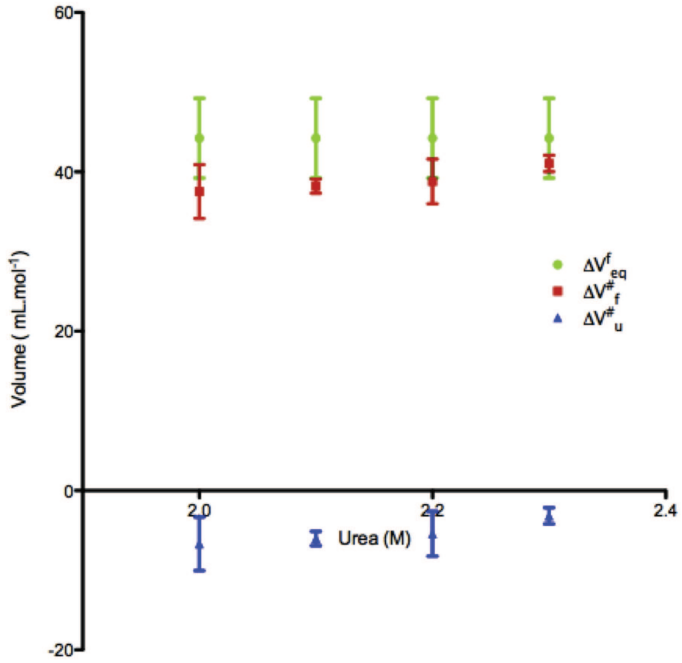

Figure 3.13: Evolution of the ΔV_{eq}^{f}, $\Delta V_{f}^{\#}$ and $\Delta V_{u}^{\#}$ for all the urea concentrations at 20°C. In green the ΔV_{eq}^{f}, in red the $\Delta V_{f}^{\#}$ and $\Delta V_{u}^{\#}$ are in blue.

is that, because of this highly similar values of the activation volume change independently of the denaturant, we will be able to determine the values of these volume changes just by duplicating the experiments for one only urea concentration (chosen to be in the middle of the urea concentration scale used), or by calculating the average.

At low temperature, the specific volume of the TSE was found to be significantly closer to that of the folded state than to that of the unfolded state ($\Delta V_{f}^{\#}$ =+50 mL/-mol, $\Delta V_{u}^{\#}$ = -18 mL/mol) at 12°C, indicating that at the barrier, a significant degree of dehydration occurs. Comparison of the temperature dependence of the activation volumes for folding and unfolding with that of the equilibrium volume change (figure 3.14) reveals that the entire temperature dependence of the ΔV_{eq}^{f} resides in the activation volume for unfolding, $\Delta V_{u}^{\#}$. Moreover, the value of $\Delta V_{u}^{\#}$ is negative at low temperature and changes sign to become positive at high temperature. Indeed, unlike denaturant based unfolding, pressure-induced unfolding need not display Chevron-like behavior, as the sign of the activation volumes can be either positive or negative. The temperature dependence of the activation volumes, like that of the equilibrium volume change, is equivalent to the difference in thermal expansivity between the two states implicated in the reaction. We observe no dependence of the $\Delta V_{f}^{\#}$ on temperature (+0.12 ± 0.54 mL/mol.K). In contrast, we observe a strong dependence of $\Delta V_{u}^{\#}$ on temperature (+2.08 ± 0.46 mL/mol.K), highlighting a significant difference in thermal expansivity between the TSE and the folded state. This value is equivalent, within error, to that (+1.72 ± 0.25 mL/mol.K) observed for the ΔV_{eq}^{f} (the difference in thermal expansivity between the unfolded and folded states). According to our model, at the barrier, the protein is

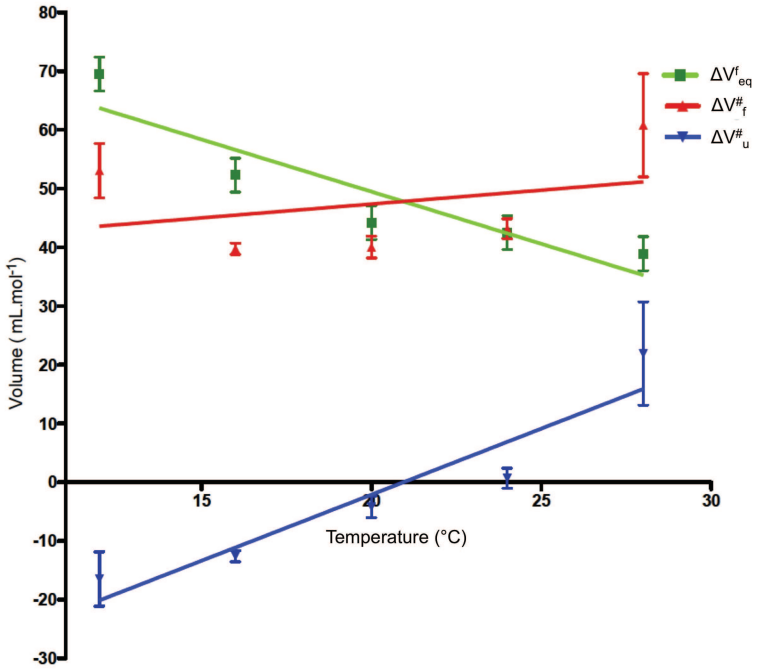

Figure 3.14: Evolution of activation and equilibrium volume changes with temperature. In green the ΔV^f_{eq}, in red the $\Delta V^\#_f$ and in blue the $\Delta V^\#_u$. The lines appearing are simple linear fit of the raw datas.

collapsed, forming a significant proportion of the native-state solvent-excluded volume, and is thus relatively "dry" (figure 3.15).

Moreover, we deduce that the thermal expansivity of the TSE is equivalent to that of the unfolded state.

In figure 3.15, we proposes another point of view on these data, more folding mechanism oriented. In this representation, we can see that at small temperatures, the $\Delta V^\#_u$ is negative because the TSE is placed in terms of volume between the unfolded state (the lower one) and the folded state (the higher one). The more we increase temperature, the more the TSE will become larger since the energetic constraints are smaller than for the folded state. Its volume hence reaches that of the folded state when the $\Delta V^\#_u$ is slightly equal to 0, and then becomes larger than the folded state at higher temperatures.

After that, the $\Delta V^\#_u$ becomes positive and will continue to increase linearly with temperature, becoming bigger and bigger in absolute value. This is the first direct observation of the change of sign of $\Delta V^\#_u$. We can see that in order to change the sign of the $\Delta V^\#_u$, we just have to change the volumetric positions of each state.

Now that we have well characterized the $Nank1-7^\star$ pressure-induced unfolding, we can presume that we will be able to analyze accurately the evolution of the ΔV^f_{eq} depending on the size and sequence, thanks to the numerous mutants we have. Indeed, within a range of 3 kbar, we managed to unfold completely this protein and obtained consistent and accurate values for all the thermodynamic parameters. Although pressure managed to unfold the protein in a highly cooperative way, it's already sure that

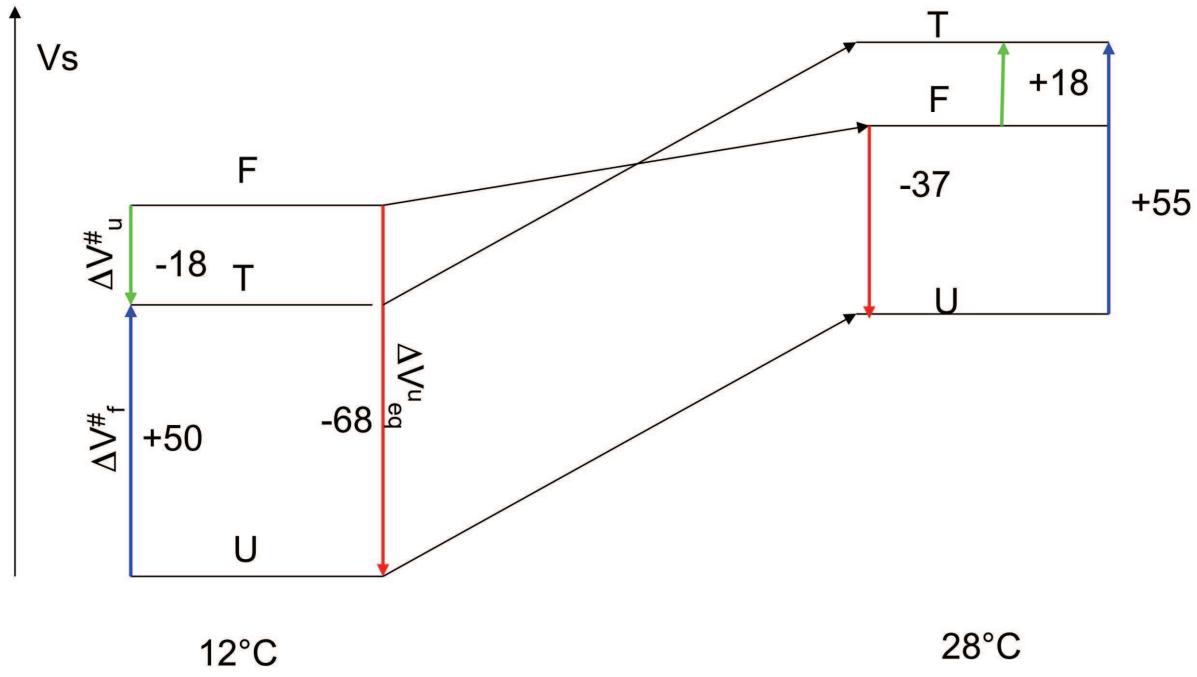

Figure 3.15: Schema of the evolution of volume changes with temperature for $Nank1-7^{\star}$. $\Delta V_f^{\#}$ is shown in blue, $\Delta V_u^{\#}$ in green, and ΔV_{eq}^{u} is shown in red.

the ΔVs we can observe comes from the 3D structuration of the protein, and that the secondary structures don't have an important impact on the ΔV_{eq}^{f} value. Finally, we can too make the hypothesis that the folding process of this protein is done by nucleation from small $\alpha-helices$ that appears early in the folding process. These data allow us to examine this protein and all it's mutants in a temperature/denaturant/pressure multifactorial approach, in order to understand how the size, the sequence and the temperature can correlate with the ΔV_{eq}^{f}.

These results make us clearly see that we don't understand yet the complexity of the TSE and of the expansivity of all the states. Indeed, even if it was shown that the TSE of $Nank1-7^{\star}$ is majorily unfolded, with only three motifs folded in the middle of the structure and not defined close contacts [176], its volume is really close to the one of the folded state, as the $\Delta V_f^{\#}$ is close to ΔV_{eq}^{f}. This can give clues on the parametrization of the forces involved in the definition of the ΔV_{eq}^{f}, and it shows us that we really need to understand where are the voids leading to the ΔV_{eq}^{f} in the structure, how they form and how they are taken into account in the global ΔV_{eq}^{f}. Moreover, the expansivity of the TSE is close to the one of the unfolded state, this is a proof that the TSE can physically be unfolded like (in terms of structure and expansivity) while thermodynamically react in a folded-like way (in terms of volume). The relationships between volume, expansivity and stability clearly are not yet defined.

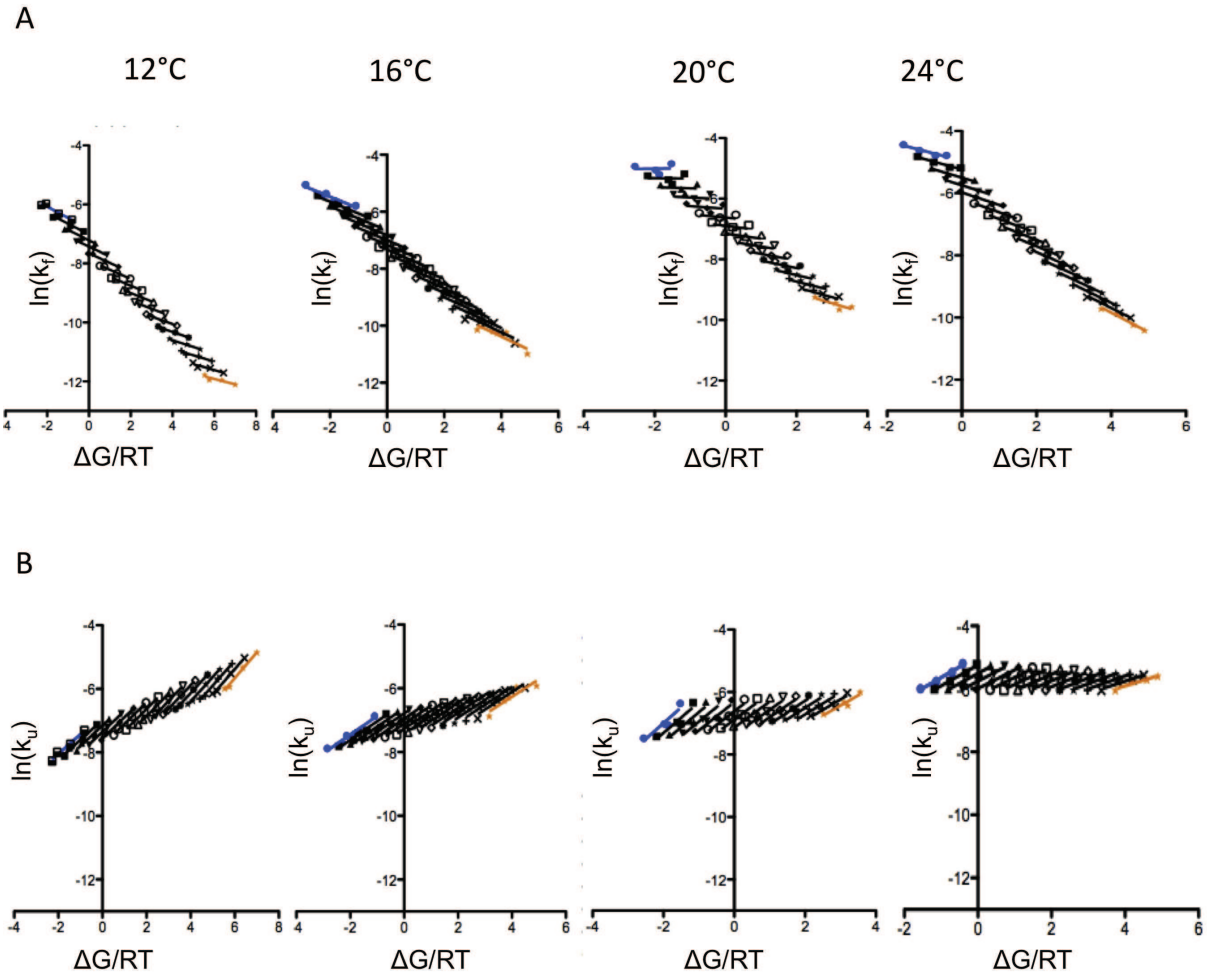

Figure 3.16: Evolution of $ln(k_f)$ and $ln(k_u)$ over $\Delta G/RT$, depending on temperature. In blue are plotted the datas at 200 bars, in orange at 3000 bars. A : $ln(k_f)$. B : $ln(k_u)$. Each data point of the same line was obtained at different urea concentration.

3.5 Pressure effects on $Nank1-7^{\star}$

From the obtained values of k_f^0 and k_u^0, we were able to determine the constant rate for folding (k_f) and unfolding (k_u), depending on the pressure applied, for each of the temperatures and at each urea concentration. In order to analyze these datas, we made the plot of the $ln(k_f)$ and $ln(k_u)$ over $\Delta G/RT$ (normalization of the protein stability by temperature) to understand how pressure acts on these two parameters, as shown in figure 3.16. These representation are called linear free energy plots and their slope represents the fractional contribution of the folding (or unfolding) rate constant to the change in stability.

This plot is the representation of the evolution of the $ln(k_f)$ and $ln(k_u)$ at each pressure, depending on the stability of the protein. Blue corresponds to 200 bar, and orange to 3 kbar. All the intermediate pressures are in black. First, by eye, we can observe that in figure 3.16 A and B, the linear regression of the change in $ln(k_u)$ and $ln(k_f)$ are not completely parallel to the Y-axis. So it is possible to reach each given stability at all the pressures, just depending on other parameters used, as urea.

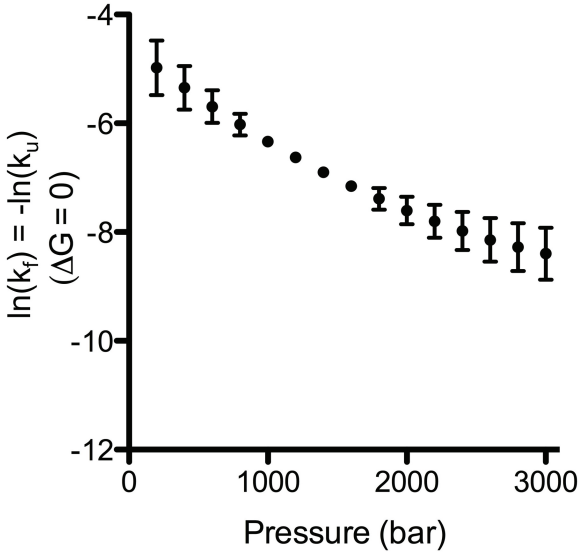

Figure 3.17: Evolution of $ln(k_u)$ (= $ln(k_f)$) with pressure when $\Delta G = 0$ for $Nank1-7^*$ at 20°C.

We can also remark that at all temperatures used, the $ln(k_f)$ (figure 3.16 A) decreases more with pressure (from -6 to -12) than the $ln(k_u)$ increase (from -8 to -5). So pressure acts majorily (2/3) on the k_f parameter, and poorly (1/3) on the k_u. Here we have a new clue on the pressure effect on this protein. As seen in figure 3.19, this implies that pressure has a big effect on the energetic level of the folded state, and the TSE, but less on the unfolded state. Indeed, the more pressure is increased, the more it will be hard for the unfolded protein to "climb" the energetic barrier, while the barrier for unfolding changes less. At high pressure, the energetic diagram will be completely inverted, with a higher-energy level for the folded state than for the unfolded state, so unfolding will be favored. The TSE will be largely higher in energetic degree, so we need to ask ourselves why the unfolded-like TSE is more pressure-sensitive than the unfolded state, even though its expansivity is similar. It is because at a given temperature (21°C), its volume is similar to the folded state.

Another interesting fact about figure 3.16 A is that if we look at the blue and orange lines, we can see that the slope of the orange line at 3 kbar is increased with regard to the blue one at 200 bar, so it means that pressure effects are not linear, but change with the stability of the protein, what we can too see with the representation of $ln(k_u)$ or $ln(k_f)$ versus pressure when $\Delta G = 0$ at 20°C for example (figure 3.17). Indeed, we can observe that in this particular condition, when $\Delta G = 0$, so when the folded and unfolded states are equiprobable, $ln(k_u)$ and $ln(k_f)$ don't have a linear evolution with pressure. Pressure seems to slow the whole process at equi-energetic values so this may represent a pressure effect on the pre-exponential factor in $k = k_0 e^{\frac{-p.\Delta V}{RT}}$. This corresponds to the diffusion coefficient on the free energy landscape. This non linear dependence of the $ln(k_u)$ and $ln(k_f)$ with pressure is best seen in figure 3.18.

We can observe that at a stable temperature, for example 10°C, and varying pres-

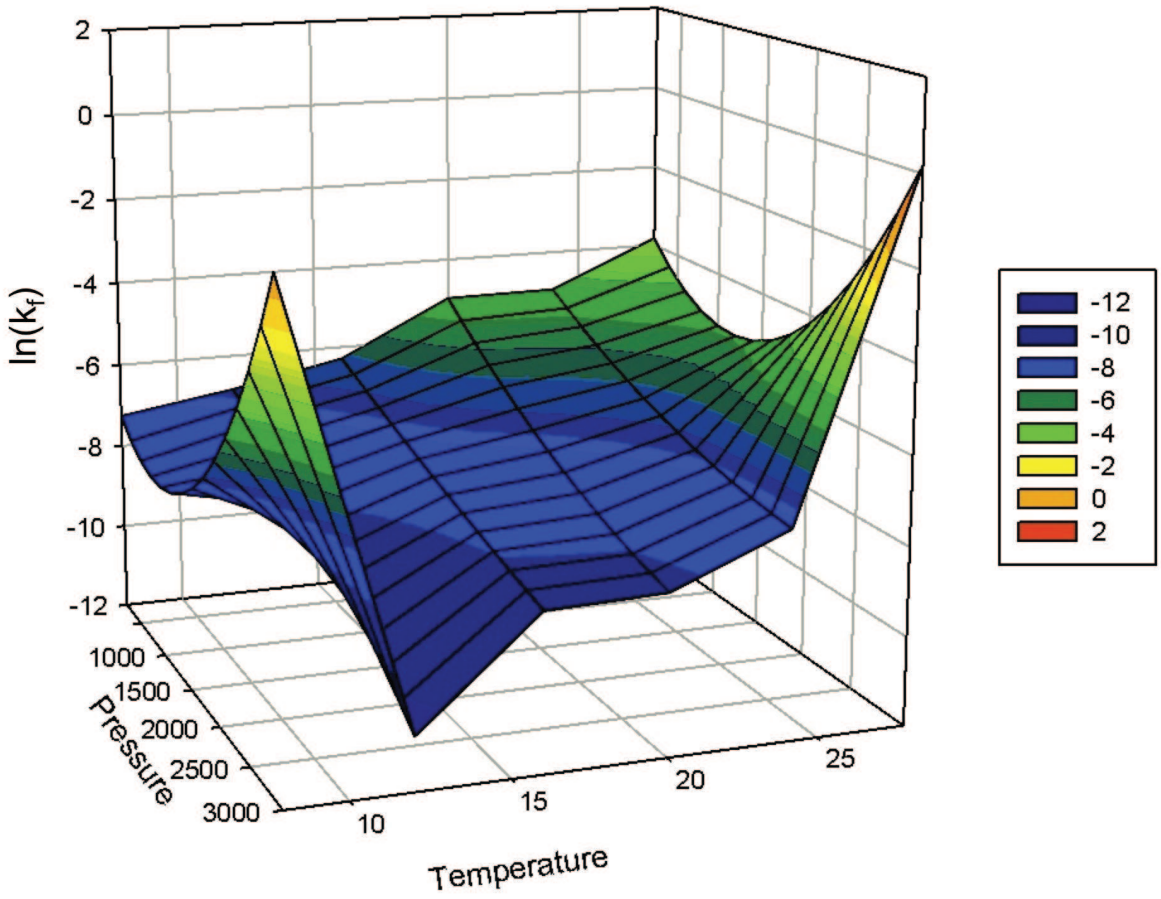

Figure 3.18: 3-dimensional representation of $ln(k_f)$ as a function of pressure and temperature. This representation uses the same set of data than in figure 3.16 but includes data at 8 and 28°C.

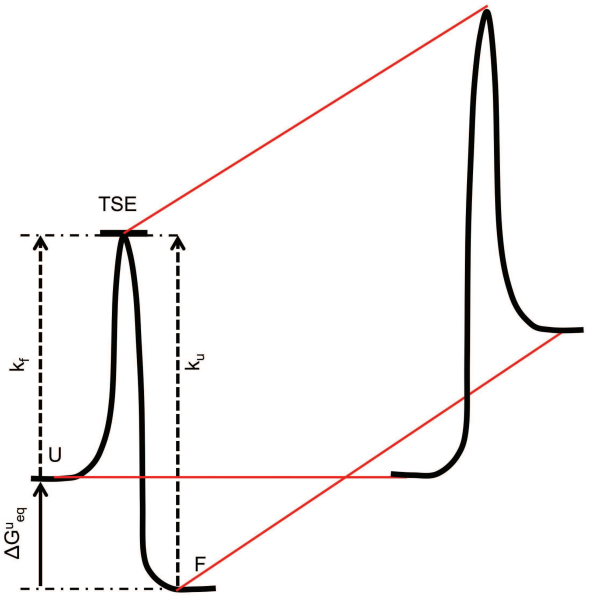

Figure 3.19: Energetic diagram of the protein folding. U means the unfolded state, F means the Folded state. Schema of an energetic diagram of a typical protein and its changes after effect of high pressures. We can observe that in this case k_u remains constant, and k_f increases.

sure, we observe important changes in the rate constants, demonstrating that pressure induces changes in the broad of the energetic barrier. At 10°C, from atmospheric pressure to 1500 bar, the rate constant increases, as $ln(k_f)$ decreases, so pressure make the energetic barrier more narrow, and from 1500 bar to 3 kbar, changes are opposite, with a highly important increase of the $ln(k_f)$, ie an important decrease of the folding rate. This increase of the folding rate means that the barrier is broader and broader with pressure. The exact same pattern can be seen at 28°C. But for intermediate temperatures, changes are more flat. From now we can't understand yet why pressure changes so much the width of the barrier at certain pressures and temperature, and not at other ones.

3.6 Short discussion

The folding energetics and kinetics of $Nank1-7^*$ have been characterized and linked to structural folding pathways [176, 184, 146], and prior results provide a solid basis for comparison and interpretation of the high-pressure studies presented here. These SAXS and fluorescence studies reveal pressure induced unfolding with a plateau near 3 kbar. $Nank1-7^*$ ΔV^f_{eq}, with a value of $\approx +$ 45 ml/mol, exhibits a rather small pressure sensitivity [93] given its size. Comparison of the fluorescence, SAXS, and FTIR data reveal significant effects of urea on the physical properties of pressure-unfolded $Nank1-7^*$. At moderate urea concentration, the pressure-unfolded ensemble, although a Gaussian chain, is not completely expanded and retains a major fraction of the native secondary structure. Indeed, the chain becomes much more expanded

and the tryptophan significantly more exposed to solvent, and the secondary structure content decreases to ≈50% that of the native protein at the high-pressure plateau reached in the presence of slightly higher urea concentrations. It is not surprising that the secondary structure persists at high pressure. It has been shown that pressure has little effect on α-helices, and hence, the loss of secondary structure under pressure typically occurs only if the helices are not intrinsically stable in the absence of tertiary interactions. Estimation of helix stability for the Nank helices using the program Agadir, revealed that in comparison to the ≈60–65% helicity of the peptide described by Baldwin and co-workers (AAAAKAAAAKAAAAKA) [228], the $Nank1-7^{\star}$ helices were highly unstable. However, the residual interactions responsible for the relative chain compaction and blue tryptophan emission at low urea likely stabilize the helices. Results of recent experimental and theoretical studies on volume changes suggest that pressure-induced unfolding results from internal voids in the native state. Since these voids are not always homogeneously distributed throughout the structure, certain regions can be less pressure-sensitive than others. It is interesting that fits of the fluorescence equilibrium profiles to a two-state unfolding model yielded equivalent volume changes for unfolding, despite the larger plateau values for the average emission wavelength of the tryptophan with increasing urea. We and others have shown that denaturant itself does not influence the value of the volume change upon unfolding [125, 213]. However, we observe here that increasing urea leads to greater solvent exposure of the tryptophan residue and a more expanded Gaussian chain at high pressure, without any change in molar volume. Such an observation is consistent with our working model, in which disappearance of internal voids upon unfolding, rather than density changes due to differential (polar versus hydrophobic) solvation, provide the bulk of the volume change upon unfolding. Once the protein has expanded sufficiently to allow occupation of these internal voids by water, pressure has no further effect. Hence, we suggest that the pressure unfolded state, which is not maximally expanded or devoid of α-helices, may more closely resemble the unfolded state that is sampled by this protein under native conditions than those unfolded by denaturant or temperature perturbation. Consistent with the lack of energetic barriers between these unfolded conformations, extrapolation to atmospheric pressure and 0 M urea of the urea- and pressure-dependent fluorescence free energies of unfolding yields values that are in reasonably good agreement with the global stability of the protein previously determined from urea denaturation profiles under similar conditions [188]. For this transition from the folded state to the compact denatured state, the p-T phase behavior of $Nank1-7^{\star}$ displays the shape observed for the few globular proteins that have been studied to date. Indeed, it has already been shown that $Nank1-7^{\star}$ exhibits a large heat capacity upon unfolding [152, 148], and we show here that it also exhibits a significant change in thermal expansivity, the second most important parameter giving rise to the curvature in the p-T plane. Our pressure-jump kinetics

measurements on *Nank*1 – 7* fit well to a two-state unfolding/refolding model. In contrast, thorough studies of *Nank*1 – 7* folding and unfolding kinetics using denaturant revealed multiple kinetic phases, resulting from denatured-state proline isomerization and an on-pathway unfolding intermediate [184]. This intermediate displayed tryptophan emission and circular dichroism similar to the folded state, a modest free-energy gain, and ≈40% loss of solvent-accessible surface area compared to the unfolded state. Proline isomerization reactions, albeit quite slow, are not rate-limiting at high pressure, due to the large positive activation volume for folding, which slows the main folding transition even further. Hence, it is not surprising that the proline-based kinetic complexity was not observed in our pressure-jump studies. Likewise, the on-pathway intermediate was not detected in our pressure jump studies, either because it was not highly populated enough to be observable in our experiments or because of the pressure-dependent increase in the main folding barrier. The TSE for the single transition that we observe exhibits a molar volume at low temperature quite close to that of the folded state. According to our working model of pressure effects, this indicates that the bulk of the solvent excluded packing defects have formed in this TSE. In contrast, the thermal expansivity of the TSE is equivalent to that of the unfolded state, indicating that most of the native contacts that limit the thermal expansion of the folded state are not formed at the barrier. Previous studies on *Nank*1 – 7* folding indicated that it folds inside-out, with the main barrier corresponding to the formation of the central repeats. Our expansivity data suggest that this TSE remains quite malleable. An expansivity for the TSE of *Nank*1 – 7* equivalent to that of the unfolded state is in contrast with that found for the TSE of Snase (the only other protein for which this quantity has been measured), which was nearly equivalent to that of the folded state, consistent with a substantial degree of structure at the barrier for this globular protein. That the tertiary interactions in *Nank*1 – 7* are not locked in at the folding barrier may derive from the elongated and modular nature of the hydrophobic core of *Nank*1 – 7*, which imposes fewer spatial constraints on the structure. In conclusion, these pressure perturbation studies on the ankyrin repeat domain of the Notch receptor reveal a high degree of plasticity in both the unfolded-state ensemble and the TSE. The physical properties of these states, such as chain condensation and thermal expansion, are strongly dependent upon conditions. Overall, the observation of a less drastic effect of pressure, as compared to chemical denaturants, can be understood in light of the different physical basis for their perturbations. Urea interacts with the protein, leading to more expanded chains in the unfolded state via an increase in the persistence length. Pressure effects arise simply due to the smaller system volume when the protein is in the unfolded state, and therefore do not necessarily result in maximal expansion of the chain, or maximal loss of helical content. Our understanding of protein folding mechanisms is dependent upon our ability to characterize the conformational states on the protein folding landscape. Here, we show that pressure

can reveal a unique set of conformational states with distinct structural and thermodynamic properties, thus increasing our knowledge and broadening our perspective of this complex process.

Now that we have characterized how $Nank1-7^{\star}$ unfolds with pressure and how its ΔV_{eq}^{f}, $\Delta V_{f}^{\#}$ and $\Delta V_{u}^{\#}$ evolve with temperature, we are able to use $Nank1-7^{\star}$ as a reference for all the mutants we have, and to make a complete study of the effects of pressure on repeat proteins.

Chapter 4

Study of Ankyrin deletion constructs

In the present study, we have sought to examine in a systematic fashion the factors affecting the value of the ΔV_{eq}^{f} for proteins using ankyrin repeat domain of the Notch receptor as a model system. Previously shown results on $Nank1-7^{\star}$ revealed that the ΔV_{eq}^{f} at 20°C is rather small, given its size. Indeed with 248 amino acids, the weak correlation of volume changes with size based on bibliographic data [108, 107, 220, 133, 105, 131, 229, 106] predicts a ΔV_{eq}^{f} of $\approx$ 125 mL/mol, whereas the measured value is only 44 mL/mol. We postulated that the small volume change might be related to the sequence and packing of the protein, which is quite different from the globular proteins that present the remaining points in the correlation. To test the correlation between the size and sequence of the protein and ΔV_{eq}^{f}, while maintaining the same fold, we have investigated the pressure effects on Nank constructs with N- and C-terminal deletions of one or more repeats (respectively $Nank2-7^{\star}$ / $Nank3-7^{\star}$ / $Nank4-7^{\star}$ and $Nank1-6^{\star}$ / $Nank1-5^{\star}$), two with an alanine to glycine mutation at analogous positions in the second and fifth repeat (NAG2 and NAG5) and finally two small constructs of two and three consensus repeats (respectively NW and NWC).

4.1 Study of the ΔV_{eq}^{f} dependence with size

4.1.1 Equilibrium unfolding at 20°C

N-terminal deletion constructs

$Nank2-7^{\star}$ contains motifs 2 to 7, so contains only completely folded motifs. This mutant constitutes an interesting base for the ensemble of this study, as it allows us to follow similar folding mechanisms with the other completely folded mutants and to observe distinctions coming only from the first partially folded repeat of $Nank1-7^{\star}$. $Nank3-7^{\star}$ is smaller than $Nank2-7^{\star}$ and contains repeats 3 to 7, which are all completely folded. The $Nank4-7^{\star}$ mutant contains four repeats and is the only protein we have that is not fully stable at ambient temperature, without urea [175]. For this particular construct, it is necessary to add 750 mM of Trimethylamine N-oxide

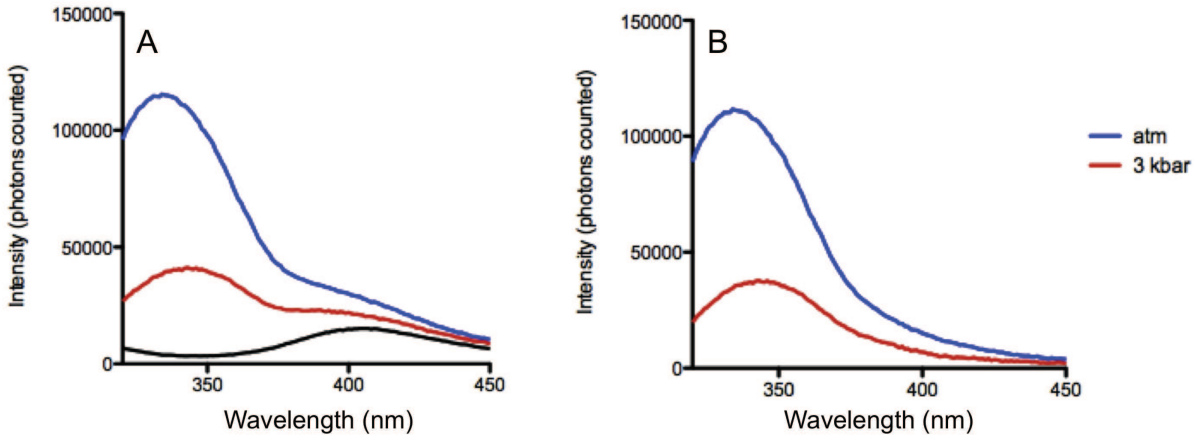

Figure 4.1: Analysis of tryptophan intrinsic fluorescence emission spectra of $Nank4-7^{*}$. A : the raw tryptophan emission spectra at atmospheric pressure (blue) and at 3 kbar (red). Buffer only is represented in black. B : Tryptophan emission spectra corrected by subtraction of the buffer.

(TMAO)[175, 215, 214, 230] to stabilize the folded structure at ambient temperature. Hence, high-pressure fluorescence measurements were done in Tris 50 mM, NaCl 150 mM, TMAO 750 mM, at various urea concentrations. In the special case of $Nank4-7^{*}$, it is particularly difficult to use it with fluorescence under pressure. Indeed, even if we bought a 99,99% pure TMAO powder, this TMAO conserves some impurities. And as we have to put our protein in a buffer containing 750 mM of TMAO, we have 0.01% of impurity, so a final concentration of impurities of 75 μM, which is $\approx$ 2 times the concentration of the protein (40 μM). It is evident that the fluorescent impurities are seen in the fluorescent assays. In figure 4.1 A, we can observe both for atmospheric pressure (in blue) and 3 kbar (in red), at 1.5M of urea, that at wavelengths above $\approx$ 375 nm the emission spectrum is modified by the background (in black). None of the filtration processes used have managed to decrease the signal contribution of the buffer impurities. In figure 4.1 B, we checked that the tryptophan emission spectra were standard while subtracting the buffer signal, and this confirms that the protein tryptophan emission is not natively perturbed for $Nank4-7^{*}$ but just because of the TMAO impurities.

However, our setup can't permit us to monitor the emission of light from both sample and buffer at all the pressures at the same time, so we were not able to check precisely if this buffer contribution changes with pressure. Hence, it was impossible for us to make an exact correction for all the equilibrium spectra at each pressure. Finally, the only rational way to compute the spectral mass of the tryptophan emission spectrum was to define the limits of the center of mass calculation on the zone where no change comes from the buffer, so from 320 to 370 nm, without subtracting the buffer signal. In this manner, we can be sure of the real change of the tryptophan environment, as represented in figure 4.1.

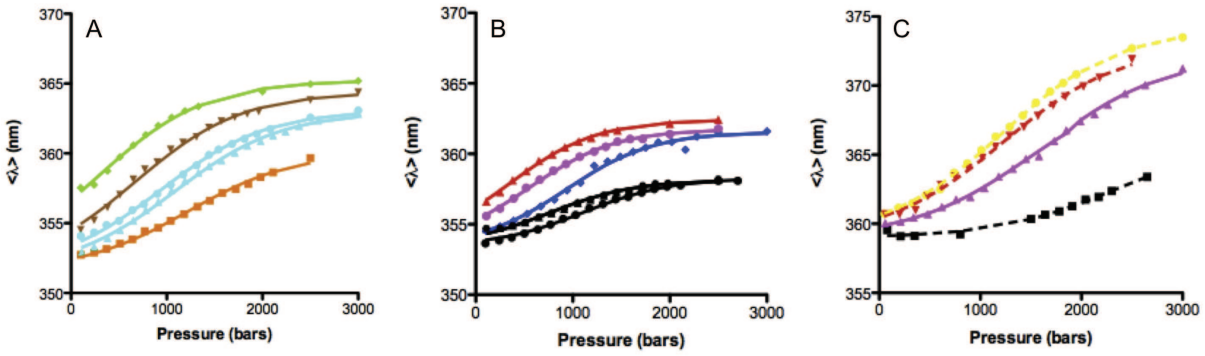

Figure 4.2: Center of the spectral mass of $Nank2-7^\star$, $Nank3-7^\star$ and $Nank4-7^\star$ at 20°C, depending on urea concentration and pressure (from 0 to 3 kbar). A, B, C are respectively at $Nank2-7^\star$, $Nank3-7^\star$ and $Nank4-7^\star$. The urea concentrations used are 1M (dotted black), 1.1M (black), 1.3M (blue), 1.5M (purple), 1.7M (red), 1.75M (dotted red), 1.8M (yellow), 1.85(dotted yellow),1.9M (orange), 2.1M (cyan), 2.3M (brown) 2.4M (green), and 2.5M (purple).

Pressure-induced unfolding of these constructs was monitored by intrinsic fluorescence of the tryptophan as for the full-length $Nank1-7^\star$, at 20°C, for several urea concentrations. For both $Nank2-7^\star$ and $Nank3-7^\star$, the center of spectral mass of each tryptophan emission spectrum was calculated from 320 to 450 nm, like for $Nank1-7^\star$. We can easily follow the maximum of the tryptophan emission peak (which is not modified by the buffer), but the first plateau (folded state) can't be compared to that of $Nank4-7^\star$ because it wasn't calculated on the same wavelength range.

The plots of the center of spectral mass versus pressure of these three mutants are shown in figure 4.2. We observe that as for $Nank1-7^\star$ the unfolding transition curves present distinct high pressure plateau values, depending on the urea concentrations used. Hence, we can hypothesize that $Nank2-7^\star$, $Nank3-7^\star$ and $Nank4-7^\star$ have a high plasticity of the unfolded state, like $Nank1-7^\star$.

For each mutant, data were analyzed as previously described for $Nank1-7^\star$, with floating ΔG^f_{eq} and linked ΔV^f_{eq} between the different urea concentrations used for each mutant. Results are summarized in table 4.1. We can observe that both $Nank2-7^\star$ and $Nank3-7^\star$ have an approximately identical value of ΔV^f_{eq}, and surprisingly only the $Nank4-7^\star$ deletion construct exhibited a ΔV^f_{eq} significantly smaller than $Nank1-7^\star$. The Gibbs free energy difference between the two states upon folding without urea (ΔG^{f0}_{eq}) was calculated by linear regression of the ΔG^f_{eq} values obtained at each urea concentration, as shown in figure 4.3. A summary of ΔG^{f0}_{eq} (with standard deviation errors) and averaged ΔV^f_{eq} is provided in table 4.2. The ΔG^u_{eq} values obtained from extrapolation of the atmospheric pressure values at each urea concentration to zero urea compared reasonably well within the error with those reported previously from urea melts, as shown in figure 4.4 [176].

Thus, the ΔG^{f0}_{eq} obtained from fluorescence of $Nank2-7^\star$ and $Nank3-7^\star$ are respectively 5.05 ± 0.04 and 1.7 ± 0.23 [176]. Here we can compare and see that even if we don't find the exact same value, we are quite close to them. We can estimate that

	$Nank2-7^{\star}$		$Nank3-7^{\star}$		$Nank4-7^{\star}$	
Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	ΔG^f_{eq}	ΔV^f_{eq}	ΔG^f_{eq}	ΔV^f_{eq}
1.0	–	–	–	–	1.94	39.4 ± 4.1
1.1	–	–	1.2	50.62 ± 5.1	–	–
1.3	–	–	1.22	50.62 ± 5.4	–	–
1.5	–	–	0.71	50.62 ± 5.6	1.0.8	39.4 ± 4.3
1.7	–	–	0.46	50.62 ± 6.7	–	–
1.75	–	–	–	–	0.84	39.4 ± 5.7
1.9	1.7	54.58 ± 5	–	–	–	–
2.1	1.52	54.58 ± 4.6	–	–	–	–
2.3	0.94	54.58 ± 3.7	–	–	–	–
2.4	0.4	54.58 ± 5.5	–	–	–	–

Table 4.1: Thermodynamic variables of $Nank2-7^{\star}$, $Nank3-7^{\star}$ and $Nank4-7^{\star}$, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$ and $\Delta V^f_{eq}(mL.mol^{-1})$ are indicated for all the urea concentrations (M).

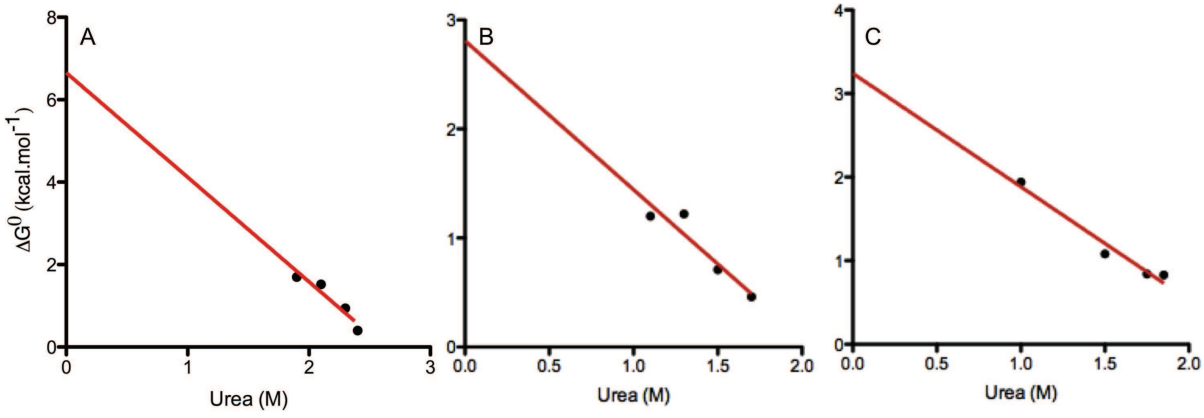

Figure 4.3: Linear regression of ΔG^f_{eq} of $Nank2-7^{\star}$, $Nank3-7^{\star}$ and $Nank4-7^{\star}$ at 20°C

Team of Doug Barrick CD			Team of Doug Barrick Fluorescence			HP Fluorescence		
Construct	ΔG°u (kcal/mol)	m (kcal/l/mol)	Construct	ΔG°u (kcal/mol)	m (kcal/l/mol)	Construct	ΔG°u (kcal/mol)	m (kcal/l/mol)
1-7	6.65	2.85	1-7	6.78	2.91	1-7	7 ± 1.2	2.6 ± 0.5
1-6	2.85	1.76	1-6	3.01	1.85	1-6	2.8 ± 0.1	1.2 ± 0.1
1-5	2.69	1.73	1-5	2.84	1.67	1-5	2.7 ± 0.2	1.2 ± 0.1
3-7	1.81	1.58	3-7	1.70	1.67	3-7	2.8 ± 0.5	1.4 ± 0.4
2-7	4.96	2.36	2-7	5.05	2.44	2-7	6.7 ± 1.3	2.5 ± 0.6

Figure 4.4: Thermodynamic parameters for folding of deletion constructs of the Notch ankyrin domain. Based on [176] and our experimental results.

Protein	ΔG^{f0}_{eq}	Averaged ΔV^f_{eq}
$Nank2-7^{\star}$	6.6 ± 1.2	54.58 ± 4.7
$Nank3-7^{\star}$	2.8 ± 0.5	50.62 ± 5.7
$Nank4-7^{\star}$	3.2 ± 0.3	39.4 ± 4.7

Table 4.2: Thermodynamic variables of $Nank2-7^{\star}$, $Nank3-7^{\star}$ and $Nank4-7^{\star}$ at 20°C. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$ are averaged.

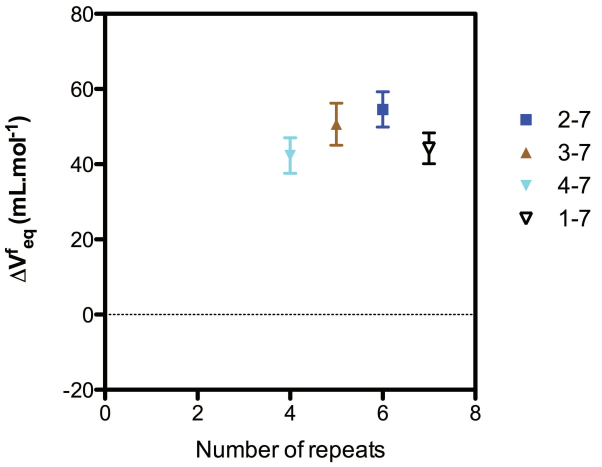

Figure 4.5: Evolutions of ΔV^f_{eq} at 20°C depending on the number of repeats for $Nank2-7^\star$, $Nank3-7^\star$, $Nank4-7^\star$, NWC and NW.

we are in the standard error (± 1), particularly given the long extrapolation of only a few points (from 0 to 2.4M and 1.75M of urea respectively). Hence, we can be confident in the results obtained from the treatment of the pressure-induced transition curves.

Short discussion

In a plot of the ΔV^f_{eq} versus the number of repeats, provided in figure 4.5, we see that there is no general dependence for the ΔV^f_{eq} on the number of repeats.

C-terminal deletion constructs

We know that the first motif seems to be highly different from the others in terms of structure and changes the overall cooperativity by a "capping" process, we want now to explore some mutant proteins with different number of motifs, but always keeping the first one in order to estimate the implication of size and sequence in the ΔV^f_{eq} determination at 20°C, and to compare it with the previously studied proteins.

As we can observe for $Nank1-6^\star$ and $Nank1-5^\star$ in the plot of the mass center of the tryptophan emission spectrum versus pressure, as shown respectively in figure 4.6 A and B, we still have the same pattern than usual for both of them, with various ending plateaus depending on urea concentrations used. The folded state plateau is well defined, and the overall change in the mass centers is similar in amplitude to that of the previous mutants. The quality of experimental measurements allowed us to obtain data with a minimum of errors at 20°C.

The global analysis of all these data was easy because of the good accuracy of the fits, and results are reported in tables 4.3 for $Nank1-6^\star$ and 4.4 for $Nank1-5^\star$.

We can be confident in the accuracy of these values, as the $\Delta G^f_{eq}0$ obtained by linear regression of the ΔG^f_{eq} at each urea concentrations (figure 4.7) are respectively of 2.68 ± 0.6 kcal/mol for $Nank1-5^\star$ and 2.76 ± 0.14 kcal/mol for $Nank1-6^\star$, compared

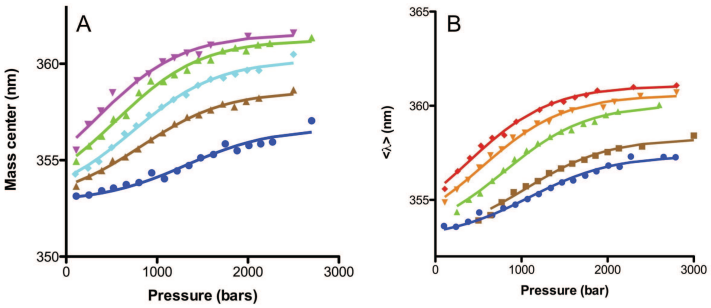

Figure 4.6: Evolution of the mass center of tryptophan emission spectra of $Nank1-6^{\star}$ at 20°C. The concentrations of urea used are : 1M (blue), 1.4M (brown), 1.6M (cyan), 1.8M (green) and 2M (purple).

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}
1	1.628	48.14 ± 8
1.4	1.01	48.14 ± 3.9
1.6	0.9	48.14 ± 3.5
1.8	0.67	48.14 ± 2.1
2	0.41	48.14 ± 9.1

Table 4.3: Thermodynamic variables of $Nank1-6^{\star}$, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$ are indicated for all the urea concentrations (M).

Urea (M)	ΔG	ΔV^f_{eq}
1.1	1.32	47.01 ± 8
1.3	1.27	47.01 ± 9
1.5	0.96	47.01 ± 10.5
1.7	0.65	47.01 ± 9
1.9	0.45	47.01 ± 8.5

Table 4.4: Thermodynamic variables of $Nank1-5^{\star}$, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$ are indicated for all the urea concentrations (M).

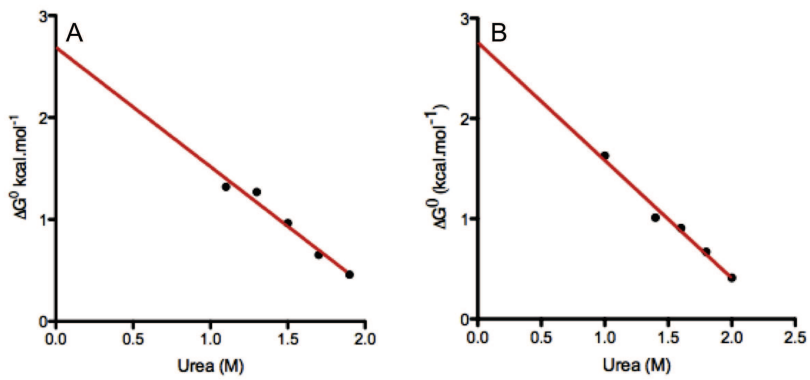

Figure 4.7: Linear regression of ΔG^f_{eq} for $Nank1-5^\star$ and $Nank1-6^\star$

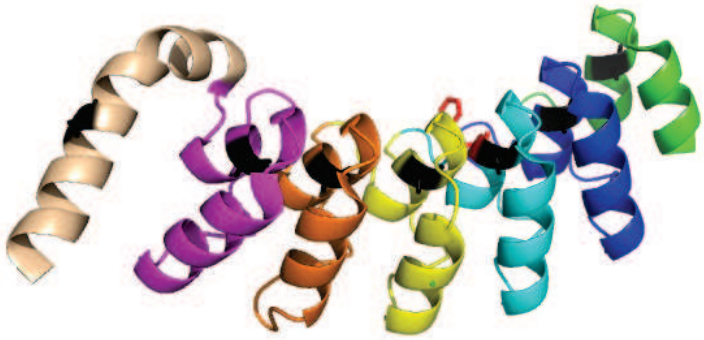

Figure 4.8: Conserved alanines allong the $Nank1-7^\star$ structure. The first motif is here colored in pink, the seventh in green. In black are the repeated conserved alanines in each repeats.

to 2.69 and 2.85 kcal/mol in literature [176].

The value of the ΔV^f_{eq} for $Nank1-5^\star$ and $Nank1-6^\star$ at 20°C are respectively 47.01 ± 9 and 48.14 ± 5.32 $mL.mol^{-1}$ at 20°C, again not significantly different from the full-length $Nank1-7^\star$ construct.

Sequence mutated constructs : NAG2 and NAG5

Next, we intended to change the stability between repeats, through the use of alanine to glycine mutants : between repeat 1 and 2 in NAG2 and between repeat 4 and 5 in NAG5. The alanines changed into a glycine are at highly conserved positions, as shown in figure 4.8.

The mutation in NAG2 strongly destabilizes the second motif, so in this mutant, repeats 1 and 2 are unfolded. In the case of NAG5, the mutation destabilizes the 5th motif but its fold is maintained.

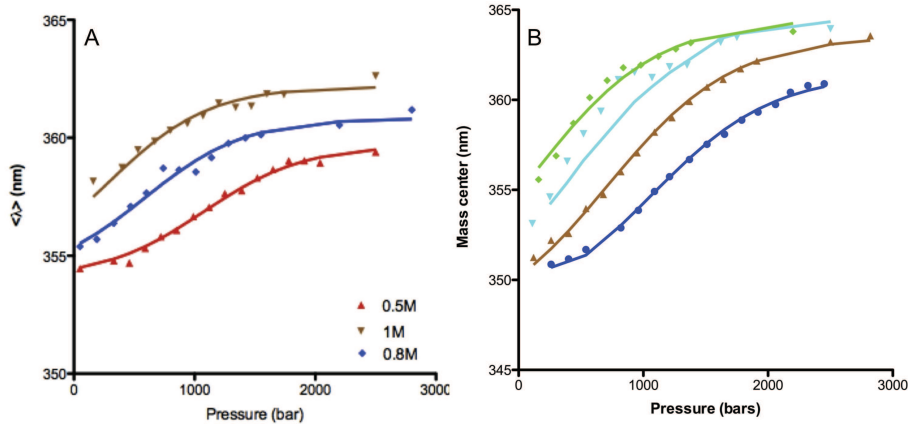

Figure 4.9: Center of spectral mass for *NankNAG2* (A) and *NAG5* (B), obtained by high-pressure fluorescence.

Urea (M)	ΔG	ΔV^f_{eq}
0.5	1.09	54.36 ± 5
0.8	0.41	54.36 ± 5
1.0	0.00	54.36 ± 5

Table 4.5: Thermodynamic variables of Nank NAG2, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

As shown in figure 4.9, the center of mass of the tryptophan emission spectra for both NAG2 (A) and NAG5 (B), obtained in Tris 50 mM, NaCl 150 mM, follow the same pattern than previously described for the other variants, defining multiple ending plateaus, and a unique folded state plateau. Calculations were done the same way than previously described, and data are respectively summarized in tables 4.5 and 4.6. The value of the ΔV^f_{eq} is of 54.36 ± 5 $mL.mol^{-1}$ at 20°C for NAG2 and 46.95 ± 5 mL/mol for NAG5. The ΔV^f_{eq} for NAG2 is significantly larger than that found for the $Nank1-7^{\star}$ construct, while that for NAG5 is found to be equivalent.

Short discussion

If we plot the ΔV^f_{eq} of all the constructs with the first repeat at 20°C ($Nank1-7^{\star}$, $Nank1-6^{\star}$, $Nank1-5^{\star}$, NAG2 and NAG5) as a function of the number of repeat,

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}
1.1	1.29	46.95 ± 5
1.3	0.95	46.95 ± 5
1.5	0.37	46.95 ± 5

Table 4.6: Thermodynamic variables of NAG5, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$ are indicated for all the urea concentrations (M).

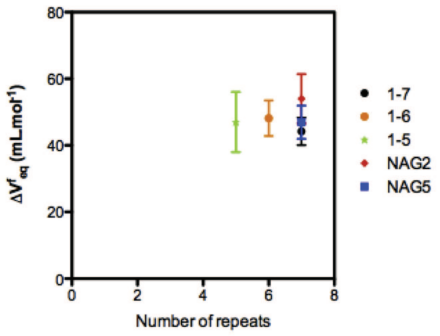

Figure 4.10: Evolution of ΔV^f_{eq} depending on the number of repeats for proteins with the first repeat

Consensus	xtx**OULHΦAΦ**xxttxt**ΦΦ**tx**LL**txtxtxxxxxx bb aaaaaaaaa aaaaaaaaaa aa
Repeat 1	cgl**TPL**m**IAA**vrggzq**VI**sd**LL**aqgaelnatmdk
Repeat 2	tge**Ts****LHLAA**rfarad**AA**kr**LL**dagadancqdn
Repeat 3	tgr**TPLHAAV**aadamg**VF**qi**LL**rnratnlnarmh
Repeat 4	dgt**TPLiLAA**rlaieg**MV**ed**L**itadadinaadn
Repeat 5	sgk**TALHWAA**avnnte**AV**ni**LL**mhhanrdaqdd
Repeat 6	kde**TPL**f**LAA**regsye**AC**ka**LL**dnfanreitdh
Repeat 7	mdrl**P**rd**VA**serl.hhd**I**vr**LL**dehvprspqml
Repeat N	dgn**TPLH**n**AA**knghaee**V**kk**LL**skgadvnarsk
Repeat W	dgw**TPLHLAA**knghae**IV**kl**LL**akgadvnarsk
Repeat C	dgn**TP**e**HLA**kknghhe**IV**kl**L**dakgadvnarsk

Figure 4.11: Sequences of all repeats used.

as shown in figure 4.10, we can easily notice that their ΔV^f_{eq} seem similar and that no clear dependence between size and ΔV^f_{eq} appear. Hence it depicts a pattern quite similar to the previous results with N-terminal deletion constructs.

Small constructs of two and three repeats

We also examined two smaller concensus repeat constructs, NW and NWC. They are composed of two (NW) and three (NWC) repeats, and do not contain the sequence of the first repeat, like the previous studied mutants. Sequences are provided in figure 4.11. These constructs are N-capped in order to stabilize them.

Experiments on NW (figure 4.12 A) and NWC (figure 4.12 B) were performed as usual at 20°C, at several GuHCl concentrations, as urea wasn't sufficient enough to perturb their stability for pressure to act. As we can see, the evolution of the center of the spectral mass with pressure is very small, indicative of a very small ΔV^f_{eq}. A very simple analysis was done, linking the ΔV^f_{eq} of each construct over all GuHCl

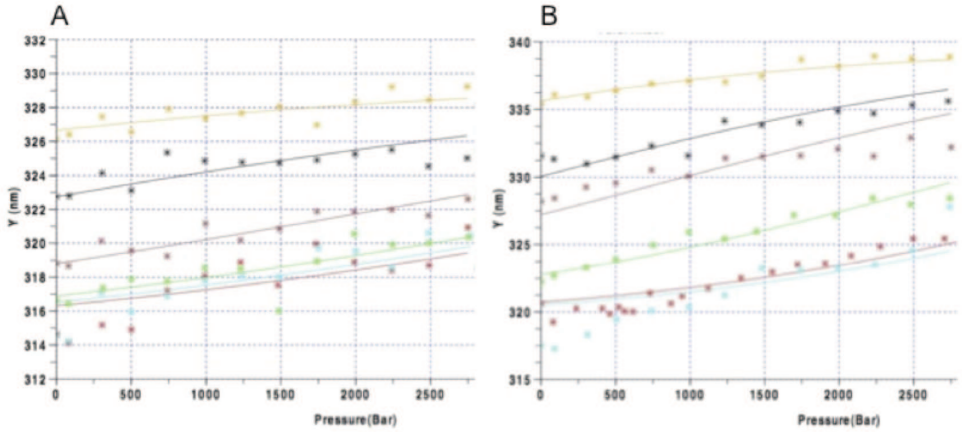

Figure 4.12: Evolution of the mass center of the spectra for NW and NWC. A : NW, and B : NWC.

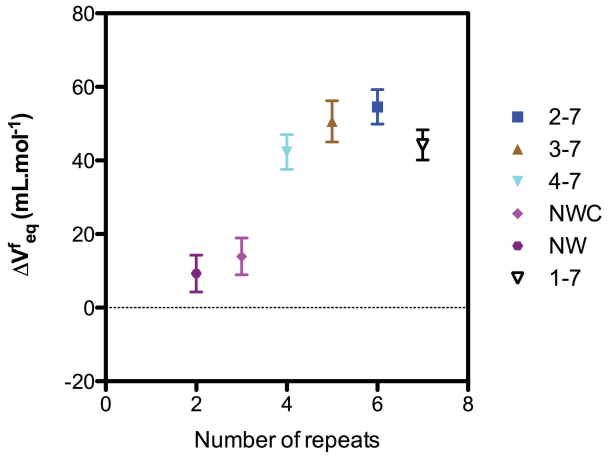

Figure 4.13: Evolution of ΔV^f_{eq} at 20°C depending on the number of repeats, including $Nank1-7^\star$, $Nank2-7^\star$, $Nank3-7^\star$, $Nank4-7^\star$, NWC and NW.

concentrations, and yielded ΔV^f_{eq} values of 9.5 and 13.52 $mL.mol^{-1}$, respectively for NW and NWC. Even if these constructs were not really easy to unfold with pressure and to compare with the other ones, we use them as an estimation of the ΔV^f_{eq} for small number of repeats. By plotting these results with previous ones for N-terminal deletion constructs, in figure 4.13, we see a general dependence for the ΔV^f_{eq} on the number of repeats, although it is not a smooth one.

4.1.2 Kinetic experiments at 20°C

We performed pressure jump relaxation kinetic measurements of the pressure-jumps in order to determine the activation volumes associated with the folding and unfolding transitions. As described in material and methods, single exponential fits of the relaxation profiles yielded pressure dependent relaxation times, τ(p). Analysis of the $ln(\tau)$ with pressure, as shown in figure 4.14, allows the calculation of the activation volumes for folding and the folding rate constant at atmospheric pressure. We constrain

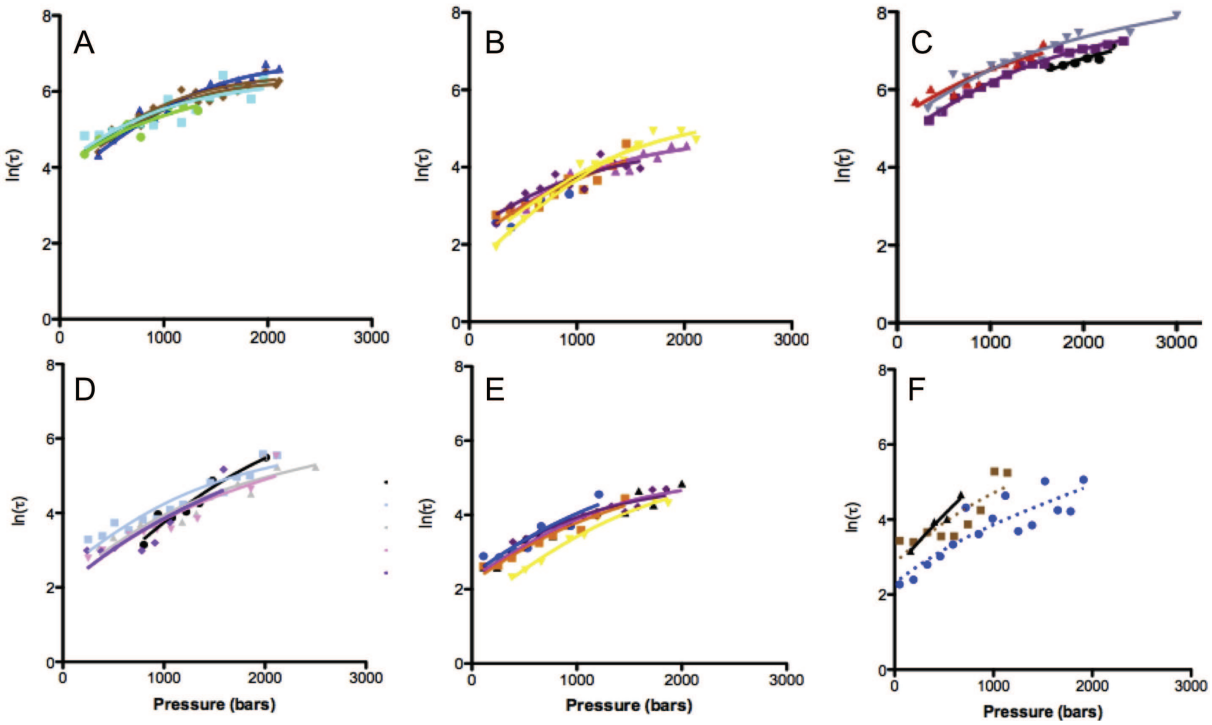

Figure 4.14: Natural logarithm of τ of the mutants at 20°C, depending on pressure and urea concentration. A : *Nank*2−7*, B : *Nank*3−7*, C : *Nank*4−7*, D : *Nank*1−6*, E : *Nank*1−5*, F : NAG2. Urea concentrations are 0.5M (dotted blue),0.8M (dotted brown), 1M (black), 1.1M (yellow), 1.3M (magenta), 1.4M (sky blue), 1.5M (purple), 1.6M (light gray), 1.7M (orange), 1.75M (red), 1.8M (pink), 1.85M (blue grey), 1.9M (blue), 2M (violet), 2.1M (brown), 2.3M (cyan), 2.4M (green)

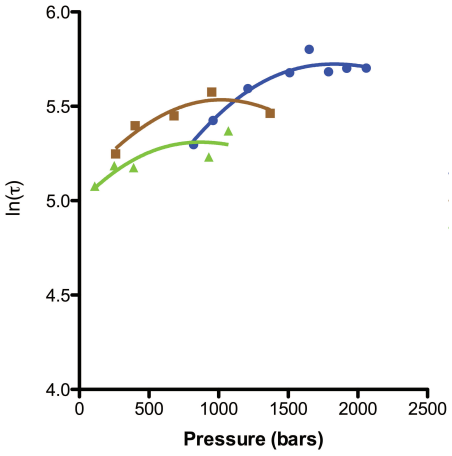

Figure 4.15: Evolution of the relaxation times for Nank NAG 5. Urea concentrations used are 1.1M (blue), 1.3M (brown), 1.5M (cyan) and 1.7M (green).

the volumes of the unfolding arm with the equilibrium constant and volume change. The $ln(\tau)$ studies depending on pressure, we know that the pattern is similar to the $Nank1-7^{\star}$ one. The k_f^0 is the only accessible data, so it's still the only rate constant that is really changed by pressure, and the k_u^0 remains quite stable on the pressure range.

NAG5 is the last mutant with an alanine to glycine mutation and the analysis of its relaxation time logarithm curves depending on pressure is shown in figure 4.15.

Results of the calculations are provided in table 4.7.

In the case of $Nank4-7^{\star}$, results at 1M urea are subjects to discussion because the $ln(\tau)$ are defined only by 6 data points between 1500 and 2500 bars, so we don't have the complete and precise knowledge of the k_f and k_u, increasing the possible errors of the fits. So we prefer to take into account only the values obtained from 1.5 to 1.85M of urea which are better defined. The average $\Delta V_f^{\#}$ is +46.12 $mL.mol^{-1}$, so bigger than the ΔV_{eq}^{f}. The $\Delta V_u^{\#}$ is +6.93 $mL.mol^{-1}$, positive. Values of all the averaged ΔV_{eq}^{f}, $\Delta V_u^{\#}$ and $\Delta V_f^{\#}$ are provided in table 4.8.

Short discussion

With the activation volumes determined at 20°C for all our proteins, we are now able to have a volumetric diagram of all these proteins at 20°C, as represented in figure 4.16. We have seen earlier that the ΔV_{eq}^{f} dependence with size seemed to be of two kind, depending on the constructs chosen. In figure 4.16 B, we can observe that between $Nank1-7^{\star}$, $Nank1-6^{\star}$ and $Nank1-5^{\star}$, the volume of the folded state remains quite constant. Interestingly, for all of them, the transition state ensemble (TSE) volume is larger than the folded state volume, whereas at this temperature the transition state volume lies below that of the folded state for full-length $Nank1-7^{\star}$. We can observe that deleting repeat 7 has the largest effect on the volume of the TSE, which is shifted to a much higher value, indicating a large expansion of the TSE. And deleting the sixth repeat, the TSE decreases its expansion, but is still more expanded than for $Nank1-7^{\star}$.

In contrast, while looking at $Nank1-7^{\star}$, $Nank2-7^{\star}$, $Nank3-7^{\star}$ and $Nank4-7^{\star}$ together, in figure 4.16 C, we observe a completely inverted pattern. Indeed, while deleting the first motif, the volume both the folded and transition states increases (this latter above that of $Nank1-7^{\star}$), and more deletions result in lower folded state volumes but equivalent TSE volumes. Here it seems that while removing the capping of the first motif, we make the protein expand more in space and deleting motifs diminishes the volume of the folded state with the number of repeats, while removing the capping of the seventh repeat and deleting C-terminal motifs do not change anything on the volume of the folded state but makes the TSE increase its expansion a lot. Finally, in figure 4.16 D, we observe that in the case of NAG2 and NAG5, even if the number of repeats is the same than for $Nank1-7^{\star}$, the single point mutation in the second motif

Nank 2-7*

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$	k^0_f	k^0_u
1.9	1.7	54.58 ± 5.0	56.29 ± 2.5	1.71 ± 2.5	0.026	0.00144
2.1	1.52	54.58 ± 4.6	54.45 ± 2.3	-0.135 ± 2.3	0.018	0.00136
2.3	0.94	54.58 ± 3.7	59.81 ± 3.5	5.23 ± 3.5	0.014	0.0029
2.4	0.4	54.58 ± 5.5	67.23 ± 4.2	12.65 ± 4.2	0.012	0.0064

Nank 3-7*

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$	k^0_f	k^0_u
1.1	1.2	50.62 ± 5.1	66.73 ± 2.5	16.11 ± 2.5	0.168	0.0216
1.3	1.22	50.62 ± 5.4	54.6 ± 1.9	4 ± 1.9	0.114	0.0141
1.5	0.71	50.62 ± 5.6	57.89 ± 4.2	7.27 ± 4.2	0.075	0.0222
1.7	0.46	50.62 ± 6.7	68.49 ± 5	17.87 ± 5	0.090	0.0413

Nank 4-7*

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$	k^0_f	k^0_u
1.0	1.94	39.44 ± 4.1	36.52 ± 4.7	-2.92 ± 4.7	0.0116	0.0004
1.5	1.08	39.44 ± 4.3	46.12 ± 0.74	6.68 ± 0.74	0.0076	0,0012
1.75	0.84	39.44 ± 5.7	44.09 ± 3.3	4.65 ± 3.3	0.0041	0,0009
1.85	0.83	39.44 ± 5.7	48.91 ± 1.6	9.47 ± 1.6	0.005	0.0012

Nank 1-6*

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$	k^0_f	k^0_u
1	1.628	48.14 ± 8	65.8 ± 5.5	17.6 ± 5.5	0.2584	0.0159
1.4	1.01	48.14 ± 3.9	58.8 ± 2.8	10.66 ± 2.8	0.0750	0.0125
1.6	0.9	48.14 ± 3.5	61.7 ± 3.1	13.56 ± 3.1	0.103	0.0218
1.8	0.67	48.14 ± 2.1	66.44 ± 4.4	18.2 ± 4.4	0.107	0.0379
2	0.41	48.14 ± 9.1	72.48 ± 10.3	24.2 ± 10.3	0.0947	0.0469

Nank 1-5*

Urea (M)	ΔG	ΔV^f_{eq}	$\Delta V^\#_u$	$\Delta V^\#_f$
1.1	1.32	47.01 ± 8	10.86 ± 2	57.87 ± 2
1.3	1.27	47.01 ± 9	3.09 ± 1.8	50.1 ± 1.8
1.5	0.96	47.01 ± 10.5	4.6 ± 2.1	51.61 ± 2.1
1.7	0.65	47.01 ± 9	13.95 ± 2.6	60.96 ± 2.6
1.9	0.45	47.01 ± 8.5	17.37 ± 6.6	64.38 ± 6.6

Nag 2

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$	k^0_f	k^0_u
0.5	1.09	54.36 ± 1	77.96 ± 4.6	+23.6 ± 4.6	0.050	0.007758
0.8	0.41	54.36 ± 5	77.66 ± 10.1	+23.3 ± 10.1	0.038	0.018854
1	0.00	54.36 ± 5.5	106.1 ± 10.6	+51.74 ± 10.6	0.035	0.035000

Nag 5

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$
1.1	1.29	46.95 ± 4.5	37.05 ± 0.8	-9.9 ± 0.8
1.3	0.95	46.95 ± 5	27.42 ± 1.1	-19.53 ± 1.1
1.5	0.37	46.95 ± 6	34.77 ± 1.9	-12.18 ± 1.9

Table 4.7: Thermodynamic variables of all the mutants at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^\#_f(mL.mol^{-1})$,$\Delta V^\#_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

Protein	ΔV^f_{eq}	$\Delta V^\#_f$	$\Delta V^\#_u$
$Nank2-7^\star$	54.58 ± 4.7	59.44 ± 3.1	4.86 ± 3.1
$Nank3-7^\star$	50.62 ± 5.7	61.9 ± 3.4	11.28 ± 3.4
$Nank4-7^\star$	39.44 ± 4.9	46.12 ± 4.7	6.93 ± 4.7
$Nank1-6^\star$	48.14 ± 5.3	65.03 ± 5.2	16.89 ± 5.2
$Nank1-5^\star$	47.01 ± 9	56.98 ± 3.0	9.97 ± 3.0
NAG2	54.36 ± 3.8	77.81 ± 8.4	23.45 ± 8.4
NAG5	46.95 ± 5.1	33.08 ± 1.2	-13.87 ± 1.2

Table 4.8: Volumes of each mutant, at 20°C. ΔV^f_{eq}, $\Delta V^\#_u$ and $\Delta V^\#_f$ are in $(mL.mol^{-1})$.

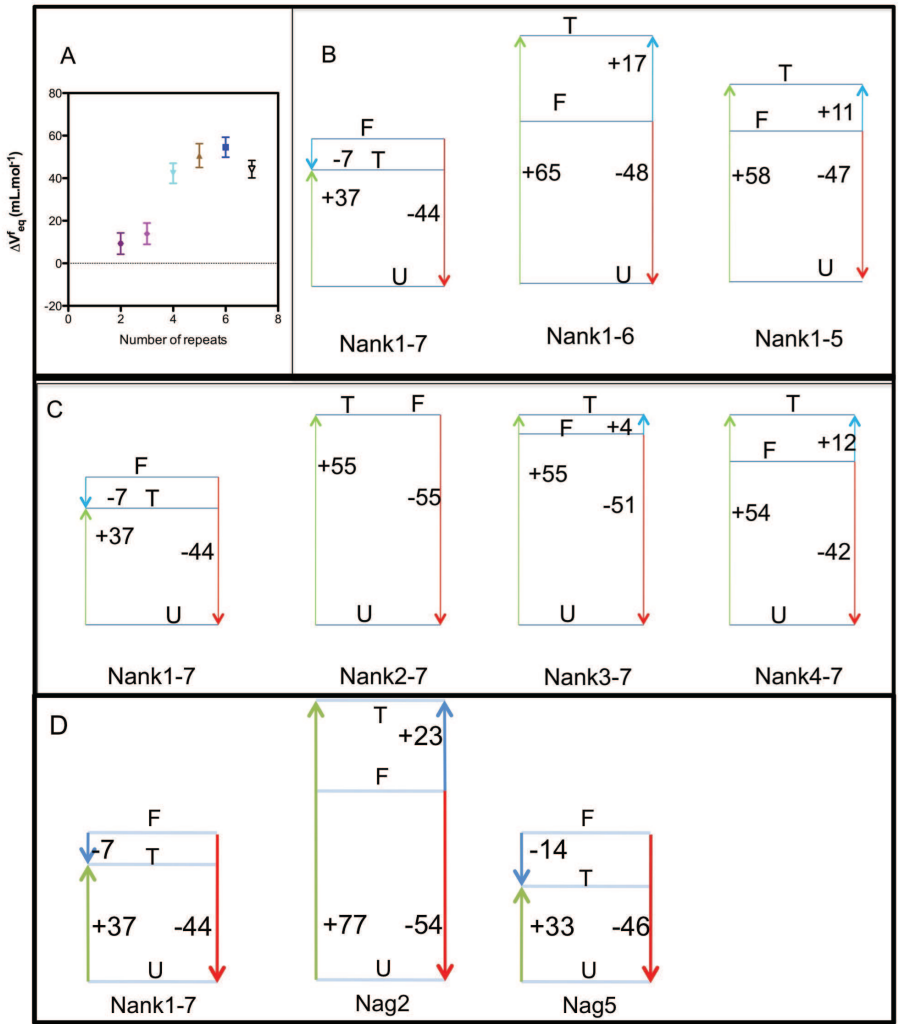

Figure 4.16: Volume changes obtained at 20°C for the Nank variants and the NW and NWC consensus constructs. A) Linear regression of the absolute value of ΔV^u_{eq} , the volume change of unfolding for the two consensus constructs, NW and NWC, as well as the N-terminal deletion constructs Nank2-7, Nank3-7 and Nank4-7 as a function of the number of repeats in the construct. The absolute values of ΔV^u_{eq} for the Nank1-7, Nank1-6 and Nank1-5 constructs are also shown for comparison., B) Volume diagrams for the Nank C-terminal deletion constructs, Nank1-6 and Nank1-5, C) for the N-terminal deletion constructs Nank2-7, Nank3-7 and Nank4-7 and D) for Nank1-7, Nag2 and Nag5. In all figures the previously reported diagram of the Nank1-7 full-length construct is given.

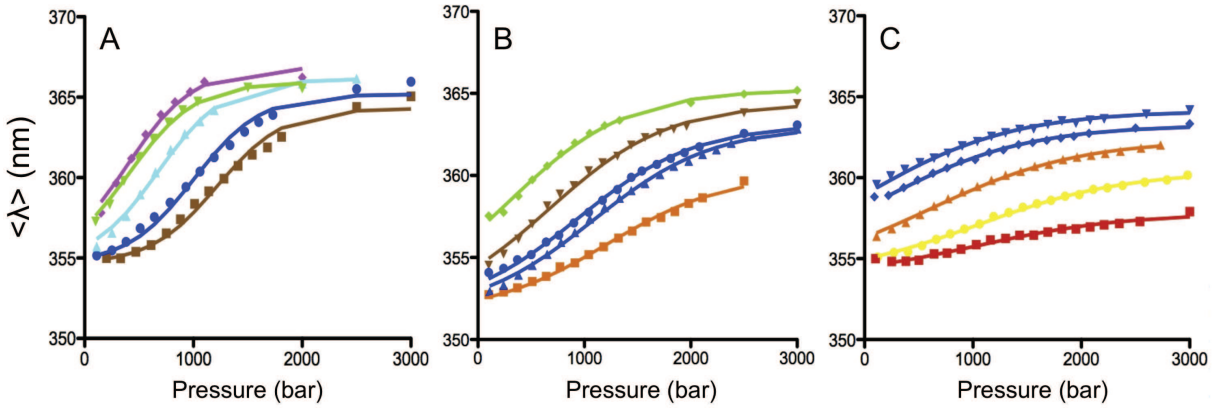

Figure 4.17: Center of the spectral mass of $Nank2-7^{\star}$ depending on the pressure applied from 0 to 3 kbar, at different temperature and urea concentration. A, B, C are respectively at 12, 20 and 28°C. The urea concentrations used are 1.7M (red), 1.8M (yellow), 1.9M (orange), 2.1M (blue), 2.3M (brown) 2.4M (green), and 2.5M (purple).

of NAG2 make increase both the folded and transition state ensemble volumes, while the mutation in the fifth repeat doesn't change the folded state volume and make that of the TSE a little lower.This makes us think that the TSE, which is always similar or superior to the folded state in terms of volume is really important for our comprehension of the ΔV^u_{eq}.

4.2 Study of the ΔV^f_{eq} dependence with temperature

4.2.1 At the equilibrium

$Nank2-7^{\star}$

To further explore the volumetric properties of the deletion variants, we examined their expansivities by pressure unfolding at different temperatures. We examined $Nank2-7^{\star}$ and $Nank3-7^{\star}$ at three different temperatures : 12°C, 20°C and 28°C. These temperatures were shown ideal to observe the evolution of the ΔVs depending on temperature in the case the complete $Nank1-7^{\star}$ protein, because of on the one hand a high signal to noise ratio for each one, and on the other hand an important and accessible effect of temperature on ΔV^f_{eq} in this range of temperatures. Moreover, high-pressure fluorescence experiments are highly time-consuming, and as we need for each protein to do a minimum of four or five experiments at each temperature, multiplying the number of temperatures reached would have been leading to more months of work, for nearly identical results.

The high-pressure fluorescence experiments $Nank2-7^{\star}$ show a pattern of unfolding transition curves depending on pressure and urea which is highly similar to $Nank1-7^{\star}$, as shown in figure 4.17 A, B and C. These unfolding transition curves at each temperature and urea concentrations show several final plateaus for unfolding, so we

	12°C		20°C		28°C	
Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	ΔG^f_{eq}	ΔV^f_{eq}	ΔG^f_{eq}	ΔV^f_{eq}
1.7	–	–	–	–	0.9	33.8 ± 3.1
1.8	–	–	–	–	0.9	33.8 ± 5
1.9	–	–	1.7	54.58 ± 5	0.52	33.8 ± 5.8
2.1	2.05	72 ± 7.0	1.52	54.58 ± 4.6	0.1	33.8 ± 4.2
2.3	1.24	72 ± 7.4	0.94	54.58 ± 3.7	–	–
2.4	0.75	72 ± 7.5	0.4	54.58 ± 5.5	–	–
2.5	0.71	72 ± 5.0	–	–	–	–

Table 4.9: Thermodynamic variables of $Nank2-7^\star$, at 12, 20 and 28°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$ and $\Delta V^f_{eq}(mL.mol^{-1})$ are indicated for all the urea concentrations (M).

can still think that in this case, as with the $Nank1-7^\star$, the protein has got an important plasticity of the unfolded state, with important urea-induced motions that have an impact on the close environment of the tryptophan. We can observe that while increasing temperature, we use lower urea concentrations to unfold $Nank2-7^\star$ by pressure than for $Nank1-7^\star$, as this protein is a little less stable than $Nank1-7^\star$ as reported during introduction.

It is interesting to note that the slope of the transitions decrease for an identical urea concentration while increasing temperature, as we can observe in figure 4.17 A, B and C for 2.1M of urea (blue curves). This decrease is relevant of the fact that the more we increase temperature, the less application of pressure helps unfolding. So by eye, we can hypothesize the ΔV^f_{eq} of $Nank2-7^\star$ to be lower in magnitude at 28°C than at 12°C.

After treatment of all this denaturation curves depending on pressure and urea concentration, by the exact same method than previously described (floating ΔG^f_{eq} and linked ΔV^f_{eq} between experiments at the same temperature), we determined for each temperature the ΔG^f_{eq} and ΔV^f_{eq}, as reported in table 4.9.

By looking at the ΔV^f_{eq} depending on temperature, as shown in figure 4.18, we can observe that the ΔV^f_{eq} value we obtained separately for each temperature are perfectly aligned. The linear regression used to estimate the correlation between temperature and ΔV^f_{eq} is nice, with a high correlation factor of 0.9974. The estimation of the temperature at which ΔV^f_{eq} is null is 42.4 ± 1.5°C. The slope of this linear regression ($\frac{\delta \Delta V^f_{eq}}{\delta T}$) is important, with a value of -2.38 ± 0.12 $mL.mol^{-1}.^\circ C^{-1}$, arising from an important $\Delta\alpha_f$.

$Nank3-7^\star$

Regarding the $Nank3-7^\star$ unfolding by pressure, we observe that the unfolding transition curves depending on urea and pressure still conserve the similar pattern that previously seen for $Nank1-7^\star$ and $Nank2-7^\star$, as shown in figure 4.19 A, B and

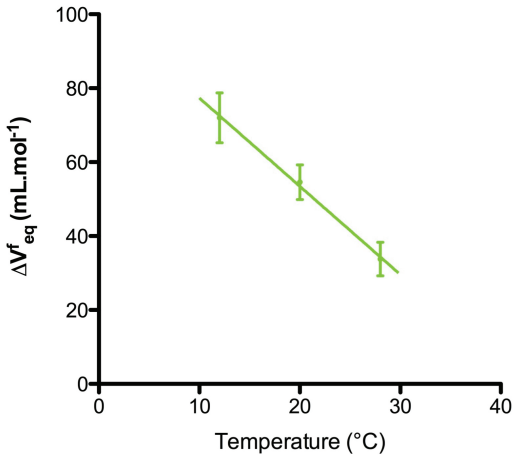

Figure 4.18: Evolution of $Nank2-7^{\star\prime}$s ΔV^f_{eq} with temperature. ΔV^f_{eq} were calculated from pressure-induced unfolding experiments at 12, 20 and 28°C

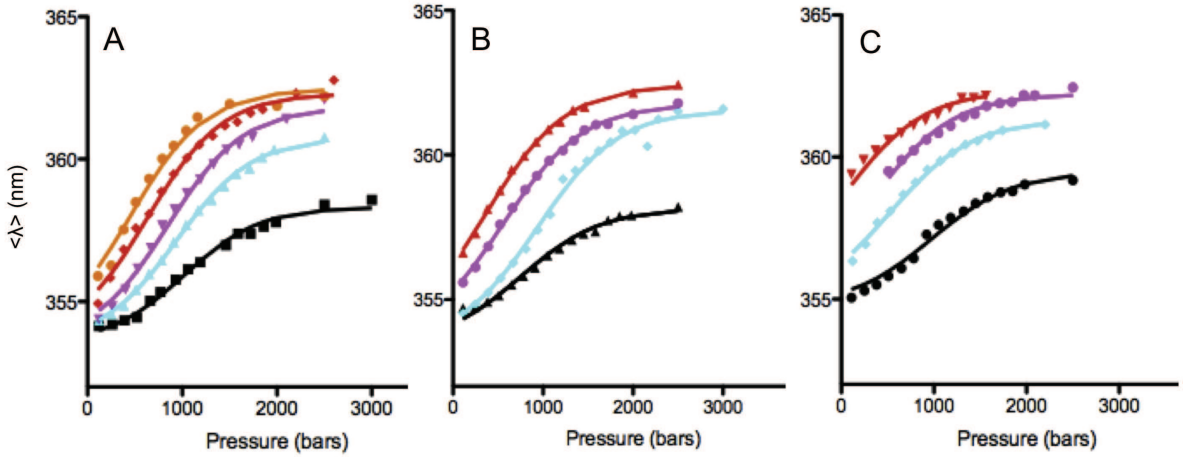

Figure 4.19: Evolution of the mass center of $Nank3-7^{\star}$ at 12, 20 and 28°C. Mass center of the spectra at (A) 12°C, (B) 20°C and (C) 28°C. Colors used are : black for 1.1M urea, blue for 1.3M urea, purple for 1.5M urea, red for 1.7M urea and orange for 1.9M urea.

C. Multiple ending plateaus are reached at different mass centers of the spectra, depending on the urea concentrations used. We can observe for example that 1.3M of urea (blue fits) permits to monitor the majority of the unfolding process at all the temperatures. The analysis of these curves was done by the same way than previously described, and results are summarized in table 4.10. We can observe in figure 4.20 that ΔV^f_{eq} changes with temperature, decreasing from +55.4 $mL.mol^{-1}$ at 12°C to +41.8 $mL.mol^{-1}$ at 28°C, with a slope of -0.84 ± 0.14 $mL.mol^{-1}.°C^{-1}$. The $\Delta\alpha$ of this mutant can be estimated really low, so the two states should have quiet similar expansivities. Thanks to the good linear regression of these data (correlation factor R of 0.9723), we can estimate that this mutant would have a null ΔV^f_{eq} above 78°C.

	12°C		20°C		28°C	
Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	ΔG^f_{eq}	ΔV^f_{eq}	ΔG^f_{eq}	ΔV^f_{eq}
1.1	1.5	55.4 ± 4.6	1.2	50.62 ± 5.1	1.32	41.83 ± 6.5
1.2	–	–	–	–	0.71	41.83 ± 6.8
1.3	1.34	55.4 ± 4	1.22	50.62 ± 5.4	–	–
1.5	1.21	55.4 ± 3.5	0.71	50.62 ± 5.6	0.43	41.83 ± 6.3
1.7	0.89	55.4 ± 7	0.46	50.62 ± 6.7	2.3E-7	41.83 ± 6.5
1.9	0.64	55.4 ± 10	–	–	–	–

Table 4.10: Thermodynamic variables of $Nank3-7^{\star}$, at 12, 20 and 28°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$ are indicated for all urea concentrations used (M).

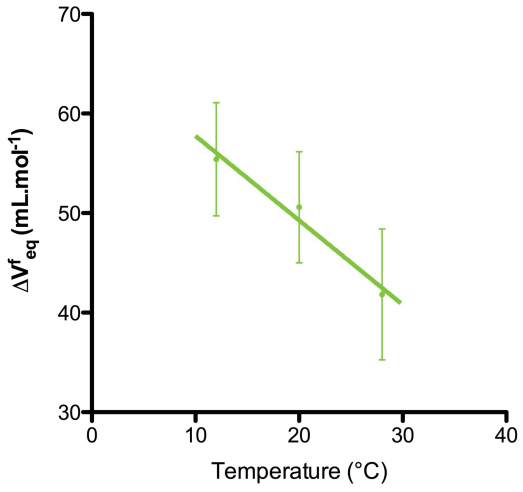

Figure 4.20: Evolution of $Nank3-7^{\star}$'s ΔV^f_{eq} with temperature. ΔV^f_{eq} were calculated from pressure-induced unfolding experiments at 12, 20 and 28°C

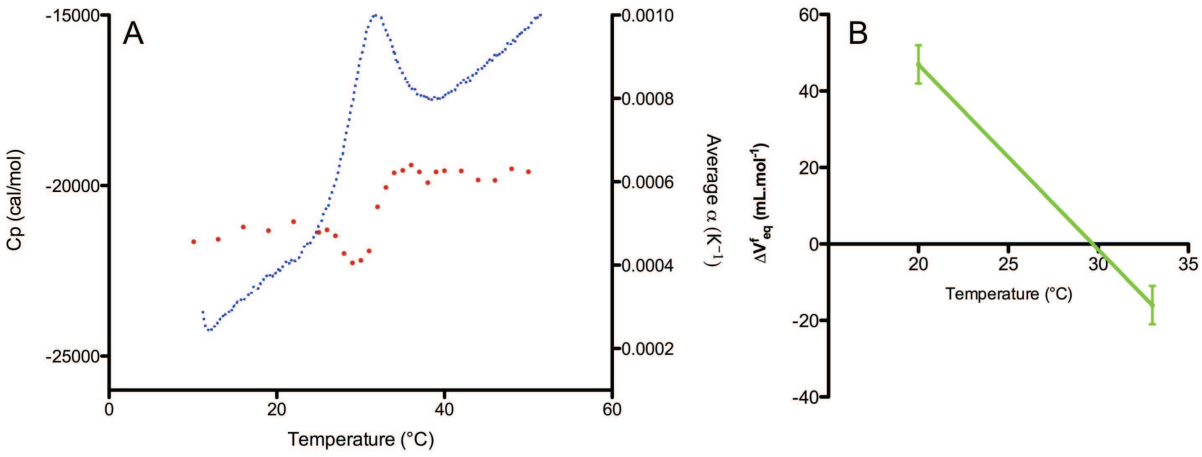

Figure 4.21: DSC and PPC of NAG5 and associated expansivity calculation. A : DSC and PPC plots. Data were collected in Phosphate buffer 50 mM, 150 mM NaCl, 1M urea, from 10 to 60°C. For DSC experiments (in blue), NAG5 was concentrated to 2 mg/mL, and 12 mg/mL for PPC experiment (in red). B : estimation of the slope of ΔV^f_{eq} on temperature.

4.2.2 Sequence mutated constructs : NAG2 and NAG5

NAG5

We performed DSC and PPC on this mutant, as shown in figure 4.21. The DSC is of good quality but we didn't manage to subtract the 1M urea buffer correctly again. The Tm was defined to be at 31°C. The PPC plot shows us that the folded and unfolded expansivities are quite flat and a large negative ΔV^f_{eq} is seen at the Tm, with a calculated value of -16 mL/mol. We see a large $\Delta\alpha$ between the folded and unfolded states expansivities. By using high pressure and PPC values of ΔV^f_{eq}, we can determine the $\frac{\delta \Delta V^f_{eq}}{\delta T}$ slope of the linear regression (data not shown) to be -5.7 mL/mol/°C, as shown in figure 4.21 B, confirming the large $\Delta\alpha$ seen in the PPC experiment.

4.2.3 Mutants with C-terminal deletion

We know that the first motif seems to be highly different from the others in terms of structure and changes the overall cooperativity by a "capping" process, we want now to explore some mutant proteins with different number of motifs, but always the first one in order to estimate the implication of size and sequence in the ΔV^f_{eq} determination at 20°C, and to compare it with the previously studied proteins.

$Nank1-6^*$

To determine the expansivities of the two states and of the ΔV^f_{eq} above 20°C, we performed DSC and PPC experiments, as shown in figure 4.22 A. The explored temperature range was from 10 to 60°C; we used 1M urea in the buffer for both DSC and PPC experiments to avoid aggregation at high temperatures, leading to a shift of the Tm to

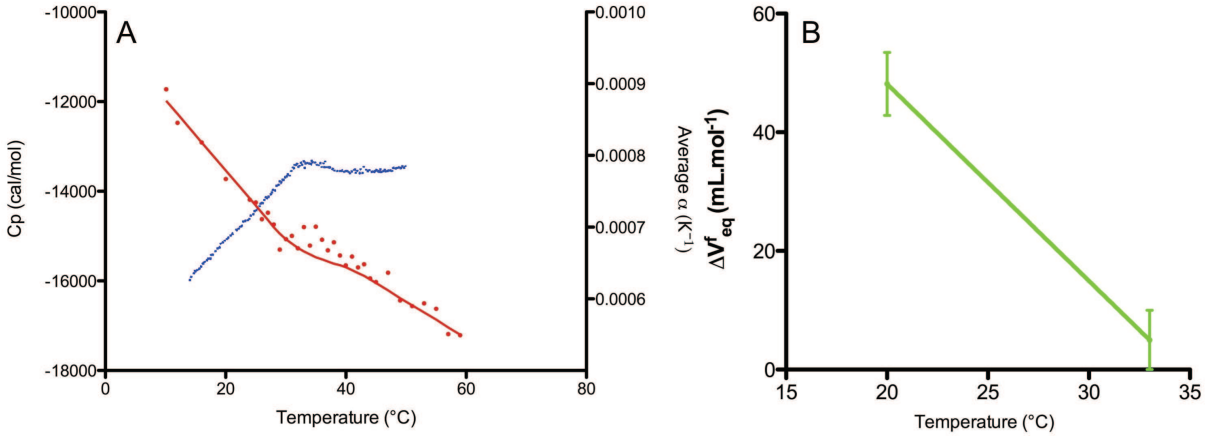

Figure 4.22: DSC PPC of $Nank1-6^{\star}$, and associated expansivity calculation. A : DSC and PPC plot. Data were collected in Phosphate buffer 50 mM, 150 mM NaCl, 1M urea, from 10 to 60°C. For DSC experiments (in blue), $Nank1-6^{\star}$ was concentrated to 2 mg/mL, and 12 mg/mL for PPC experiment (in red). B : estimation of the slope of ΔV^f_{eq} on temperature.

lower temperatures. From DSC data, the Tm was defined at $\approx$ 33°C. We experienced troubles with the baseline of the buffer, as urea was viscuous, so we didn't manage to subtract it well from the buffer-protein scan. This problem explains why the first part of the DSC curve is linear, resulting in a non-defined plateau, and why further determination of the ΔC_p wasn't possible. To the contrary, thanks to the use of 1M urea, the unfolded plateau was well defined, without any aggregation. DSC plot is used here for the Tm determination. As we can observe for the PPC experiment, the expansivity of the folded state of the protein is $0.9.10^{-3}K^{-1}$ at 10°C, and the unfolded state expansivity is of $0.55.10^{-3}K^{-1}$ at 60°C, so in the range of values known from standard proteins. At the moment of the transition, we can see that data are noisy, leading to a not well defined transition curve, but ΔV^f_{eq} was estimated to be + 5 mL/mol $\pm$ 5 at 33°C. We can observe in figure 4.22 B than from the ΔV^f_{eq} known from high pressure experiments and PPC, the dependency of ΔV^f_{eq} with temperature for this protein can be estimated to be - 3.3 mL/mol/°C, arising from a large $\Delta\alpha$ between the two states.

$Nank1-5^{\star}$

As for $Nank1-6^{\star}$, we performed PPC in 1M urea to avoid aggregations. Results were surprising, as shown in figure 4.23. Concerning the DSC experiment, the Tm was found to be at about 33°C, and like for $Nank1-6^{\star}$, the DSC didn't show a stable C_p for the folded state because of buffer problems induced by urea, so calculation of the ΔC_p wasn't possible. DSC is here indicated to estimate the Tm. A highly interesting thing is the expansivity of $Nank1-5^{\star}$ coming from PPC experiments. Contrary to usual cases, the expansivity of the folded state increases a little while increasing temperature, and the expansivity of the unfolded state seems to increase too. As we can observe here, at around 33°C, there's no change in the shape of the expansivities, so ΔV^f_{eq} should be

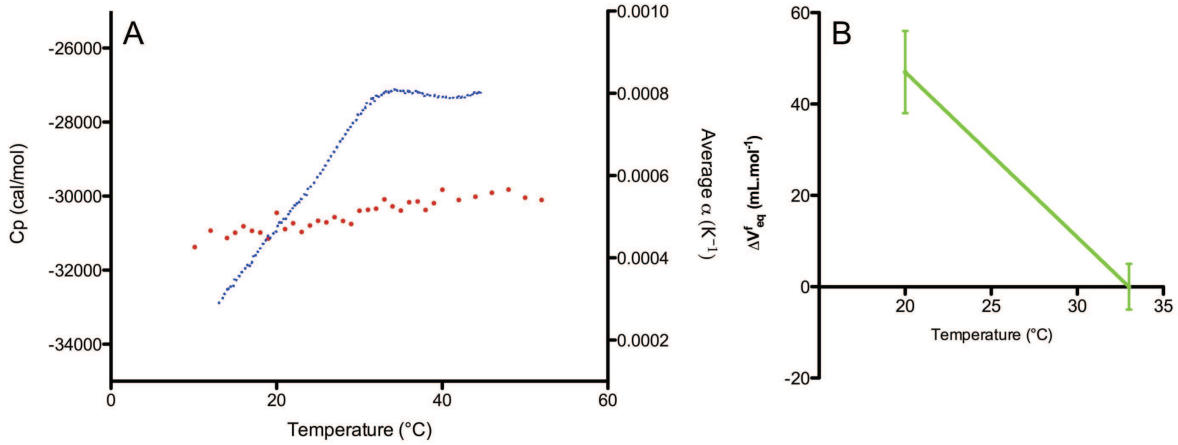

Figure 4.23: DSC PPC of $Nank1-5^\star$, and associated expansivity calculation. A : DSC and PPC plot. Data were collected in Phosphate buffer 50 mM, 150 mM NaCl, 1M urea, from 10 to 60°C. For DSC experiments (in blue), $Nank1-5^\star$ was concentrated to 2 mg/mL, and 12 mg/mL for PPC experiment (in red). B : estimation of the slope of ΔV^f_{eq} on temperature.

equal to ≈ 0 within the error, explaining why we didn't managed to unfold the protein with pressure at 28°C. By using ΔV^f_{eq} at 20°C from high pressure experiments, and at 33°C from PPC, we can estimate the slope $\frac{\delta \Delta V^f_{eq}}{\delta T}$ to be -3.6 mL/mol/°C, as shown in figure 4.23 B.

4.2.4 Kinetic experiments

Mutants with N-terminal deletion

$Nank2-7^\star$

The high pressure jump kinetics of $Nank2-7^\star$ were obtained as described in material and methods, and treated to analyze the $ln(\tau)$ dependency with pressure, as shown in figure 4.24. We can observe, as for $Nank1-7^\star$, an acceleration of the unfolding process while increasing temperature. From the $ln(\tau)$ studies depending on pressure, we know that the pattern is similar to the $Nank1-7^\star$ one. The k^0_f is the only accessible data, so it's still the only rate constant that is really changed by pressure, and the k^0_u remains quiet stable on the pressure range.

The volumetric data obtained by treatment of the whole experimental datasets allows us to perform a complete analysis of the evolution of the different volume changes with temperature for this protein. Results are summarized for 12, 20 and 28°C respectively in table 4.11, 4.12 and 4.13.

As for the $Nank1-7^\star$, we can observe a decrease of the ΔV^f_{eq} while increasing temperature, as shown in figure 4.25 for 2.1M urea. The $\Delta V^\#_u$ remains negative at low temperatures and becomes positive at higher temperatures. The $\Delta V^\#_u$ is equal to zero at 20°C exactly, so at lower temperature than the $Nank1-7^\star$, as shown in figure 4.25. At 28°C, surprisingly, we can observe that the TSE seems to be twice the volume of

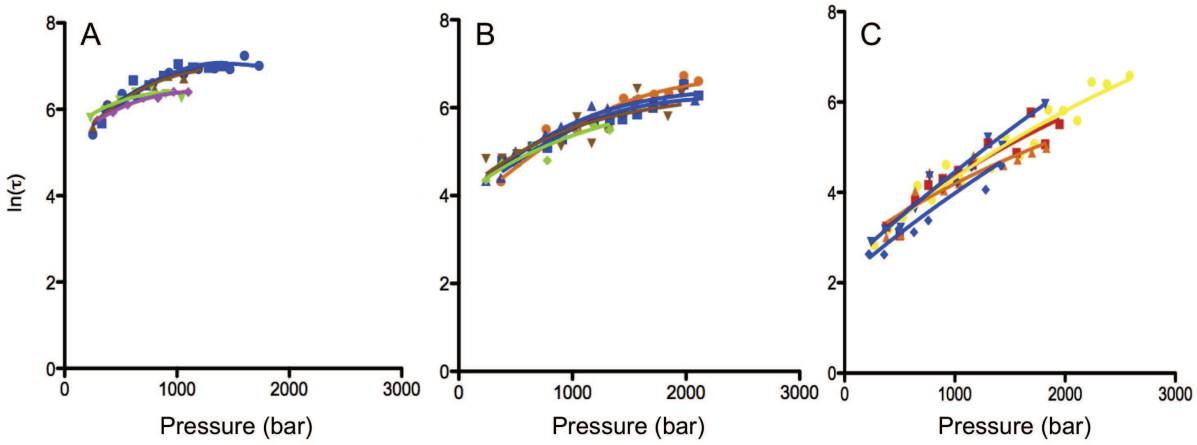

Figure 4.24: Natural logarithm of τ of $Nank2-7^{\star}$, depending on pressure, temperature and urea concentration. A, B and C are respectively at 12°C, 20°C and 28°C. The urea concentration used are 1.7M (red), 1.8M (yellow), 1.9M (orange), 2.1M (blue), 2.3M (brown) 2.4M (green), and 2.5M (purple).

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^{\#}_f$	$\Delta V^{\#}_u$	k^0_f	k^0_u
2.1	2.05	72 ± 7	50.22 ± 2.1	-21.75 ± 2.1	0.0052	0.000157
2.3	1.24	72 ± 7.4	67.86 ± 2.7	-4.14 ± 2.7	0.0056	0.00067
2.4	0.75	72 ± 7.5	61.48 ± 3.6	-10.52 ± 3.6	0.0033	0.00091
2.5	0.71	72 ± 5	69.51 ± 1	-2.49 ± 1	0.0043	0.0013

Table 4.11: Thermodynamic variables of $Nank2-7^{\star}$, at 12°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^{\#}_f$	$\Delta V^{\#}_u$	k^0_f	k^0_u
1.9	1.7	54.58 ± 5.0	56.29 ± 2.5	1.71 ± 2.5	0.026	0.00144
2.1	1.52	54.58 ± 4.6	54.45 ± 2.3	-0.135 ± 2.3	0.018	0.00136
2.3	0.94	54.58 ± 3.7	59.81 ± 3.5	5.23 ± 3.5	0.014	0.0029
2.4	0.4	54.58 ± 5.5	67.23 ± 4.2	12.65 ± 4.2	0.012	0.0064

Table 4.12: Thermodynamic variables of $Nank2-7^{\star}$, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^{\#}_f$	$\Delta V^{\#}_u$	k^0_f	k^0_u
1.7	0.9	33.8 ± 4.1	54.25 ± 4.1	20.45 ± 4.1	0.0624	0.0133
1.8	0.9	33.8 ± 5	57.7 ± 2.3	23.9 ± 2.3	0.0736	0.015
1.9	0.52	33.8 ± 5.8	51.9 ± 5.1	18.1 ± 5.1	0.0465	0.019
2.1	0.1	33.8 ± 4.2	71.83 ± 3.3	38.03 ± 3.3	0.0572	0.048

Table 4.13: Thermodynamic variables of $Nank2-7^{\star}$, at 28°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

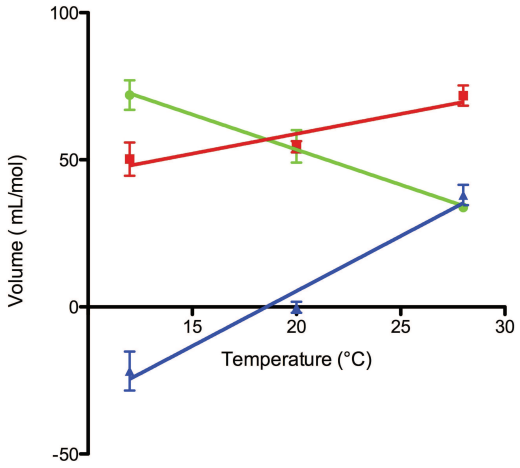

Figure 4.25: Evolution of the volume changes of $Nank2-7^{\star}$ depending on temperature. The ΔV^f_{eq} is represented in green circle, the $\Delta V^{\#}_f$ in red square, and the $\Delta V^{\#}_u$ in blue triangle.

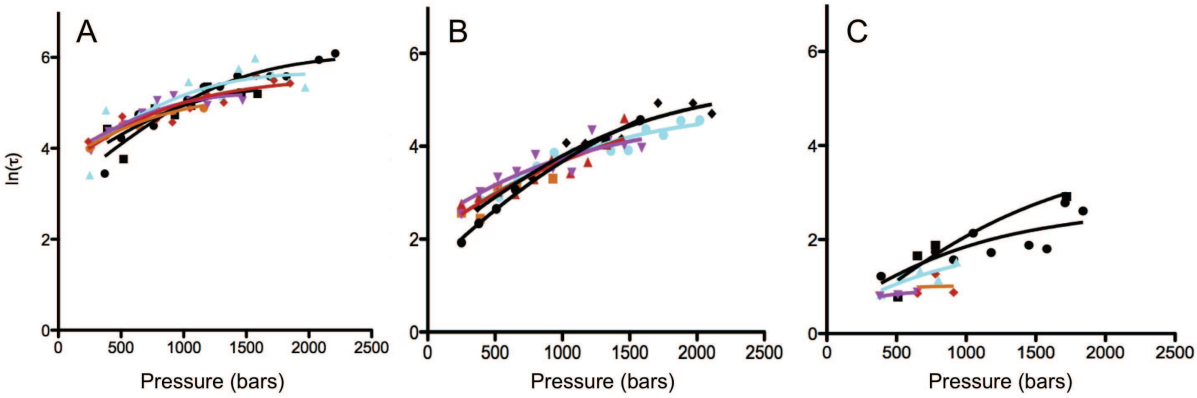

Figure 4.26: Evolution of the $ln(\tau)$ of $Nank3-7^{\star}$ at 12 (A), 20 (B) and 28°C. (C) : $ln(\tau)$ depending on the final pressure of the pressure jumps. Colors used are : black for 1.1M urea, blue for 1.3M urea, purple for 1.5M urea, red for 1.7M urea and orange for 1.9M urea.

the folded state, as the $\Delta V^{\#}_f$ (activation volume difference between the unfolded state and the TSE) is twice bigger than the ΔV^f_{eq}. The slope of $\frac{\delta \Delta V^{\#}_f}{\delta T}$ give us a $\Delta \alpha_{f \rightarrow TSE}$ of $+3.7 \pm 0.8$ mL/mol/°C, so a strong dependence with temperature.

From all these results, we can conclude here that there's still an inversion of the $\Delta V^{\#}_u$ with the temperature.

$Nank3-7^{\star}$

The analysis of the kinetics of pressure jumps (figure 4.26 B) were done by the same way than previously described and helped us determining the $\Delta V^{\#}_f$ and $\Delta V^{\#}_u$ at 12, 20 and 28°C. All the results are summarized in table 4.14, 4.15, 4.16.

Again, the analysis of the $ln(\tau)$ show us an important action of pressure on the k^0_f and not on the k^0_u, as represented in figure 4.26 B. By this analysis, we have seen that there's still a change of sign of the $\Delta V^{\#}_u$, beginning slightly negative at 12°C and

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^{\#}_f$	$\Delta V^{\#}_u$	k^0_f	k^0_u
1.1	1.5	55.4 ± 4.6	52.48 ± 2.9	-2.92 ± 2.9	0.038	0.00294
1.3	1.34	55.4 ± 4	52.68 ± 3.8	-2.72 ± 3.8	0.025	0.00260
1.5	1.21	55.4 ± 3.5	46.93 ± 3.5	-8.47 ± 3.5	0.021	0.00266
1.7	0.89	55.4 ± 7	58.24 ± 2.7	2.84 ± 2.7	0.024	0.00525
1.9	0.64	55.4 ± 10	57.44 ± 2.8	2.04 ± 2.8	0.019	0.00657

Table 4.14: Thermodynamic variables of $Nank3-7^\star$, at 12°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^{\#}_f$	$\Delta V^{\#}_u$	k^0_f	k^0_u
1.1	1.2	50.62 ± 5.1	66.73 ± 2.5	16.11 ± 2.5	0.168	0.0216
1.3	1.22	50.62 ± 5.4	54.6 ± 1.9	4 ± 1.9	0.114	0.0141
1.5	0.71	50.62 ± 5.6	57.89 ± 4.2	7.27 ± 4.2	0.075	0.0222
1.7	0.46	50.62 ± 6.7	68.49 ± 5	17.87 ± 5	0.090	0.0413

Table 4.15: Thermodynamic variables of $Nank3-7^\star$, at 20°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

Urea (M)	ΔG^f_{eq}	ΔV^f_{eq}	$\Delta V^{\#}_f$	$\Delta V^{\#}_u$	k^0_f	k^0_u
1.1	1.32	41.83 ± 6.5	54.97 ± 5	13.14 ± 5	0.54	0.083
1.2	0.71	41.83 ± 6.8	43.93 ± 9.8	2.1 ± 9.8	0.49	0.014
1.5	0.43	41.83 ± 6.3	30.25 ± 2.2	-11.58 ± 2.2	0.37	0.018

Table 4.16: Thermodynamic variables of $Nank3-7^\star$, at 28°C, depending on urea concentration. $\Delta G^f_{eq}(kcal.mol^{-1})$, $\Delta V^f_{eq}(mL.mol^{-1})$, $\Delta V^{\#}_f(mL.mol^{-1})$,$\Delta V^{\#}_u(mL.mol^{-1})$, $k^0_f(s^{-1})$ and $k^0_u(s^{-1})$ are indicated for all the urea concentrations (M).

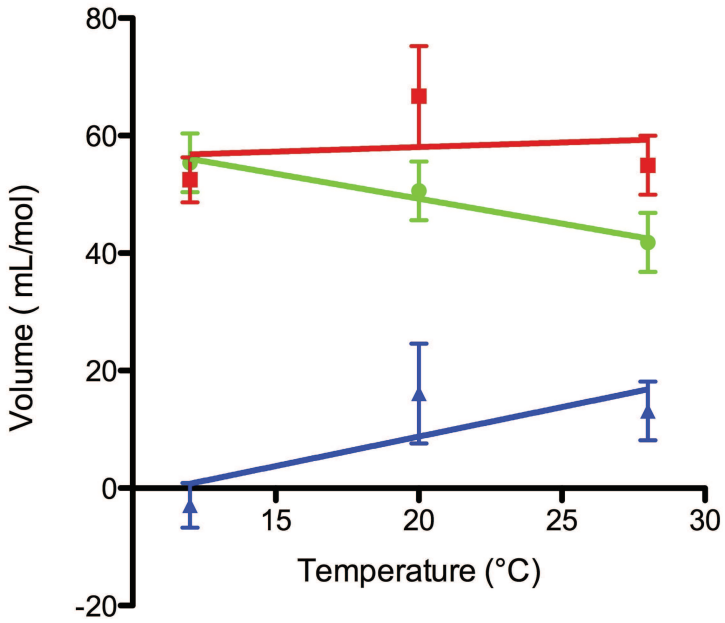

Figure 4.27: Evolution of the volume changes of $Nank3-7^{\star}$ at 12, 20 and 28°C. In green the ΔV^f_{eq}, in red the $\Delta V^{\#}_f$ and in blue the $\Delta V^{\#}_u$

becoming positive at 20°C. The $\frac{\delta \Delta V^{\#}_u}{\delta T}$ is of $+1 \pm 0.8$ mL/mol/°C. This is still proving that the TSE becomes bigger and bigger with temperature, until becoming bigger than the folded state. However, we can observe that the temperature at which the TSE is equal in volume to the folded state is lower than previously seen with the other mutants, at around 15°C, so at lower temperature than for $Nank1-7^{\star}$ and $Nank2-7^{\star}$.

4.2.5 Molecular Dynamic Simulations

We wondered why the volume difference and activation volumes didn't react the same way for all our constructs, so we decided to perform Molecular Dynamic simulations in order to understand where the cavities were in the structure.

Cavity determination in the crystallographic structure of $Nank1-7^{\star}$

In order to have a starting point for our molecular dynamics simulations, Julien Roche, another PhD student from the lab, and I used the known crystallographic structure of $Nank1-7^{\star}$, as found in the pdb under the name "1OT8". The goal was to determine the cavity positions in the 3D structures at the atomic level, in order to determine if all the cavities are taken into account into the definition of the ΔV^f_{eq}, and next to understand how these cavities can contribute to the global ΔV^f_{eq}. We decided to determine the apparent cavities in this structure using the recently published McVol algorithm [231]. In figure 4.28 A, B and C are shown in blue the apparent voids within this crystallographic structure, ie positions with no protein atoms nor water molecules, without any molecular dynamic energy minimization. We can clearly see here that all the voids are

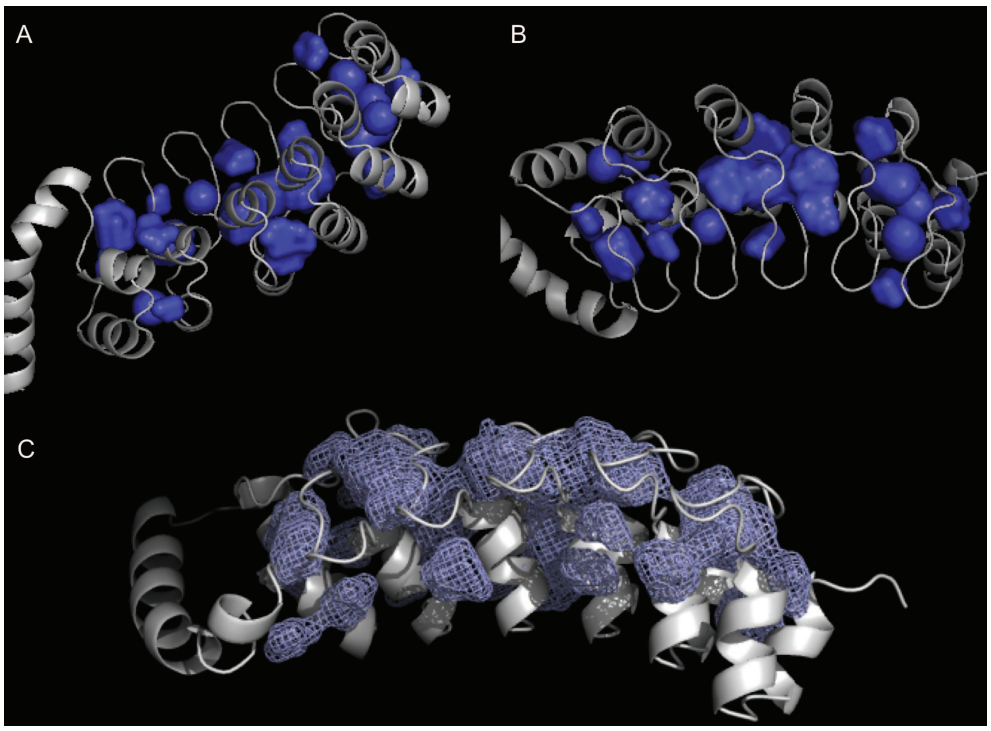

Figure 4.28: Overview of cavities in the *Nank*1−7* crystallographic structure. Protein structure is represented in grey, and cavities in blue. A : view from the bottom of the protein, showing the small linkers between helices of the same motif. N-terminal part of the protein is on the left. B : view from the top, showing the loops linking two distinct motifs. N-terminal part of the protein at the left. C : front view of the protein and the cavities.

found to be placed at the interface between motifs or in their external loops (figure 4.28 C). The overall calculation of the distances from each Monte Carlo point to the nearest water molecule showed us for this only structure that the total apparent voids were of approximatively 160 $mL.mol^{-1}$ at 20°C, so much larger than the ones found by pressure experiments.

This was at least a good starting assumption, leading us to perform advanced molecular dynamic simulations to understand how cavities will behave during time.

Defining dynamics of the apparent cavities

Because one of the limitations of molecular dynamics is CPU-time, we decided first to dissociate our *Nank*1 − 7* structure into three distinct parts, as shown in figure 4.29 in order to decrease the time needed to perform the simulations, and to check if the inner movements are similar between different motifs if the possible capping of repeats 1 and 7 are not present.

Those three smaller theoretical constructs were then used for molecular dynamics simulations, in order to dissociate the effects of the first partly folded motif on the structure of the other motifs. During the first preliminary tests of 10 ns on the 1-3 construct, we saw that the first motif had a lot of internal motions that changed it's

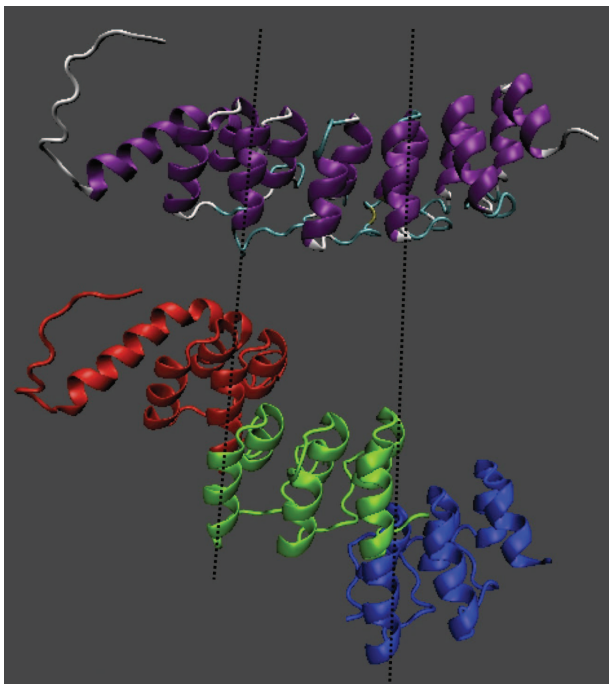

Figure 4.29: Cutting of $Nank1-7^{*}$ structure. The complete structure is shown in purple, and was cut into three parts that always share one motif. The red part contains motifs 1 to 3, the green one the motifs 3 to 5, and the blue one the motifs 5 to 7.

conformation, as shown in figure 4.30. These motions were not seen for the other repeats during the 10 ns of simulations, as we can see in the RMSD plot for each repeat, in figure 4.31. In this figure are reported the RMSD for each repeat separately, during the time of the simulation. As we can observe, from the second to the seventh repeat, all the RMSD are almost linear and without any big changes, always between 0.03 and 0.1 nm of difference with the known structure. But for the first motif, the RMSD changes a lot, from 0.1 nm to around 0.3 nm of difference with the known structure. We can see too that these motions are not evolving smoothly, but with very abrupt changes and no real continuous stabilization period.

We were interested in exploring the dynamic hydration of the cavities. Hence, we performed a water distance analysis of these cavities. In figure 4.32 A, we can observe the distribution of water molecules that are within 3 Angströms of any protein atom, defining a hydration shell. When we calculated the voids inside one structure, by conserving all the Monte Carlo points that were inside the protein, we checked for each Monte Carlo point the distance between each point and the closest water molecule. We use here a color scale from blue, where water molecules are within 2 Angströms, to red, where water molecules are at more than 6 Angströms of the Monte Carlo point, as shown in figure 4.32 B.

As we can see here, it seems that in the 1-3 construct, there's a small part of the whole voids that is highly dehydrated between the second and the third motif, and not

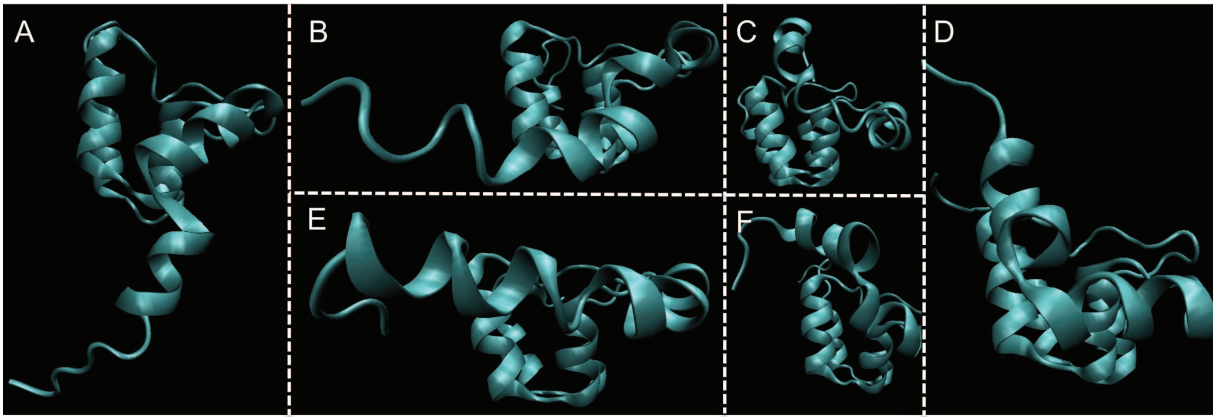

Figure 4.30: Snapshots of the movements of the first motifs in MD simulations

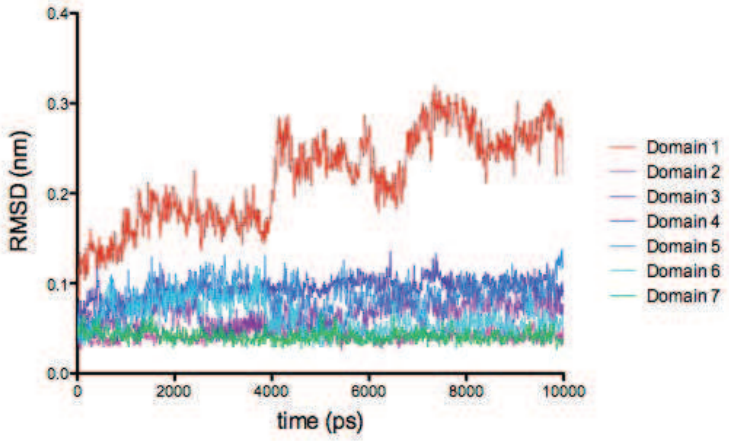

Figure 4.31: Evolution of the RMSD of each motif during the simulation on $Nank1-7^{*}$.

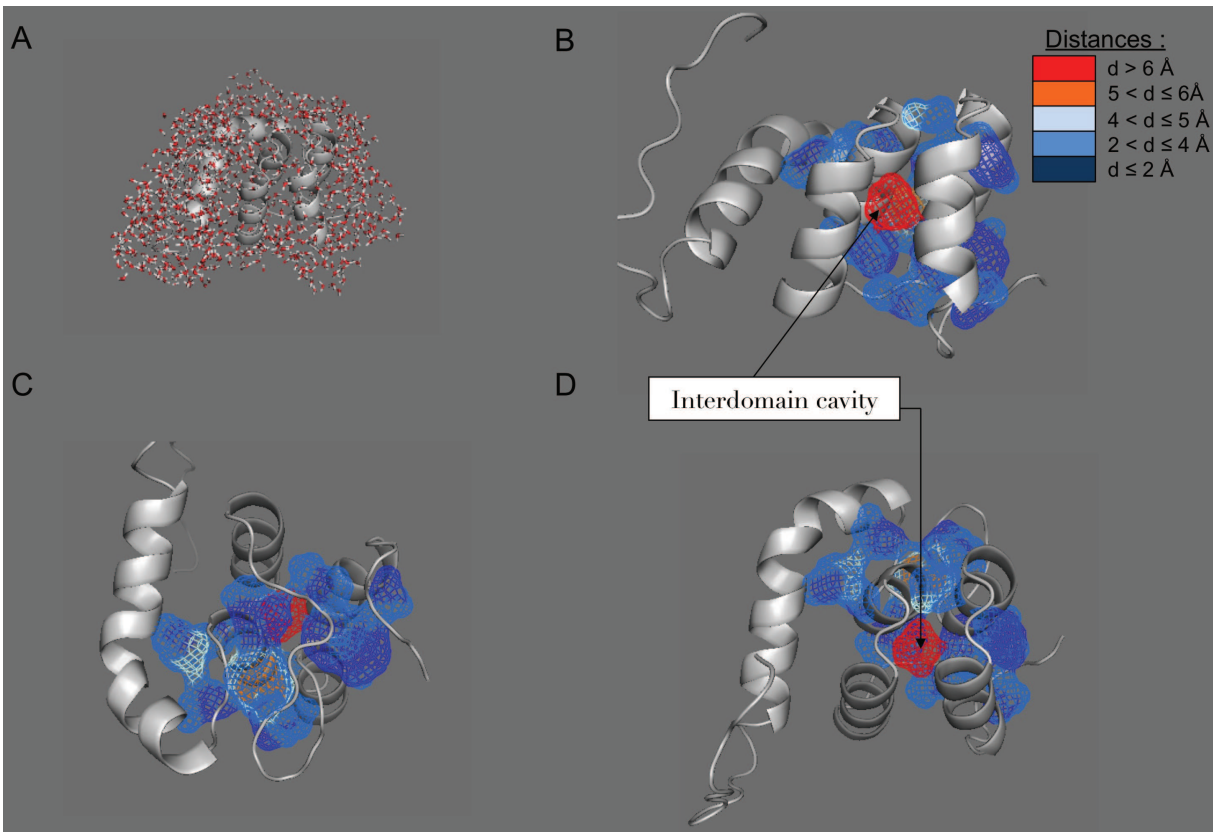

Figure 4.32: Overview of cavities in the 1-2-3 construct. Representative view of a single MD configuration extracted of our MD simulations. A : structure of the 1-3 construct surrounded by water molecules. B : Voids appearing in the structure, see by profile. In blue are shown all the voids, in red the only cavity that doesn't have water molecule close. C : Same structure, seen from the bottom. D : Same structure, seen from the top.

between the first and the second. The fact that this red cavity didn't appear between the first and the second motif comes from the first motif that moves a lot during the simulations, as such movements can just make water enter at the interface. The red cavity appearing in the second interface is about 10 $mL.mol^{-1}$.

In figure C and D are represented the same structure but by different point of view to have a better sight of the position of this red cavity. This dehydrated void, what we will call the interdomain cavity, is placed in the middle of the interface between the two repeats, in the upper part of the interface, far from the interdomain loop.

Contrary to the 1-3 construct, we didn't observe any big changes in structure of a motif during the whole simulations for the 3-5 and 5-7 construct. This is relevant of the fact that the first motif is highly different from the other ones in terms of stability.

All the results obtained the 3-5 and 5-7 constructs tend to the same red interdomain cavity that seems to be completely dehydrated in the upper part of the apparent cavities at the interface between motifs, and for which the volume is always about 10 mL/mol.

Exploring the voids

From all the results we have produced by MD simulations until this point, we always had the impression that there were no real voids between two helices of the same motif. So we decided to confirm our results with the 3-5 and 5-7 constructs, in order to determine more accurately where were the voids that remain hydrated or not within the structure during the 10 ns of time used.
In order to check the proximity of water to the inner voids, we placed four probes of a 4 Angströms radius inside the structures taken from the $Nank1-7^{\star}$ coordinates, as shown in figure 4.33 A and C respectively.

We can see that for both 3-5 and 5-7 constructs, the probes were placed at about the same positions. The two red probes (cavity 1 and 3) were placed at the upper part of the interface between two motifs in order to fit the theoretical position of the dehydrated void we saw in previous calculations. The cavities 2 and 4 were placed between the two helices of the central motif, were we previously saw poor hydration. With this systematization of the probes, we expected to be able to see during all the MD simulations the amount of water that enters into the structure and fill the voids.

We know from simulations that the bottom part of the helices of the motifs show big motions, contrary to the upper parts that show only small motions, so we estimated that the lower part of the interfaces between two motifs, where are located the interdomain loops, are transiently hydrated, and that the cavities that we can see at these positions will just be filled with water and won't have a big importance in the calculations of the ΔV^{f}_{eq}.

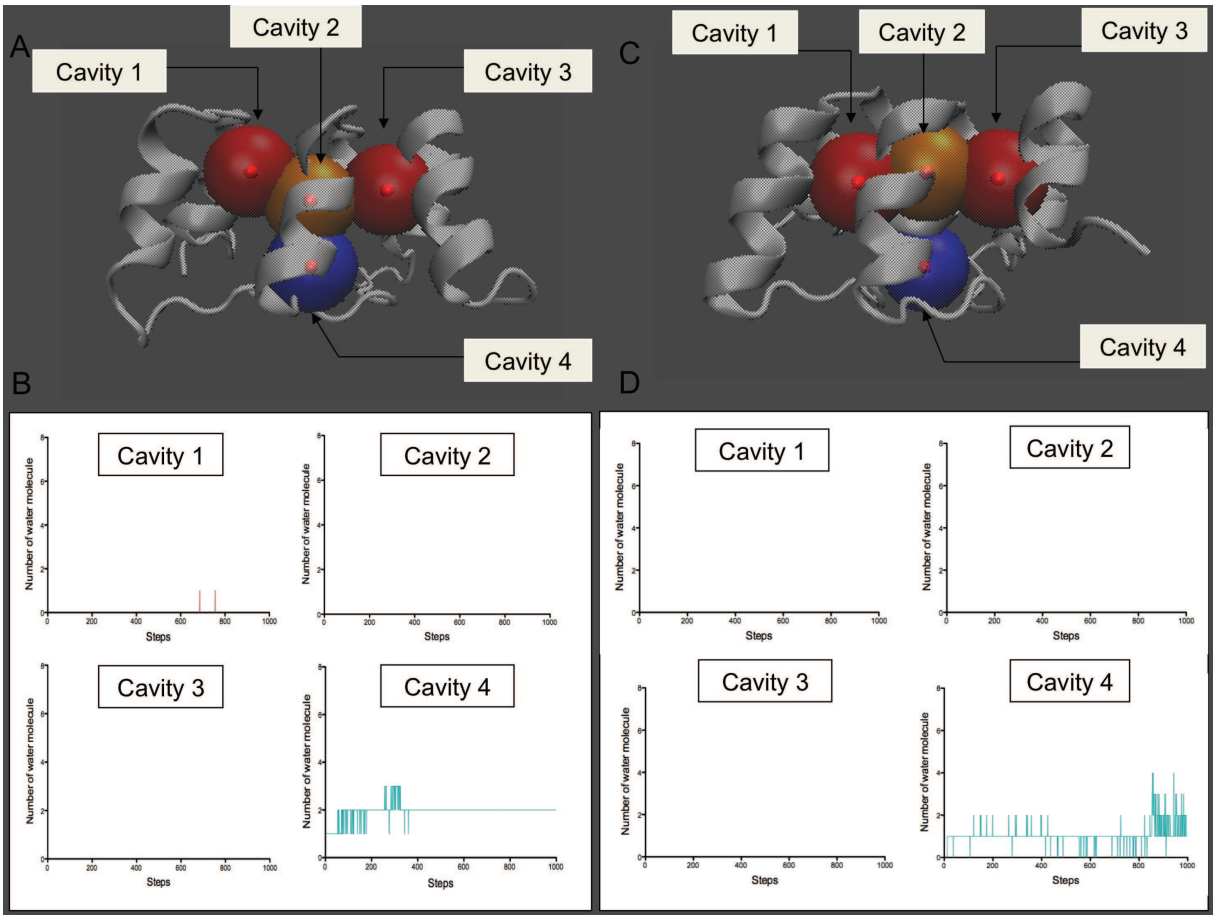

Figure 4.33: Cavity probes and cavity hydration in MD simulations for 3-5 and 5-7 constructs. A : probes placed in the 3-5 structure. B : Number of water molecules that enter into the probes environment of 3-5 construct during 10 ns of simulation (1000 steps of 10 ps each), for each cavity. C : probes placed in the 5-7 structure. D : Number of water molecules that enter into the probes environment of 5-7 construct during 10 ns of simulation (1000 steps of 10 ps each), for each cavity.

As we can see in figure 4.33 B for the 3-5 construct, the cavity 1 only accept two transient water molecules in it's surrounding during the 1000 steps, and it's homologous cavity 3 is completely dehydrated during all the 1000 steps. This leads to the conclusion that the voids appearing at these positions remains completely dehydrated in all the cases. We can too observe that in cavity 2, we don't see any water molecule in the surrounding, as expected from the fact that in figure 4.28 we didn't see any voids at these positions. Indeed, as there's no void, all the surrounding of these probes are filled with the amino acid side chains and backbone. The cavity 2 is so used as a negative control of these calculations.

Contrary to the other cavities, the cavity 4, which is at the bottom of the two helices motifs, within the motif, is highly populated by water during all the simulation, because of the important motions of this part of the protein, and is used here as a positive control.

In figure 4.33 D, we observe the exact same pattern for all the cavity probes in the 5-7 construct. We can conclude from these calculations that the global pattern of the voids are almost identical within all the constructs.

Another interesting thing is the NAG5 protein. An alanine was replaced by a glycine, theoretically leading to an increase of the void, because of the lack of the CH_3 of the lateral chain, as shown in figure 4.34 A, B and C. As we can see, this mutation increase the cavity size in a zone where water is in proximity, far from the red cavity. So this mutation brings us a way to confirm or reject our hypothesis about the importance of only a small part of the global void as a parameter of the ΔV^f_{eq}, as we will see during discussion.

Total amino acid hydration analysis

Finally we performed a simulation of the complete $Nank1-7^*$ protein, and we calculated the density of Monte Carlo points and water molecules, respectively in figure 4.35 A and B, within 4 Angströms of each amino acid $C\alpha$. In a second time, we made the same void analysis than previously shown in one single structure but in the $Nank1-7^*$ structure, as shown in figure 4.35 C.

We determined the void and water density around the Cα of each residue in the structure of $Nank1-7^*$. The void density is determined as the number of Monte Carlo points within a 4 Angströms sphere around each Cα, so it describes how each Cα is surrounded by voids. The water density is determined as the number of oxygen atoms (from water molecules) within a 4 Angströms sphere around each Cα, so it describes how close each Cα is from water. We can observe in figure 4.35 A that the density of the voids within all the 1000 structures obtained by MD simulations are dispersed in the interface between all the repeats, in the interior of the protein, from the sec-

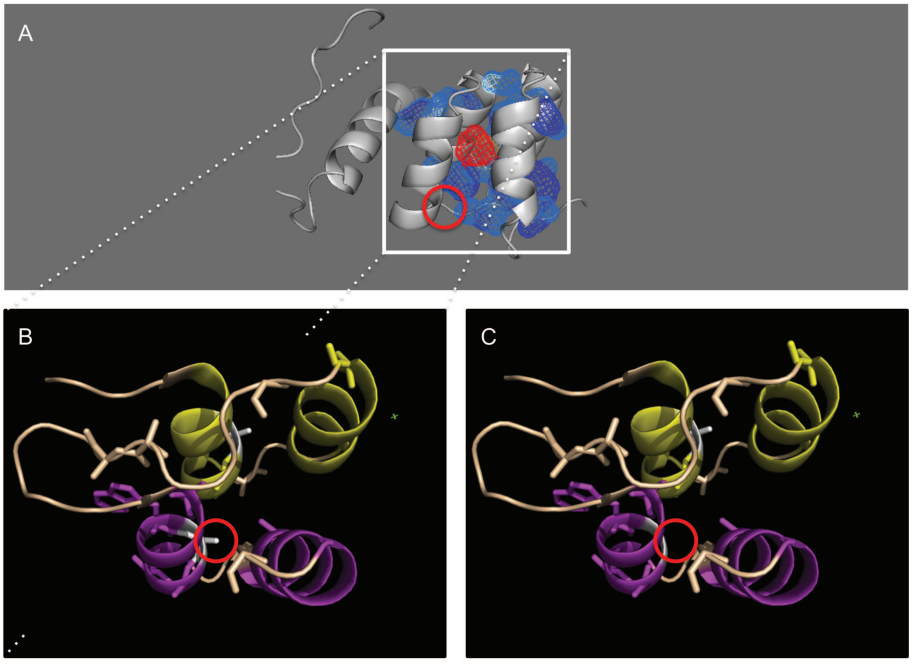

Figure 4.34: Cavities presented and attended respectively in $Nank1-7^*$ and NAG5. The red cavity corresponds to the cavity that seems important from MD simulations. The red circle corresponds to the position of the side chain of the mutation (alanine to glycine). A : structure of 3-5 MD construct with cavities represented. B : zoom on the fourth and fifth motif, with alanine circled. C : zoom on the fourth and fifth motif with glycine instead of alanine circled.

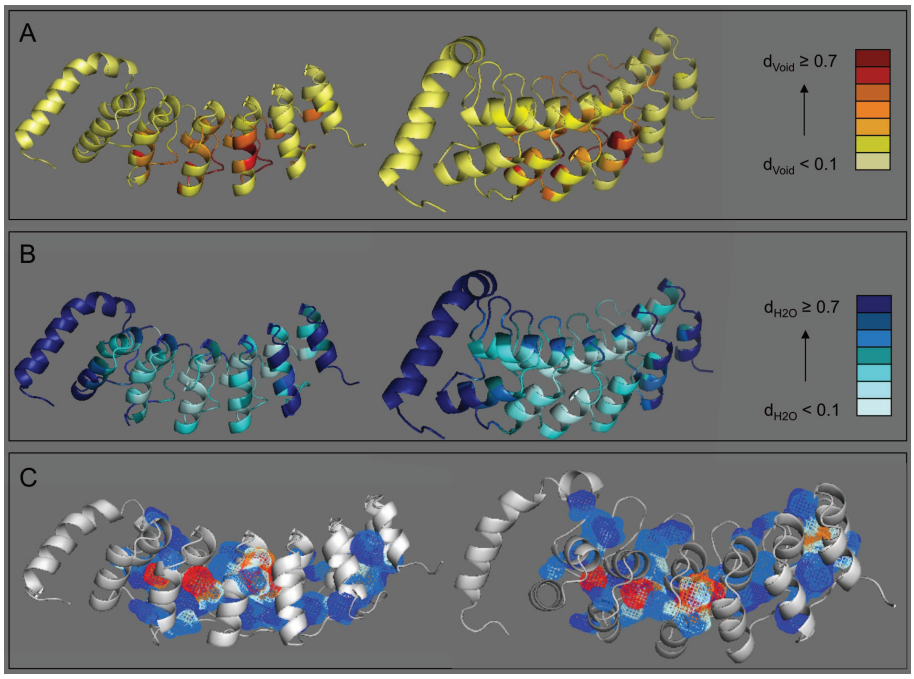

Figure 4.35: MD simulations of the whole $Nank1-7^*$ protein and calculations of voids and hydration. A : calculations of the density of the voids for each amino acids. The color turns from yellow to dark red while increasing the density of the voids. B : calculations of the density of waters in the surrounding of each amino acid, colored from light blue to dark blue while increasing the density of water. C : estimation in one single structure of the overall voids and of the highly dehydrated voids.

ond motif to the seventh one. It is important to note that the first motif seems not to have any void in it's environment, and seems not to define a void in it's interface with the second motif. Another point of interest is the highly dehydrated void that can be seen in the direct environment of the tryptophan in the fifth motif (red spot), which is highly consistent with all our fluorescence experiments, where we were able to see a significantly buried tryptophan, even if it's side chains points out of the structure of the protein.

All these observations are in agreement with 4.35 B, where we can see the density of water around the amino acids of this protein. We can clearly observe that the first motif is highly hydrated, making impossible the presence of any contributing void in its surrounding, and the motifs 3 to 5 are highly dehydrated. We can observe too that the amino acids that are buried inside of the protein are highly dehydrated in average. The motifs 2, 6 and 7 are, contrary to the motifs 3 to 5, more hydrated at the exterior of the structure but remains dehydrated inside the motifs. All these results indicate that the voids are buried within the structure at the interface between two motifs, and that the hydration is highly dependent on the internal motions of the protein, driven by its cooperativity of folding.

Last but not least, we can observe that while we have a lot of motifs, the C-terminal ones are a little less sensitive to pressure because they are more hydrated. Finally, we can observe that even if the repeats 6 and 7 seem to be a little less packed than the other ones, and more sensitive to hydration, the amino acids that are inside the interfaces are highly dehydrated, leading to possible dehydrated voids.

In 4.35 C, we took a representative structure from the MD simulations, and calculated the distance of the Monte-Carlo points from the closest water molecule, and found again that in the upper part of the motifs, a highly dehydrated cavity where the distance to the first water molecule is of more than 6 Angströms. Within the two N-terminal interfaces, we can see that because of this higher hydration of the amino acids, the voids are more hydrated, but there's still one cavity quite dehydrated (about 4 Angströms from water molecules) remaining between repeats 6 and 7.
From MD simulations, we can propose two hypotheses. First, all the cavities shouldn't be important for the ΔV_{eq}^{f} parametrization, but only a small part of them, which is highly dehydrated. Second, it seems that these highly dehydrated cavities are only located between motifs 3-4, 4-5 and a little bit between motifs 2-3.

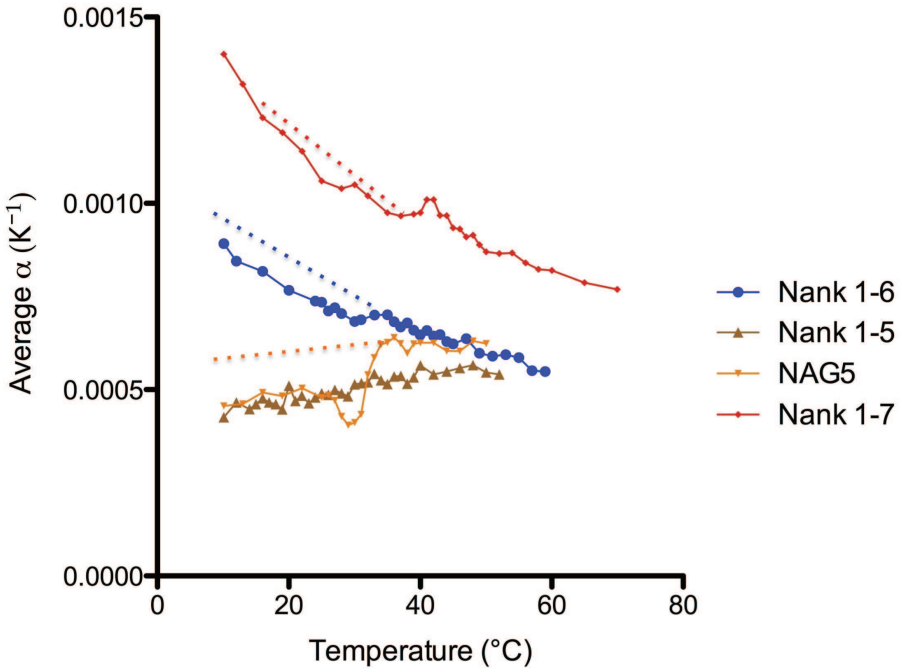

Figure 4.36: Pressure perturbation calorimetry plots of $Nank1-7^{\star}$, $Nank1-6^{\star}$, $Nank1-5^{\star}$ and NAG5. Expansivity was plotted versus temperature. Dotted lines are putative unfolded state expansivity at low temperature, based on the high temperature unfolded state expansivity and on the fact that folded and unfolded state expansivity are parallels.

4.3 Pressure perturbation calorimetry study

4.3.1 Ankyrin repeat constructs

We have plotted in figure 4.36 the PPC datasets of $Nank1-7^{\star}$, $Nank1-6^{\star}$, $Nank1-5^{\star}$ and NAG5 as a function of temperature. We can remark first that the expansivity of the unfolded state seems similar for $Nank1-6^{\star}$, $Nank1-5^{\star}$ and NAG5 at 50-60°C within the error. This is a clue saying that the unfolded state expansivity doesn't depend on the exact sequence, as NAG5 and $Nank1-7^{\star}$ differs only by one substitution, and as NAG5 and $Nank1-5^{\star}$ have two repeats of difference. We can observe also that the differences between the folded and unfolded states expansivities of $Nank1-6^{\star}$ and $Nank1-7^{\star}$ are quite low. This $\Delta\alpha$ is more important in the case of $Nank1-6^{\star}$ (unfolded state expansivity at low temperature is symbolized by a dotted blue line), and seems to be twice that of $Nank1-7^{\star}$. We can observe that the $\Delta\alpha$ of NAG5 is about twice those two previous ones. NAG5 contains all the repeats, like $Nank1-7^{\star}$, so it is highly surprising that NAG5 don't have the same expansivity than $Nank1-7^{\star}$, and more surprising that NAG5 has got the same expansivity of the folded state than $Nank1-5^{\star}$. Finally, we have determined experimentally that $\frac{\delta\Delta V}{\delta T}$ of $Nank2-7^{\star}$ and $Nank3-7^{\star}$, which are respectively of -2.38 and -0.84 mL/mol/°C. Here we can see that $Nank2-7^{\star}$ has its $\Delta\alpha$ between that of $Nank1-7^{\star}$ and $Nank1-6^{\star}$, and that $Nank3-7^{\star}$ has a $\Delta\alpha$ intermediate between that of $Nank1-5^{\star}$ and $Nank2-7^{\star}$. So it seems that $\Delta\alpha$ do

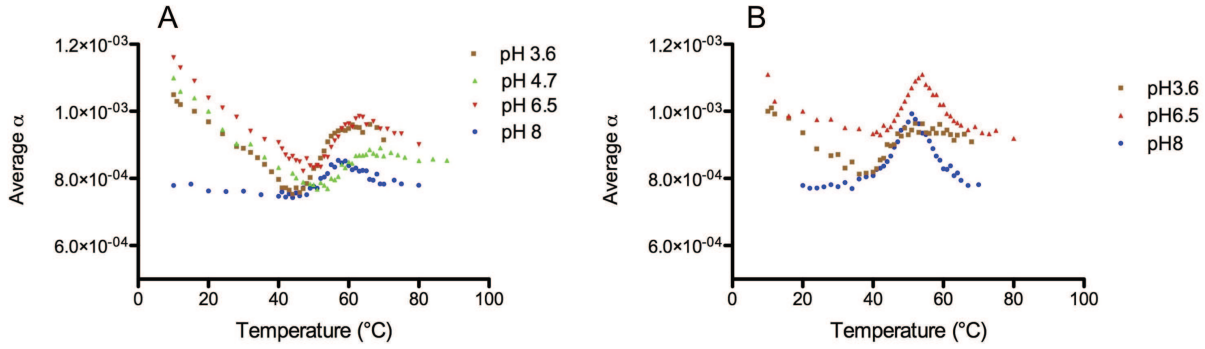

Figure 4.37: PPC of BPTI proteins. A : PPC of BPTI5,55A14GA38V at pH 3.6, 4.7, 6.5, and 8. B : PPC of BPTI5,5521A14GA38V at pH 3.6, 6.5 and 8.

The first thing we can note is a strong dependence of both folded and unfolded state expansivity with pH, as the more we increase pH the lower the expansivities are. We can note to that at pH 8, for both of them, the expansivity of the folded state is flat, as previously shown for Ankyrin $Nank1-5^{\star}$. Another aspect really interesting is comparing PPC of each of these proteins, pH per pH, as shown in figure 4.38 A, B and C.

not change in a size dependent fashion, indicating that we really don't understand its properties. The thermal expansivity of the transition state of the full-length construct is equivalent to that of the unfolded state and much larger than that of the folded state; i.e., the values of $\Delta\alpha_u$ and $\Delta\alpha_u^{\#}$ were equivalent. The $Nank3-7^{\star}$ construct exhibited a difference in expansivity of the transition and unfolded states relative to the folded state that was somewhat smaller than for $Nank1-7^{\star}$. The value of $\Delta\alpha_f^{\#}$ and $\Delta\alpha_u$ were lower for $Nank3-7^{\star}$, compared to that of $Nank1-7^{\star}$. Interestingly, the $Nank2-7^{\star}$ construct exhibited a significant difference in expansivity between the transition state and the unfolded state; $\Delta\alpha_u^{\#}$ was found to be larger than the value of zero for the $Nank1-7^{\star}$ and $Nank3-7^{\star}$ constructs. In addition, compared to $Nank1-7^{\star}$, the $Nank2-7^{\star}$ construct exhibited a much larger value of the expansivity of the transition state relative to the folded state. This very large expansivity of the transition state of $Nank2-7^{\star}$ is consistent with the increase in its molar volume upon deletion of the N-terminal repeat.

4.3.2 PPC study on BPTI

To better understand the folded state expansivity, we decided to use another model : the Bovine Pancreatic Trypsin Inhibitor (BPTI), for which a lot of mutants exist. We had been given two mutants from the team of Yutaka Kuroda : BPTI5,55A14GA38V and BPTI5,5521A14GA38V (a mutant based on BPTI5,55A14GA38V for which 21 amino acids of the surface were changed into alanines with changing the 3D structure). I performed PPC on these two mutants for various pH, and results are shown in figure 4.37 A for BPTI5,55A14GA38V and B for BPTI5,5521A14GA38V. In A, we can see that at pH 3.6, for both proteins, the expansivity of the folded state is identical, and so for the

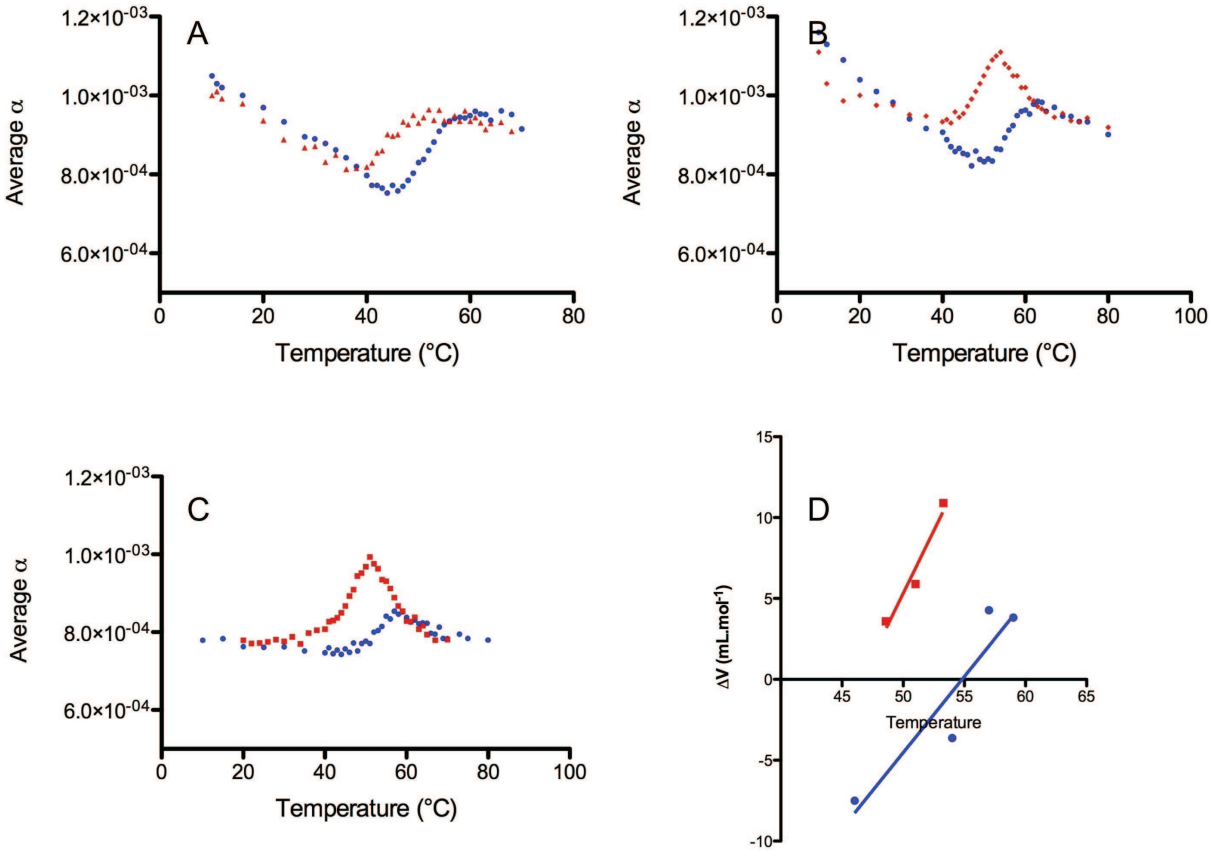

Figure 4.38: PPC of BPTI mutants depending on pH and associated ΔV^u_{eq}. PPC of BPTI5,55A14GA38V (blue) and BPTI5,5521A14GA38V (red) at pH 3.6 (A), 6.5 (B) and 8 (C). D : plot of the determined ΔV^u_{eq} versus temperature for both proteins.

unfolded state. This similarity is seen at each pH (figure 4.38 B and C. Treating these PPC datasets was easier than for ankyrin repeat proteins, as the buffer was a standard one (citrate-phosphate 50 mM) and estimated ΔV^u_{eq} were plotted in figure 4.38 D. As we can observe, BPTI5,55A14GA38V has always a lower ΔV^u_{eq}, from -7 mL/mol at 47°C to $\approx$ 5 mL/mol at 58°C. Here again we can see a ΔV^u_{eq} change of sign with temperature. It appeared for BPTI5,5521A14GA38V that the ΔV^u_{eq} increased a lot, become always positive from 45 to 55 °C.

Part IV

Discussion and conclusions

Chapter 5

Discussion and conclusions

5.1 Discussion : Volume change for Ankyrin repeat protein folding

The ankyrin repeat proteins are elongated and constructed of repeated units of same structure. In such a model, we can imagine that some aspects of the 3D structure are repeated. Thus, if the origin of ΔV_{eq}^{f} is hydration of newly exposed peptide bonds upon unfolding, then deleting repeats will decrease the magnitude of the ΔV_{eq}^{f}. If ΔV_{eq}^{f} is determined by the degree of differential hydration of exposed amino acid side chains then the nature of the exposed amino acid residues in deleted motifs should make a large difference in the value of ΔV_{eq}^{f}. Finally, if ΔV_{eq}^{f} comes from internal solvent excluded voids, then the effects of deleting specific repeats will depend upon their individual packing characteristics.

5.1.1 How do size and sequence affect the volume change upon unfolding ?

How does size contribute to ΔV_{eq}^{f} ?

In order to study the ΔV_{eq}^{f} dependence with the size of the protein, we performed experiments at 20°C on several deletion constructs based on the full-length $Nank1-7^{\star}$: $Nank1-6^{\star}$, $Nank1-5^{\star}$, $Nank2-7^{\star}$, $Nank3-7^{\star}$ and $Nank4-7^{\star}$. Upon deletion of the first motif, in $Nank2-7^{\star}$, we observed a ΔV_{eq}^{f} of slightly larger magnitude than that of $Nank1-7^{\star}$. And further deletion of repeat 2 in $Nank3-7^{\star}$ had no further effect on ΔV_{eq}^{f}. Interestingly, deletion of motif 3 in $Nank4-7^{\star}$ lead to a small decrease of the ΔV_{eq}^{f} compared to $Nank2-7^{\star}$ and $Nank3-7^{\star}$, but the value of this ΔV_{eq}^{f} is similar to that of $Nank1-7^{\star}$. The $Nank1-6^{\star}$ and $Nank1-5^{\star}$ constructs also show no difference in the value of ΔV_{eq}^{f} as compared to the full-length $Nank1-7^{\star}$.

These data show us clearly that deletion of the exterior ankyrin repeats has no significant effect on ΔV_{eq}^{f}, except a small unexpected increase in the case of

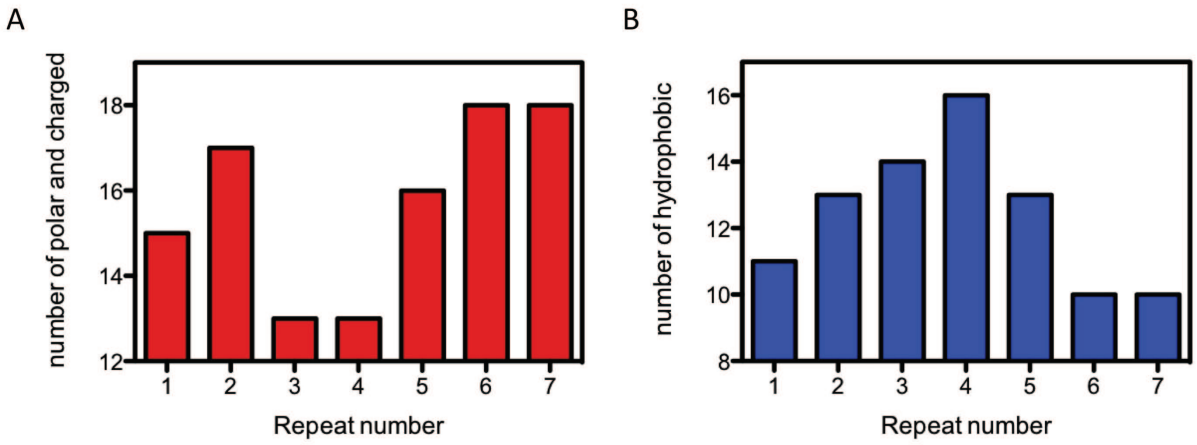

Figure 5.1: Sequence analysis of repeats, by counting the number of polar and charged (A) or hydrophobic (B) residues per repeat.

$Nank2-7^{\star}$.

The volume of atoms does not change, of course upon unfolding, so the origin of the ΔV^{f}_{eq} must be in either the hydration of peptide bonds or side chains or in disappearance of voids. Deleting up to 3/7 of the peptide bonds has no impact on the ΔV^{f}_{eq}. Hence, hydration of exposed peptide bonds in the unfolded state cannot account for the different contributions of each repeat to the volume change, since these repeats are of nearly identical structure and size, and hence, involve the exposure of a similar number of peptide bonds. Clearly, peptide bonds do not add any value to ΔV^{f}_{eq} for ankyrin repeat proteins, contrary to the hypothesis that hydration of peptide bonds makes a major contribution to pressure effects on proteins [232].

How does sequence contributes to ΔV^{f}_{eq} ?

Another parameter we wanted to explore was the contribution of differential hydration of exposed amino acid side chains to the magnitude of the ΔV^{f}_{eq}. Specific changes in the content of each kind of residues among the different repeats may be important. Examination, repeat per repeat, of the number of polar/charged and hydrophobic residues (as reported in figure 5.1) shows that the distribution of polar/charged and hydrophobic residues is not homogeneous in the sequence.

Differential hydration contributions involve volumetric effects on the unfolded state, since they arise from changes in the density of water molecules that move from the bulk to interaction with the protein moeities upon their exposure to solvent upon unfolding. Thermal expansivity involves a release of hydration water to the bulk upon heating. The thermal expansivity of polar compounds in water is large and positive at low to moderate temperatures. Therefore, hydrating newly exposed polar residues is linked to an increase in density (decrease in volume) for the newly interacting water molecules compared to their volume in bulk. Hence hydration of newly exposed polar residues should contribute negatively to ΔV^{u}_{eq}, such that their removal upon deletion of a repeat

should lead to a smaller absolute value for ΔV^u_{eq}. In our case, polar and charged amino acids are in majority in repeats 1, 2, 5, 6 and 7. This means that deleting these repeats should lead to a large decrease in the magnitude of ΔV^f_{eq}. In contrast to this expectation, deletion constructs where these motifs are deleted, $Nank2-7^\star$, $Nank3-7^\star$, $Nank1-6^\star$ and $Nank1-5^\star$, systematically show similar ΔV^u_{eq}.

The opposite is true for hydration of hydrophobic moieties. The density of water hydrating polar residues is lower than in bulk (the thermal expansivity of hydrophobic compounds in water is negative at low to moderate temperature). Hence, hydration of exposed hydrophobic residues should contribute positively to the value of the molar volume of the unfolded state at low to moderate temperature, and hence their deletion should in principle lead to a smaller molar volume of the unfolded state, and hence a larger absolute value for ΔV^f_{eq}. The hydrophobic residues are numerous in repeats 2, 3, 4 and 5, but mutants associated with deletion of repeats 2 and 3 have similar ΔV^f_{eq}.

Thus, arguing against a major role for differential hydration as the basis for pressure effects are the sequence differences between the repeats. Indeed, the central repeats contain the largest number of hydrophobic residues (that should diminish the overall magnitude of ΔV^u_{eq}) and the smallest number of polar and charged residues (those that should increase the magnitude of ΔV^u_{eq}). Whereas we have found that most of the magnitude of ΔV^f_{eq} for $Nank1-7^\star$ resides in the central repeats 3, 4 and 5. Moreover, volumetric effects associated with differential hydration should be strongly temperature dependent. One would therefore expect for the deletion constructs, a larger difference in expansivity between the unfolded compared to the folded state, $\Delta\alpha_u$, compared to $Nank1-7^\star$. The $\Delta\alpha_u$ for $Nank2-7^\star$ compared to $Nank1-7^\star$ are consistent with this hypothesis. However, this is not true for the $Nank3-7^\star$ construct, whose value of ΔV^u_{eq} is nonetheless significantly larger in absolute value than that of $Nank1-7^\star$.

Thus, based on the negative correlation between the sequence differences between the repeats and their predicted effects on ΔV^u_{eq}, as well as on the contradictory temperature effects, we conclude that the volumetric properties of the Nank deletion constructs cannot arise from differential hydration effects.

To this conclusion, we can add some more clues. Indeed, the full-length construct $Nank1-7^\star$ was studied by high-pressure fluorescence, SAXS, FTIR, DSC and PPC to allow us a global overview of its volumetric properties. High pressure fluorescence experiments and SAXS under pressure revealed pressure induced unfolding with a plateau near 3 kbar. Pressure-induced unfolding experiments by fluorescence, SAXS and FTIR were done at several urea concentrations and showed us that urea has significant effects on the physical properties of the pressure-induced $Nank1-7^\star$. SAXS showed us that at moderate urea concentration, the pressure-unfolded ensemble wasn't completely expanded at 3 kbar, with a pair distance distribution function profile less large than at higher urea concentrations. Even if we saw that the pair distance distribution profile were highly pressure dependent, at 3 kbar they are different between 2

and 2.2M of urea, even if the pressure-induced unfolding is complete. Hence, increasing urea leads to greater solvent exposure of the tryptophan residue and a more expanded Gaussian chain at high-pressure, without any change in molar volume. Changes in hydration induced by urea seem not to change the ΔV_{eq}^{f}.

At 20°C, kinetic experiments showed us on $Nank1-7^{\star}$ that the ΔV_{eq}^{f} is similar to $\Delta V_{f}^{\#}$, so that the difference between the volume of transition state ensemble and the unfolded state is similar to the difference between than of the unfolded and the folded states. This means that the volume of the TSE is nearly as large as than of the folded state, so, according to our model, the bulk of the solvent excluded packing defects have formed in the TSE of $Nank1-7^{\star}$ at several temperatures.

In contrast, our analysis at several temperatures of $Nank1-7^{\star}$ revealed that increasing temperature makes the TSE increase its volume compared to the folded one, even becoming larger in volume than the folded state. We determined that the dependance of the $\Delta V_{u}^{\#}$ and ΔV_{eq}^{f} with temperature were similar in absolute value, contrary to that of the $\Delta V_{f}^{\#}$ which is much lower. Hence, the thermal expansivity of the TSE is equivalent to that of the unfolded state, so the TSE is malleable, as the unfolded state, but is equivalent in volume to the folded state.

This high plasticity of the TSE implies that the tertiary interactions are not locked-in in the TSE, so that the TSE is unfolded like. Hence, the hydration difference between the TSE and the unfolded state should be minimum while their volumetric properties are completely different. This, adding to the fact that increasing urea, which interacts with the protein backbone, does not change the ΔV_{eq}^{f}, learn us that for $Nank1-7^{\star}$ the ΔV_{eq}^{f} can't come from differential hydration.

Definitively, differential hydration effects do not contribute appreciably to the value of ΔV_{eq}^{f} for ankyrin repeat proteins.

5.1.2 How do cavities contribute to ΔV_{eq}^{f} ?

As neither peptide bonds or differential hydration seem to make a significant contribution to ΔV_{eq}^{f}, we can emit the hypothesis that ΔV_{eq}^{f} comes from cavities in the structure.

α-helices and ΔV_{eq}^{f}

While looking at the FTIR data on $Nank1-7^{\star}$ at 20°C, we saw that at 2M urea, a loss in α-helical content occurred at 3kbar, showing only $\approx$ 50 % of α-helices (contrary to 70% in the native state), contrary to 2.35M urea where the helical content is 30%. At 3 kbar and 2M of urea, both SAXS and high-pressure fluorescence showed a pressure-unfolded ensemble, as the tryptophan is as red-shifted as possible and the chain condensation is at its minimum. All these data are consistent with a pressure-unfolded ensemble that retains about 50% of native α-helical structure. This high α-helical content is not

surprising as pressure has little effects on α-helix stability. Moreover, it is interesting to note that these experiments showed identical ΔV_{eq}^{f} between the urea concentrations. Thus, α-helices do not add any contribution to ΔV_{eq}^{f}. This is not entirely surprising since pressure has been shown by experiment and simulations to have a negligible effect on α-helices [233, 227].

Localization of cavities in the structure thanks to experimental results and MD simulations

As previously explained in introduction, we have strong evidence that the sixth repeat is unfolded when the seventh motif is deleted, so that the folded structure of $Nank1-6^{\star}$ may be the same than that of $Nank1-5^{\star}$. As the sixth motif is unfolded, it supports the fact that the sixth motif does not change the ΔV_{eq}^{f}, as the packing and the number of cavities should be the exact same as for $Nank1-5^{\star}$, as no cavity can be formed between repeats 5 and 6 when repeat 7 is missing. Moreover, we can obviously see that $Nank1-7^{\star}$ has roughly the same ΔV_{eq}^{f} than $Nank1-6^{\star}$ and $Nank1-5^{\star}$, indicating that repeats 6 and 7 do not add any value to ΔV_{eq}^{f}.

We know that folded helices have no voids in their structure, so unfolding of helices by the use of pressure don't change ΔV_{eq}^{f} in any way. In the case of $Nank1-7^{\star}$, the motifs 2 to 7 are well folded and packed, so, comparing to $Nank1-6^{\star}$ and $Nank1-5^{\star}$, we can hypothesize that the interfaces between repeat 5-6 and 6-7 should not be involved in the ΔV_{eq}^{f} in order to explain their identical ΔV_{eq}^{f}. If these interfaces are not involved in the ΔV_{eq}^{f}, it can mean that there may be no solvent-excluded void at these interfaces. We can observe too that $Nank2-7^{\star}$ has the higher ΔV_{eq}^{f}, even if its ΔV_{eq}^{f} value is similar to that of $Nank1-7^{\star}$. Finally, we observed that $Nank2-7^{\star}$ and NAG2 share the exact same ΔV_{eq}^{u} so that mutation of a single alanine in the second motif has the same effect on ΔV_{eq}^{u} than deleting the first motif. Moreover, the volume of the TSE of NAG2 is significantly larger than that of both $Nank2-7^{\star}$ and $Nank1-7^{\star}$. This indicates that the large volume and expansivity of the TSE of $Nank1-7^{\star}$ come from the second repeat or the interface between repeats 2 and 3. However, the major contributing voids to the equilibrium ΔV_{eq}^{f} must be more centrally located. We performed all-atom MD simulations in order to try to identify these solvent excluded voids.

Our molecular dynamic simulations revealed that repeat 1 undergoes a lot of thermal motions, due to its low stability. Contrary to other repeats which are stable in terms of structure, repeat 1 is continuously changing its shape and never maintain one fold more than 2 ns, as we can observe from the RMSD plot of repeat 1 compared to that of the other repeats. These high motions of the first repeat can't accept the idea that a stable cavity can be formed between motifs 1 and 2, and we can easily imagine that the presence of this motif can perturb too the interface between repeats 2 and 3.

From MD simulations, we determined that the whole cavity volume of $Nank1-7^{\star}$

was of about 160 mL/mol, which is 4 times the experimental one. The evident conclusion is that all the cavities do not account in the ΔV_{eq}^{f}. We can imagine easily that parts of the cavities which are at the interface between motifs will be moving with dynamics but can remain constant on average, and that the external ones at the level of the loops, or at the exterior part of the interfaces, can disappear depending on the frames during the simulations, because water will come in and out, which can explain in a way that these cavities can have on average a smaller volume.

By looking at the water density around the repeats, we have seen that the first, second, sixth and seventh repeats are highly hydrated at the exterior of the structure. And as they are highly hydrated, the interdomain cavities associated with them should be close to water. Previously, we defined the limits of these interdomain cavities leading to the ΔV_{eq}^{f} by a distance to the closest water molecule of 6 Angströms and more. By taking the high exterior hydration level of motifs 1, 2, 6 and 7, we can observe that for these repeats the places where are located the interdomain cavities are too close to water. Hence, empty interdomain cavities do not exist between repeats 1-2, 6-7, what was confirmed by the fact that MD simulations show us interdomain cavities only between repeats 3-4, 5-6 and a little between repeats 2-3. This is consistent with the fact that $Nank1-7^{\star}$, $Nank1-6^{\star}$, $Nank1-5^{\star}$, $Nank2-7^{\star}$ and $Nank3-7^{\star}$ share the same ΔV_{eq}^{f}, as they all contain these cavities. Morevover, it explains too why $Nank4-7^{\star}$ has a lower ΔV_{eq}^{f} than the others, as the cavity between repeats 2 and 3 can't exist.

From MD simulations, we learned that the empty cavities that lead to the ΔV_{eq}^{f} are present only between motifs 3-4, 4-5 and a bit between repeats 2-3, that they are placed only in the upper part of the interface, at about 6 Angströms and more of the closer water molecule. Slight differences between the ΔV_{eq}^{f} of all these proteins, even if they own the same cavities, can be reported to changes in interaction energy between repeats because of the stability difference between each other. We can add to this analysis that thermal motions of the first repeat may change the overall shape of the protein. Like this, when the first motif is present, the protein is more constrained, and water can enter more into the structure, leading to a decrease of the volume of the interdomain cavities, and so to a lower ΔV_{eq}^{f}; and when the first motif is deleted, the protein expands itself, leading to a bigger ΔV_{eq}^{f}.

It is interesting to note that the positions of the interdomain cavities correspond to the interfaces between repeats that are supposed folded in the TSE, as we know from literature [176]. So we can imagine that these cavities appear early in the folding, at the TSE. More, we can hypothesize that the ΔV_{eq}^{f} may have a role in the folding process.

We have seen that there are voids between helices of the same repeat from MD simulations, but that these voids are in majority hydrated. Only a small part of these voids appear to be at more than 6 Angströms of the closest water molecule. It appears

highly coherent to think that the ΔV_{eq}^{f} is made by cavities at the interfaces between repeats 3-4, 4-5 and 2-3, and that these cavities, even slightly changing with the thermal motions of the proteins, are in average of around 10 $mL.mol^{-1}$, as found during the MD simulation analysis.

We have explained in the introduction that NAG5 is a mutant where an alanine was turned into a glycine in the bottom part of the interface between repeats 4 and 5. This mutation increases the void at the interface, but not in the determined interdomain cavity. We have previously observed than ΔV_{eq}^{f} of $Nank1-7^{\star}$ is exactly the same than that of NAG5 (its mutation should increase the cavity at the bottom part of the interface), so again, this provides one more clue that all the cavities are not important, but just the highly dehydrated ones, ie those interdomain cavities between repeats 3,4 and 5, in the central part of the protein. **From all these MD simulations, we can now summarize the hypothesis that the ΔV_{eq}^{f} we observe during the unfolding of these proteins are probably coming from the highly dehydrated cavities, and not from the overall amount of cavities.**

Transition State Ensemble

Analysis of pressure-jump kinetics in terms of the activation volume and rate constant for folding (constraining the unfolding constants by the equilibrium constant obtained from the steady-state data) yielded the activation volumes. All of the deletion constructs exhibited a TSE with a molar volume either equal to, or in most cases, larger than the volume of the folded state at 20°C, whereas the volume of the TSE for $Nank1-7^{\star}$ at this temperature lies between the folded and unfolded state, and only becomes larger than that of the folded state at higher temperature.

The effect is largest upon deletion of either the first N-terminal or C-terminal domain. Indeed, we have observed that deleting the first motif leads to an increase of both the folded and transition states, and deleting repeat 7 from $Nank1-7^{\star}$ leads to an significant increase of the TSE but not of the folded state volume. Further deletions result in small decrease in the volume of the TSE ($Nank1-6^{\star}$ and $Nank1-5^{\star}$) or of the folded state ($Nank2-7^{\star}$, $Nank3-7^{\star}$, $Nank4-7^{\star}$), but they all remain larger than the TSE of $Nank1-7^{\star}$.

Thermal expansivity is a key parameter for understanding volumetric properties of proteins. The temperature dependence of ΔV_{eq}^{u} and the activation volumes, $\Delta V_{f}^{\#}$ and $\Delta V_{u}^{\#}$ corresponds to $\Delta\alpha$, the difference in thermal expansivity between the folded and unfolded state ($\delta\Delta V_{eq}^{u}/\delta T = \Delta\alpha_{u}$), between the transition state and the unfolded state ($\delta\Delta V_{f}^{\#}/\delta T = \Delta\alpha_{f}^{\#}$), and the transition state and the folded state ($\delta\Delta V_{u}^{\#}/\delta T = \Delta\alpha_{u}^{\#}$).

We find the expansivity α of the TSE for $Nank1-7^{\star}$ and $Nank2-7^{\star}$ to be close to that of the unfolded state, while that of $Nank3-7^{\star}$ is equivalent to that of the unfolded state. Hence, the origin of this large TSE expansivity is in repeat 2 or the interface

between repeats 2 and 3. Moreover, the volumes of the TSEs for $Nank2-7^{\star}$ and the NAG2 variant are much larger at 20°C than for $Nank1-7^{\star}$. Even the ΔV_{eq}^{f} is larger for $Nank2-7^{\star}$. This indicates that when the first motif is present, the protein ΔV_{eq}^{f} is restricted as if capping of the first repeat compacts the whole protein leading to smaller volume of the folded state, and that when we delete it, the folded protein and even more the TSE expands as internal constraining interactions are disrupted. This means that in this case, the first motif can change the interaction energy in repeat 2 or between repeats 2 and 3, leading to a compaction of the protein when it is present.

It is now evident that the stability of the protein, and the associated expansivities, also contribute the ΔV_{eq}^{f}.

Here we have important clues to say modulation of the internal void volume occurs in the TSE of these proteins, and that clearly the first motif restrains the structure by a capping effect, for both the folded state structure and the TSE structure while the capping of the seventh repeat restrains the volume of the TSE only. As the capping process involves stability changes, we can see here a correlation between stability and the ΔV_{eq}^{f}. Moreover, the TSE can have a higher volume than the folded state due to its lack of strong internal interactions. Energetically (expansion-wise), the TSE can behave like the unfolded protein, whereas volume-wise it is folded like, i.e. its internal cavities are formed, but the key constraining interactions are not locked-in.

Analysis of the localization of cavities and amino acid sequence

In order to understand why interdomain cavities are located only at certain positions within the structure, we intended an analysis of the correlation between the amino acid sequence and their position. As shown in figure 5.1 A, we can observe that the number of polar and charged amino acids in each repeat are high for repeats 1, 2, 5, 6 and 7, so that they are high around the interfaces 1-2, 5-6 and 6-7, where no cavity takes place.

More, we observe in figure 5.1 B that the hydrophobic amino acids are numerous in motifs 2, 3, 4 and 5, so around the interfaces 2-3, 3-4 and 4-5 where we found the interdomain dehydrated cavities. From this analysis, we can emit the hypothesis that the ΔV_{eq}^{f} is highly correlated with the number of hydrophobic amino acids delimiting the empty (ie not solvated) cavities, and so that the contributions to ΔV_{eq}^{f} are likely to originate in the most hydrophobic regions of proteins.

5.1.3 Pressure effects on the pre-exponential factor

During this study, we have tried to define the physical basis for pressure effects on proteins. To evaluate possible effects of pressure on the pre-exponential term in the rate constant we reformulated our results as linear free energy or Leffler plots.

At $\Delta G = 0$, the folded and unfolded states are energetically identical, so $ln(k_f) =$

- $ln(k_u)$. In this particular way, differences that can happen in $ln(k_f)$ don't come from differences in the energetic level of the energetic barrier, as the system is maintained stable at $\Delta G = 0$, but comes from a pre-exponential term that is relevant to the diffusion coefficient at the energetic barrier [234]. It as been shown that with various perturbations (denaturant, temperature, pH), one can observe changes in broadness of the energetic barrier, that can't be seen with standard experiments based on mutation-induced structural changes. Our linear free energy diagrams revealed significant effects of pressure on the pre-exponential term. Such effects have been interpreted as arising from changes in the shape or broadness of the barrier, or the energetic roughness at the barrier. Further work needs to be done to better understand it, but from now, we have a good starting point with these data for such studies.

5.2 Conclusions

The first clear conclusion of the present work is that hydration of peptide bonds and differential hydration of amino acids do not appear to contribute significantly to the ΔV^f_{eq} for Ankyrin repeat proteins. To the contrary, the main contributing factor to ΔV^f_{eq} seems to be internal solvent excluded voids that have both specific structural and dynamic origins. We have shown, as well, that removal of a repeat has dynamic consequences on the remaining repeats.

We have found that the largest and "driest" voids in $Nank1-7^\star$ are found at the interfaces between repeats 2 and 3, 3 and 4 and 4 and 5, entirely consistent with the preponderant contribution of the central region of the protein to the magnitude of the volume change of unfolding. We interpret the large increase in the magnitude of ΔV^u_{eq} upon deletion of domain 1, as arising from a destabilization of the interactions at the interface between domains 2 and 3, in the interior of domain 2, or both, leading to an expansion of the structure that is manifested by a significant increase in the magnitude of ΔV^u_{eq}. An expansion of protein structure, with decreased tertiary interactions, has been proposed to constitute the first step in their unfolding [235]. The effect of the Nank deletions on the folded state is limited to the repeats 1 and 2, and while the volumetric effects on the TSE are significant, those on the equilibrium ΔV are small. Likewise, the energetic consequences are likely to be rather minor, consistent with the applicability of a simple Ising model to explain the free energy landscape of the full-length protein based on the stabilities of the deletion variants. The increased sensitivity of the volumetric properties of the transition state to the deletions is consistent with our general hypothesis (based on our observation of generally large positive activation volumes for protein folding) that the transition states of proteins contain regions in which key constraining interactions are missing, and that in some cases their overall molar volume under certain conditions of temperature can be larger than the molar volume of the folded state. This was observed for Nank1-7 at 28°C and for most of the

variants. This systematic investigation of the role of size and sequence in determining the magnitude of pressure effects on protein structure and stability provides strong evidence in support of the notion that the major contributing factor to pressure effects is their imperfect internal packing. These packing defects appear to be specifically localized in the 3D structure, rendering pressure effects rather idiosyncratic in character compared to temperature or denaturant effects that depend upon uniformly distributed properties of the structure. Thus, the degree to which pressure disrupts interactions throughout the structure provides information concerning the intrinsic cooperativity of unfolding.

Interestingly, we have also found that the energetics and dynamics of the protein's conformations can significantly modulate their volumetric properties. We show here that loss of interactions can lead to expansion of internal voids, while increased dynamics can minimize internal voids via occlusion by protein side-chains or incursions by solvent. While MD simulations can be used to locate the solvent excluded voids fairly well, it will prove much more difficult to estimate their size, and hence the absolute magnitude of ΔV^f_{eq} since this is strongly dependent on expansivity, dynamics and energetics. Thus, we have determined that contrary to the usual point of view regarding the ΔV^f_{eq}, we have to take this ΔV^f_{eq} into account with a highly dynamic point of view.

In this study, thanks to high-pressure fluorescence and densitometry measurements, we have been able to explain the Chalikian's "Volume Change Paradox", by seeing that the change of sign of ΔV^f_{eq} comes simply from differences in thermal expansivity between the folded and unfolded states. Indeed, the change of sign of ΔV^f_{eq} just depends on the expansivities of the unfolded and folded state that can just cross themselves with temperature. Unfolded state expansivity is well explained by hydration and low energetic and structure constraints [110], while the folded state expansivity appears to result from strong energetic and structural constraints.

Even if this story is far from its end, this work contributes significantly to the understanding of pressure effects.

5.3 Perspectives

5.3.1 Pressure perturbation calorimetry

Clearly, a deeper understanding of the factors determining expansivity of proteins will be required to completely understand pressure effects. Hence, PPC measurements will help. We have only preliminary results for the ankyrin repeat proteins, which were not as well-behaved as we could hope. Despite the fact that $Nank1-7^{\star}$ showed maybe aggregation at the high concentration used for PPC experiments, the other proteins showed conventional α. Because of the aggregation of $Nank1-7^{\star}$, we can postulate all the expansivity values are just shifted to uper values, as the expansivity

of the aggregated unfolded state is higher than if it wasn't aggregated. We can remark first that the expansivity of the unfolded state seems to be identical for $Nank1-6^{\star}$, $Nank1-5^{\star}$ and NAG5 at 50-60°C within the error, so the unfolded state of these three mutants share the same propensity to expansion, even if their sizes are different. This is a clue saying that the unfolded state expansivity don't depends on the exact sequence, as NAG5 and $Nank1-6^{\star}$ differs by one repeat, and as NAG5 and $Nank1-5^{\star}$ have two repeats of difference. Indeed, the expansivity of unfolded states of proteins is rather similar at least at high temperature [109, 83, 110, 236]. We can observe too that the differences between the folded and unfolded states expansivities of $Nank1-6^{\star}$ and $Nank1-7^{\star}$ are quite low. This $\Delta\alpha$ is more important in the case of $Nank1-6^{\star}$, and seems to be twice that of $Nank1-7^{\star}$, confirming the experimental $\frac{\delta\Delta V}{\delta T}$ values obtained by analysis of unfolding monitored by fluorescence and by PPC.

We can see here that all our data combine to the same conclusions : depending on the size of the mutants, $\Delta\alpha$ can change a lot, and, more, can double. The reason of this high degree of changes is not clear for the moment, and may arise from the difference in terms of size, or stability.

NAG5 contains all the repeats, like $Nank1-7^{\star}$, so it is highly surprising that NAG5 does not have the same expansivity than $Nank1-7^{\star}$, and more surprising that NAG5 exhibit the same expansivity of the folded state than $Nank1-5^{\star}$. More, as $Nank1-5^{\star}$ and $Nank1-6^{\star}$ have the same ΔV_{eq}^{f}, it is surprising that they don't share the same expansivity of the folded state. Again, these differences may have an energetic origin.

We have seen that $Nank2-7^{\star}$ has its $\Delta\alpha$ between that of $Nank1-7^{\star}$ and $Nank1-6^{\star}$, and that $Nank3-7^{\star}$ has a $\Delta\alpha$ intermediate between that of $Nank1-5^{\star}$ and $Nank2-7^{\star}$. So it seems that $\Delta\alpha$ does not evolve depending on the size or sequence. This was also true for ΔV_{eq}^{f} and, so that cavities may play a role in determining the expansivity as well as the magnitude of ΔV_{eq}^{f}. This would not be particularly surprising since the two are linked.

All we can affirm is that clearly, the ΔV_{eq}^{f} dependence with temperature comes from the $\Delta\alpha$. And moreover, the calculated values of $\frac{\delta\Delta V}{\delta T}$ agree well with the $\Delta\alpha$ observed in the PPC, but that we don't understand what are the factors of the $\Delta\alpha$. As $\Delta\alpha$ does not change in a manner correlated with the size, we can hypothesize that perhaps cavities may be important.

Indeed, our results on BPTI are consistent with a negligeable contribution of hydration to the magnitude of the folded state expansivity.

As ankyrin repeat proteins were shown not to have a correlation between ΔV_{eq}^{f} and hydration, we decided to use another model : the Bovine Pancreatic Trypsin Inhibitor (BPTI), for which a lot of mutants exists. We had been given two mutants from the lab team of Yutaka Kuroda : BPTI5,55A14GA38V and BPTI5,5521A14GA38V (a mutant based on BPTI5,55A14GA38V for which 21 amino acids of the surface were changed into alanines with changing the 3D structure), thus considerably modulating the po-

larity of its surface and hence its hydration characteristics. I performed PPC on these two mutants for various pH for both BPTI5,55A14GA38V and BPTI5,5521A14GA38V. No differences in the folded state expansivity were observed but ΔV^f_{eq} were different. As the alanines are at the surface, and not buried in the structure, and as the structure of those two proteins are highly similar (we made MD simulations on both and defined the cavities were of identical volumes), it means that alanine clusters can push water away from the protein surface, displacing the first water hydration shell away, inducing a void at this position.

5.3.2 High pressure NMR

Now that we know in the structure where are the real voids leading to a ΔV^f_{eq}, we'd like in the future to understand more how pressure acts at the atomic level on the entire structure to unfold it, and see how these cavities react with pressure. This knowledge, more than giving us answers and parametrization of pressure effects, will provide wealth of knowledge concerning cooperativity. To do so, I began NMR experiments on $Nank2-7^\star$ and $Nank4-7^\star$. $Nank2-7^\star$ was chosen because of the lack of the first motif, which can make troubles during NMR because of its continuous configurational changes, and $Nank4-7^\star$ was chosen because it is the smaller construct of known structure we have. Indeed, ankyrin repeat proteins are known to form multimers at high concentrations, such as the one of NMR experiments, and as all the motifs are of same structure, we hypothesized conserved amino acids can have similar chemical shifts in different repeats, leading to difficulties in the attribution of HSQC. We effectively had aggregation troubles that we partially solved by use of several condition tests of pH, salt concentration and temperature. But HSQC spectrum are still incomplete and further work on it needs to be done. I managed to obtain an incomplete HSQC of $Nank2-7^\star$ and a 70% attribution of $Nank4-7^\star$ (data not shown). As soon this work will be ended, we will use these proteins in NMR with our new pressure system, to follow the pressure-dependent chemical shifts, across the entire structure.

Finally, during this work I used several techniques (high-pressure fluorescence, DSC, PPC) that are all linked with ΔV^u_{eq}, but manage to give data on distinct parameters and temperature scales. PPC shows data about α, ΔV^u_{eq} and $\Delta\alpha$, DSC shows data about ΔC_p, ΔH and ΔS, and high pressure fluorescence shows us data about ΔV^u_{eq} and $\Delta\alpha$. By linking during treatments all these data, we should obtain a complete overview of the whole p,T phase diagram, and a better comprehension of the pressure effects on proteins. This work can be done nowadays but is highly time consuming, as unfortunately no program permits to treat data from all these distinct techniques at the same time by sharing the parameters between each other. The complete p,T phase diagrams in literature can be counted by hand, and their lack is one of the reason why this kind of study is difficult to do and to generalize. In order to make this work easier in the future, and to increase the number of possible studies and researchers

using these techniques, I began to write a program based on work by Makhatadze [109, 83, 110] under MATLAB (screenshot in appendix A). This program already permits to treat DSC and PPC at the same time by sharing parameters and performs treatments automatically. I have already scripted the whole treatment of high pressure fluorescence experiments in a way than now the user just have to put datasets and to configure the treatment to obtain results within a few seconds. This script already saved me days of analysis, and I'm ready to implement it in the MATLAB part of the program.

Part V

Appendix

Appendix A

Developing new tools

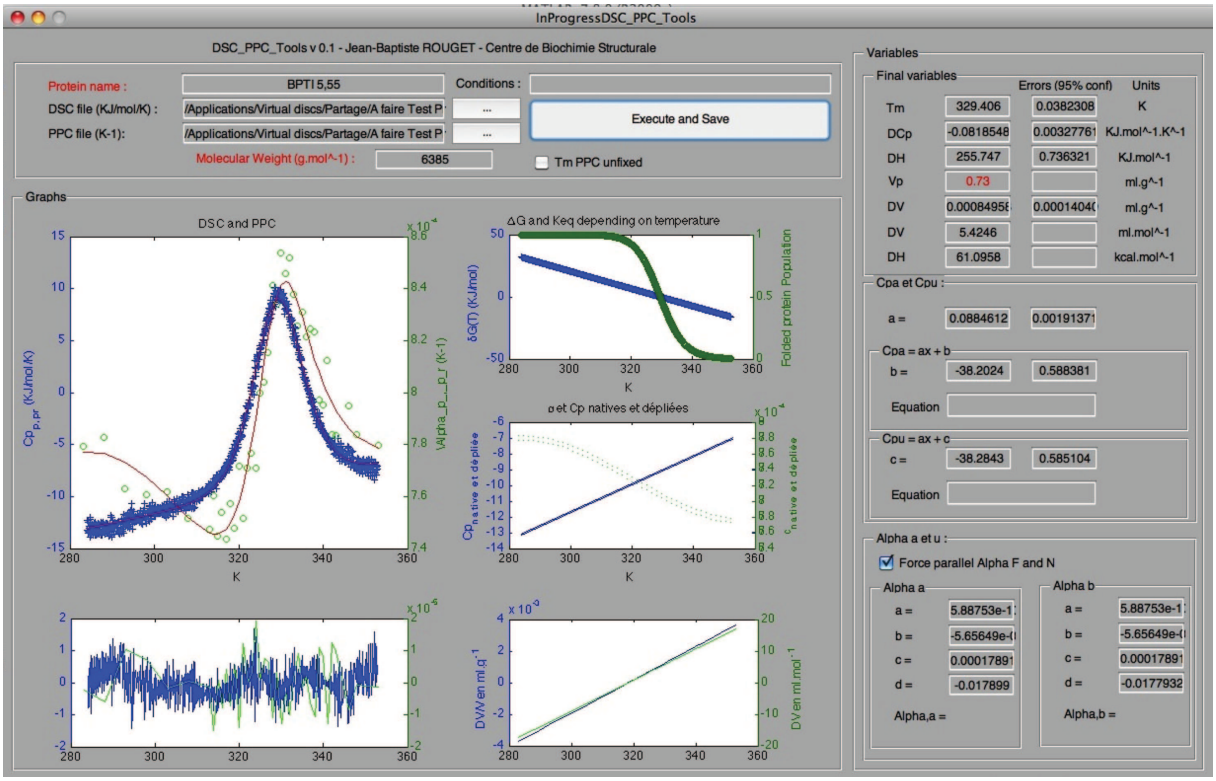

Figure A.1: Screenshot of home-made Matlab program for DSC-PPC treatments.

Part VI

Contributions

J|A|C|S
COMMUNICATIONS

Published on Web 10/26/2007

V_i -Value Analysis: A Pressure-Based Method for Mapping the Folding Transition State Ensemble of Proteins

Lally Mitra,[†] Kazumi Hata,[‡] Ryohei Kono,[‡] Akihiro Maeno,[‡] Daniel Isom,[§] Jean-Baptiste Rouget,[¶,#] Roland Winter,[†] Kazuyuki Akasaka,[‡] Bertrand García-Moreno,[§] and Catherine A. Royer*[,¶,#]

Department of Chemistry, Physical Chemistry I, Otto-Hahn Str. 6, University of Dortmund, D-44227 Dortmund, Germany, Department of Biotechnological Science, School of Biology-Oriented Science & Technology, Kinki University, Kinokawa, Wakayama, 649-6493 Japan, Department of Biophysics, Johns Hopkins University, Baltimore, Maryland 21218, CNRS, UMR5048, Centre de Biochimie Structurale, Montpellier F-34090, France, and INSERM, U554, Montpellier F-34090, France

Received June 1, 2007; E-mail: catherine.royer@cbs.cnrs.fr

We introduce an approach to mapping the structure of the folding transition state ensemble (TSE) of proteins, based on the pressure dependence of protein stability. Pressure leads to the unfolding of most proteins at ambient temperatures because the specific volume of the unfolded state is smaller than that of the folded state. Although a fully quantitative parametrization of volume changes accompanying protein conformational changes remains one of our main objectives, recent advances based on pressure perturbation calorimetric studies[1] have confirmed our earlier conclusions[2] that the relatively small, negative ΔV of unfolding measured near ambient temperatures is the result of positive contributions from increased hydrophobic hydration and negative contributions from increased hydration of polar moieties and the hydration of cavities that are opened to solvent upon unfolding. Thus, overall, the volume changes reflect the change in hydration of the protein upon unfolding. The effect of pressure on the folding and unfolding rate constants depends upon the activation volumes and hence on the relative hydration of the TSE with respect to the folded and unfolded states. Here we use Staphylococcal nuclease (SNase) to illustrate how pressure can be used evaluate the state of hydration of the folding TSE and hence to map its structure.

Since it was first introduced in the late 1980s, ϕ-value analysis has provided considerable insight into the structure of the TSE in the protein folding reaction and into the overall energetics and landscape of protein folding.[3−6] The ϕ-value analysis requires that the mutations destabilize the TSE enough for the effect to be measurable, but not sufficiently to alter the structure of the TSE significantly. In contrast, our approach based on activation volumes relies on significant disruption of the TSE structure by mutations to identify structured regions in the WT TSE.

Previously,[2,7] we have shown that the activation volume for folding of WT SNase is large and positive, and hence that the rate-limiting step, as for a number of other proteins,[8−14] in folding involves significant dehydration. We refer to this measure of the fractional dehydration of the TSE with respect to the unfolded state as the V value. It is calculated as the ratio of the folding activation volume, ΔV^{*}_{f}, over the total equilibrium volume change upon folding, ΔV°_{f}.

$$V = \Delta V^{*}_{f}/\Delta V^{\circ}_{f} \quad (1)$$

Unlike m values, which rely on denaturant binding to estimate solvent-exposed surface area, this V value is directly related to the fractional change in hydration. A V value near 1 corresponds to a TSE that is as dehydrated as the folded state.

Recent pressure perturbation studies on a series of variants of a stabilized form of SNase,[15−17] in which valine 66, which is buried in the hydrophobic core of the protein, was substituted by ionizable residues, revealed that their TSE, unlike that of WT SNase, was highly hydrated. For example, the variant with the V66K substitution has a V value ~0.25 compared to 0.88 for WT SNase.[18] To compare these variants with the WT, we introduce the V_i value, defined as the ratio of the V values for the variant with the internal ionizable group and for the WT protein:

$$V_i = V^{x}/V^{\mathrm{WT}} \quad (2)$$

We have now performed a systematic study of variants in which several internal positions were substituted with Lys. Under the conditions of pH that were used for these studies, the Lys residues are neutral when they are internal and charged when they are in water (see pK_a values in Table 1, Supporting Information). Substitution of internal positions that are ordered in the WT TSE with Lys is expected to disrupt the TSE structure because the Lys residues near neutral pH are charged, thus they promote structural states where they can remain charged by virtue of being exposed to bulk water. This should lead to completely different pressure-dependent kinetics. The sites that were studied were selected to sample different interfaces between elements of secondary structure (Figure 1A). V23 and V74 are both in the β-barrel; their side chains are in the most hydrophobic region of SNase. T62 and V66 are on the same α-helix, and they are also members of the hydrophobic core. L36 and L103 are in the region between the α- and β-subdomains, and L125 is on the C-terminal α-helix. By far UV−CD and/or X-ray crystallography, these variants have been shown to have structures that are indistinguishable from the WT.[16,17,19−21] They undergo cooperative unfolding transitions.[22] In the variants studied here, the internal Lys residues titrate with highly perturbed pK_a values, consistent with the neutral side chain of the Lys residues being buried inside the protein.

The pressure-jump fluorescence relaxation profiles for the variants were analyzed with single exponential decays. The plots of ln τ versus pressure (τ, relaxation time) were fitted for the folding (or unfolding) rate constant and the folding (or unfolding) activation volume, using as constraints the equilibrium constant and volume change obtained from the equilibrium unfolding profile, as described previously[17] (e.g., Figure 1C and Table 1, Supporting Information). Because the variant proteins have high stability, GuHCl was added to some of them at low concentration to allow unfolding in the range of P that is experimentally accessible. We have shown previously that the equilibrium and activation volume changes for SNase and many of its variants are independent of pH and GuHCl up to 1 M.[23,24] The equilibrium ΔV°_{f} values for all the variants are within the range of 65 ± 6 mL/mol except for those at positions 62 and 36. Relaxation at position 62 was extremely slow (several hours), and hence the results were somewhat less accurate, but clearly all of the effect of pressure was on the unfolding rate constant, which increased with pressure. The response of the V74K

[†] University of Dortmund.
[‡] Kinki University.
[§] Johns Hopkins University.
[¶] Centre de Biochimie Structurale.
[#] INSERM.

10.1021/ja073576y CCC: $37.00 © 2007 American Chemical Society

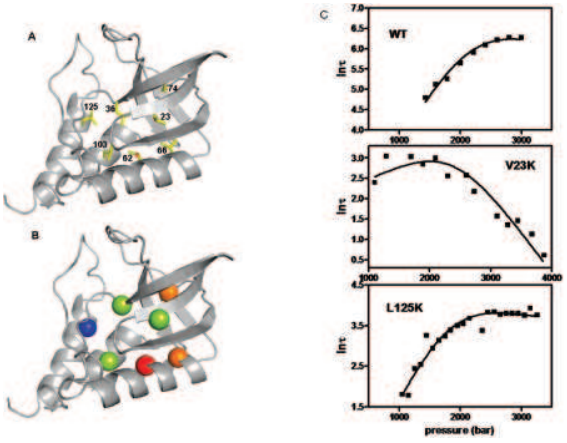

Figure 1. (A) Positions in SNase that were substituted with Lys residues. (B) High (blue, 1), low (red, orange, 0−0.3), and intermediate V_i values (green, 0.5−0.7). (C) Plots of the ln τ versus pressure for WT and two of the Lys-containing variants. Lines through points correspond to fits of the data for the activation volumes and rate constants as described in ref 17. All experiments except WT were performed with a hyperstable variant of SNase known as Δ+PHS. Structures were rendered using PyMol 0.99c6.

variant to pressure perturbation was comparable to that of V66K.[18] Like the T62K variant, the large negative activation volume for unfolding for both V74K and V66K, as opposed to the large positive activation volume for folding observed for WT SNase, implies that, in terms of hydration, the folding TSEs of these variants are much closer to the unfolded than to the native state in terms of their fractional hydration. The L36K, V23K, and L103K variants exhibited an intermediate behavior, with pressure leading to both an increase in the unfolding rate constant and a decrease in the folding rate constant. At position 36, even if there is some ground state (unfolded state) perturbation by the substitution, as suggested by the much larger negative equilibrium volume change upon unfolding, this would only lead to a small decrease in the calculated V_i value and not alter the intermediate nature of the effect of this substitution. Finally, the L125K variant exhibited behavior similar to that of WT SNase, indicating that its TSE is nearly identical to that of the wild type protein, that is, largely dehydrated relative to the unfolded state.

The V_i values are superimposed on the structure of SNase in Figure 1B. It can be seen that the sites with the lowest V_i values (red, red-orange), corresponding to the highest degree of perturbation of the TSE by the internal Lys residue, are in the β-barrel and in the α-helix that packs against it. The burial of unpaired ionizable residues by substitution of one member of a buried ion pair with a neutral side chain likewise destabilized the TSE of barnase.[25] Substitution of residues 23, 36, and 103 with Lys disrupts the TSE to a lesser extent than substitutions in the core (green in Figure 1B). Lys125 (in subdomain 2) has no effect on the structure of the TSE (blue), suggesting that the C-terminal helix of SNase does not participate in the transition state, and that it folds against the more stable β-subdomain after passage of the barrier. These studies suggest that the structure of the TSE of WT SNase includes part of the β-barrel and part of the first α-helix (residues 55−67). In the WT TSE, these elements of secondary structure are sufficiently well ordered to exclude solvent from the interior of the hydrophobic core. The region linking this hydrophobic core to subdomain 2 appears to adopt an ensemble of conformations with an intermediate probability to be folded in the TSE, while the C-terminal α-helix remains disordered.

Our approach to mapping the TSE of SNase differs from the much more widely used ϕ-value analysis in that the goal of ϕ-value analysis is to perturb the TSE energetically without perturbing its structure significantly. More often than not, this is actually possible, and for this reason, ϕ-value analysis has proven very useful. An alternative to ϕ-value analysis is ψ-value analysis[26] in which the TSE structure is stabilized by cross-linking, which is considered to occur only when the structure pre-exists. In the V_i-value analysis introduced in this paper, the goal is to make substitutions that are sufficiently destabilizing to disrupt the TSE in the mutant, especially at sites that would be ordered in the TSE of the wild type protein. Use of ionizable residues was essential for this purpose. This study was possible owing to the availability of the highly stable Δ+PHS variant of SNase, which tolerates substitutions of internal positions with ionizable groups.[22] In other protein systems, a V_i-value analysis will likewise require enhancement of the stability of the protein by mutation or solution conditions. The unique advantage of V_i-value analysis is that it conveys direct information about the state of hydration of the TSE, which has been recognized as a key factor in the protein folding transition.[27]

Acknowledgment. This work was supported by NSF Grant MCB-0212414 and NIH GM-061597 to B.G.-M., by INSERM (C.A.R.), by DFG and EU/Land NRW to R.W., and by the JSPS Core-to-Core Program (K.A.).

Supporting Information Available: Data acquisition protocol, data analysis, complete results table including the rate constants, examples of the fits of the relaxation profiles to single exponential decays, and ln τ versus pressure plots for all of the tested variants. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Mitra, L.; Smolin, N.; Ravindra, R.; Royer, C.; Winter, R. *Phys. Chem. Chem. Phys.* **2006**, *8* (11), 1249−1265.
(2) Panick, G.; Vidugiris, G. J.; Malessa, R.; Rapp, G.; Winter, R.; Royer, C. A. *Biochemistry* **1999**, *38* (13), 4157−4164.
(3) Fersht, A. R.; Matouschek, A.; Serrano, L. *J. Mol. Biol.* **1992**, *224* (3), 771−782.
(4) Fersht, A. R. *Philos. Trans. R. Soc. London, Ser. B* **1995**, *348* (1323), 11−15.
(5) Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254* (2), 260−288.
(6) Serrano, L.; Matouschek, A.; Fersht, A. R. *J. Mol. Biol.* **1992**, *224* (3), 805−818.
(7) Vidugiris, G. J.; Markley, J. L.; Royer, C. A. *Biochemistry* **1995**, *34* (15), 4909−4912.
(8) Desai, G.; Panick, G.; Zein, M.; Winter, R.; Royer, C. A. *J. Mol. Biol.* **1999**, *288* (3), 461−475.
(9) Jacob, M. H.; Saudan, C.; Holtermann, G.; Martin, A.; Perl, D.; Merbach, A. E.; Schmid, F. X. *J. Mol. Biol.* **2002**, *318* (3), 837−845.
(10) Kitahara, R.; Royer, C.; Yamada, H.; Boyer, M.; Saldana, J. L.; Akasaka, K.; Roumestand, C. *J. Mol. Biol.* **2002**, *320* (3), 609−628.
(11) Mohana-Borges, R.; Silva, J. L.; Ruiz-Sanz, J.; de Prat-Gay, G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (14), 7888−7893.
(12) Pappenberger, G.; Saudan, C.; Becker, M.; Merbach, A. E.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (1), 17−22.
(13) Silva, J. L.; Foguel, D.; Royer, C. A. *Trends Biochem. Sci.* **2001**, *26* (10), 612−618.
(14) Tan, C. Y.; Xu, C. H.; Wong, J.; Shen, J. R.; Sakuma, S.; Yamamoto, Y.; Lange, R.; Balny, C.; Ruan, K. C. *Biophys. J.* **2005**, *88* (2), 1264−1275.
(15) Denisov, V. P.; Schlessman, J. L.; García-Moreno, E. B.; Halle, B. *Biophys. J.* **2004**, *87* (6), 3982−3994.
(16) García-Moreno, B.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spencer, D. S.; Stites, W. E. *Biophys. Chem.* **1997**, *64* (1−3), 211−224.
(17) Stites, W. E.; Gittis, A. G.; Lattman, E. E.; Shortle, D. *J. Mol. Biol.* **1991**, *221* (1), 7−14.
(18) Brun, L.; Isom, D. G.; Velu, P.; García-Moreno, B.; Royer, C. A. *Biochemistry* **2006**, *45* (11), 3473−3480.
(19) Dwyer, J. J.; Gittis, A. G.; Karp, D. A.; Lattman, E. E.; Spencer, D. S.; Stites, W. E.; García-Moreno, E. B. *Biophys. J.* **2000**, *79* (3), 1610−1620.
(20) Karp, D. A.; Gittis, A. G.; Stahley, M. R.; Fitch, C. A.; Stites, W. E.; García-Moreno, E. B. *Biophys. J.* **2007**, *92* (6), 2041−2053.
(21) Nguyen, D. M.; Leila, R. R.; Gittis, A. G.; Lattman, E. E. *J. Mol. Biol.* **2004**, *341* (2), 565−574.
(22) Isom, D. G. Ph.D. Dissertation, Johns Hopkins University, 2007.
(23) Vidugiris, G. J.; Truckses, D. M.; Markley, J. L.; Royer, C. A. *Biochemistry* **1996**, *35* (12), 3857−3864.
(24) Royer, C. A.; Hinck, A. P.; Loh, S. N.; Prehoda, K. E.; Peng, X.; Jonas, J.; Markley, J. L. *Biochemistry* **1993**, *32* (19), 5222−5232.
(25) Tissot, A. C.; Vuilleumier, S.; Fersht, A. R. *Biochemistry* **1996**, *35* (21), 6786−6794.
(26) Sosnick, T. R.; Krantz, B. A.; Dothager, R. S.; Baxa, M. *Chem. Rev.* **2006**, *106* (5), 1862−1876.
(27) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389−415.

JA073576Y

DOI: 10.1002/cphc.200((will be filled in by the editorial staff))

Towards a Quantitative Understanding of Protein Hydration and Volumetric Properties

Lally Mitra [a], Jean-Baptiste Rouget[c], Bertrand Garcia-Moreno [b], Catherine A. Royer [c], and Roland Winter [a]

In this work, we probe by pressure perturbation calorimetry (PPC), the coefficient of thermal expansion, the volumetric and the hydration properties of variants of a hyperstable variant of staphylococcal nuclease (SNase), Δ+PHS. The temperature dependent volumetric properties of the folded and unfolded states of the wild type protein were calculated with previously published data. The present PPC results were then used to interpret the volume diagram and expansivity at a molecular level. We conclude that the expansivity of the unfolded state is, to a first approximation, temperature independent, while that of the folded state decreases with increasing temperature. Our data suggest that at low temperature the defining contribution to ΔV comes mainly from excluded volume differences and ΔV for unfolding is negative. In contrast, at high temperatures, differential solvation due to the increased exposed surface area of the unfolded state and, in particular, its larger thermal volume linked to the increased conformational dynamics of the unfolded state ensemble takes over and ΔV for unfolding eventually becomes positive.

Introduction

The physical-chemical properties of proteins determine their function and as such have been the object of intense research for many years. Indeed, major progress in the understanding of protein structure, thermodynamics, and dynamics as well as their interactions has been made thanks to advances in experimental and computational approaches. [1] Despite this gain in fundamental understanding, a complete description of the factors that control these properties has not been achieved. In particular, understanding of the dependence of protein stability on pressure, as well as the characterization of the role of solvent in controlling protein conformational transitions and stability remain to be accomplished. [1-5]

Clearly, the exquisite specificity of protein structure, dynamics and function is intimately linked to the interactions of the protein with its solvent, water. [1] During the 1970's and 80's, the fundamental basis for the temperature dependence of protein stability and conformational changes, the heat capacity change upon unfolding, ΔC_p, was revealed. [2] Heat and cold denaturation were clearly attributed to the significant change in ΔC_p, leading to entropy driven unfolding at high temperature and enthalpy driven unfolding at low temperature. The amount of hydrophobic surface area that is removed from interaction with water was shown to be proportional to the magnitude of the loss of heat capacity associated with structural transitions involved in protein folding or function. [3] Interestingly, it has been shown that negative ΔC_p effects are a ubiquitous feature of the thermodynamics of biomolecular interactions in aqueous solution as well. [6]

Much less is known about the volumetric properties of proteins, however, and these govern the effects of pressure on stability. Differences in hydration that accompany conformational changes in proteins are known to give rise to differences in their volumetric properties. [3-11] Hence, the study of pressure effects on proteins should, in principle, provide insight into these changes in hydration and the energetics and structural role of the solvent in determining the energy landscape of proteins. [12] Despite these prospects, using volumetric properties to understand the role of solvent in protein conformational transitions has been hindered due to a lack of fundamental and quantitative understanding of volume changes. Indeed, while hydrophobic model compounds capture to a large extent the effects of temperature on protein structure, which are linked to the changes in heat capacity that accompany order to disorder transitions of proteins, [2] they are inappropriate to explain volumetric properties.

The specific volume of the folded state of proteins is generally larger than that of other more disordered and hydrated conformations, including the unfolded state ensemble. [3,13] Recently, evaluation of protein hydration by pressure perturbation calorimetry (PPC) revealed the strong temperature dependence of the thermal expansivity of proteins, as well as the opposing contributions of hydrophobic and polar hydration to the expansivity term. [7-9] However, no details are known about the general behavior of the temperature dependence of the volume of a protein in its folded and unfolded state and hence also not on the temperature dependence of the volume change which occurs upon unfolding, ΔV.

In this work, we probe the volumetric and hydration properties of variants of staphylococcal nuclease (SNase) by PPC. Recent studies on a series of variants with internal ionizable residues at position 66 in the hydrophobic core of a hyperstable SNase known as, Δ+PHS, demonstrated essentially identical structures for the folded form [14] (Figure 1). Under conditions of pH where

[a] *Prof. Dr. R. Winter, L. Mitra*
Dortmund University of Technology
Department of Chemistry
Physical Chemistry I – Biophysical Chemistry
Otto-Hahn Str. 6
D-44227 Dortmund, Germany
Fax: +49 231 755 3901
E-mail: roland.winter@tu-dortmund.de
[b] *Prof. Dr. B. Garcia-Moreno*
Department of Biophysics
Johns Hopkins University, Baltimore MD 21218, USA
Fax: (410) 516-4118
E-mail: bertrand@jhu.edu
[c] *Prof. Dr. C. Royer, J.-B. Rouget*
3 CNRS, UMR5048, Centre de Biochimie Structurale
F-34090 Montpellier, France
Fax: +33 467 417913
E-mail: catherine.royer@cbs.cnrs.fr

the internal ionizable group at position 66 is neutral, it is buried in the hydrophobic core. Changes in pH that result in ionization of these residues result in local modifications of the structure. [15] We use our present PPC results on these variants, along with previous results on WT SNase [17] to calculate quantitatively the temperature dependent volumetric properties of the folded and unfolded states of SNase.

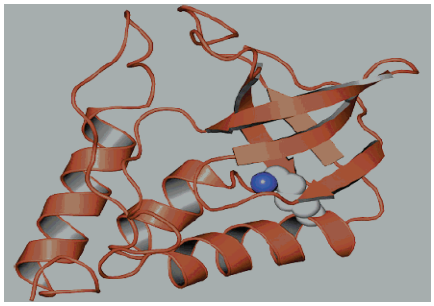

Figure 1. Schematic of the structure of the SNase V66K variant with K66 in a space-filling representation. [16]

Results

For comparison, Fig. 2 presents previously published DSC and PPC curves of wild type (WT) SNase dissolved in 10 mM phosphate buffer at pH 5.5, from which $\alpha(T)$, the midpoint of thermal unfolding (T_m) and the relative volume change upon unfolding, $\Delta V/V$, are obtained (see Experimental Section). [17] T_m is also obtained from the DSC experiments. All thermodynamic quantities are summarized in Table 1. The protein begins to unfold at ~43 °C and the temperature of the unfolding midpoint, T_m, is 51.5 °C.

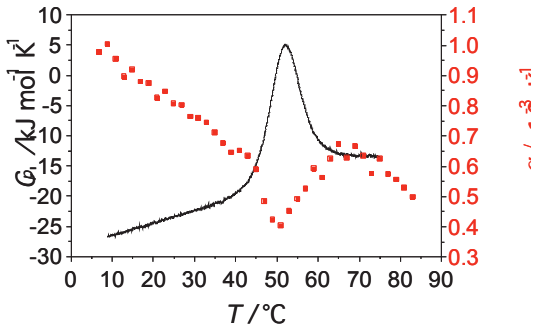

Figure 2. In black: DSC trace of a 0.2 wt% (2 mg·mL^{-1}) solution of SNase (background corrected) at pH 5.5 in 10 mM phosphate buffer (scan rate: 40 K·h^{-1}). In red: Temperature dependence of the apparent thermal expansion coefficient, $\alpha(T)$, of SNase in 10 mM phosphate buffer at pH 5.5 as obtained from the PPC experiments recorded at a protein concentration of 4 mg·mL^{-1}. The thermal expansivity is calculated from the heat released upon small pressure perturbations in the context of a differential scanning calorimetry experiment (see Experimental Section).

In the native state of the protein, between 10 and 40 °C, the apparent thermal expansion coefficient of the protein, $\alpha(T)$, decreases almost linearly from $1.0 \cdot 10^{-3}$ K^{-1} to a value near $0.65 \cdot 10^{-3}$ K^{-1}. At approximately 51.7 °C, where the DSC curve $C_p(T)$ exhibits its maximum, a distinct dip is observed. Above that temperature, α increases again to a relatively constant level in the post-transition temperature range. Above 70 °C, α then decreases again.

High α values at lower temperatures and their steep decrease up to 40 °C are indicative of the presence of a large number of structure-breaking hydrophilic side chains on the surface of SNase. [7] As the temperature increases, water molecules are released from the protein surface and the value of the thermal expansivity is reduced. Since SNase has an extraordinarily large number of ionisable residues (39 %), it exhibits large values of α and $d\alpha/dT$ in its native state. The relative volume change upon unfolding, $\Delta V/V$, as obtained from the area under the transitional peak, is calculated by integrating the $\alpha(T)$ transition curve (Eq. 4); $\Delta V/V = -1.5 \cdot 10^{-3}$. The absolute volume change upon unfolding, ΔV, can be calculated from $\Delta V/V$ using the molar mass of SNase (16812 Da) and its partial specific volume (0.754 mL·g^{-1}), which yields $\Delta V = -19$ mL·mol^{-1} and is of the order of –0.15 %.

Figure 3 exhibits the PPC profiles of the corresponding SNase Δ+PHS variants at different pH values. For better visualization of the unfolding region, the temperature region of the unfolding transition, as obtained by (the more accurate) DSC traces, is indicated in the plots as well. The thermal expansivity profiles of the V66K and V66E variants of Δ+PHS at pH values where the internal ionisable residues is neutral, 7 and 4.5, respectively, are very close to that of WT SNase. In contrast, at pH values where the internal Lys-66 (pH 4.5) or Glu-66 (pH 9) are charged, the α values at low temperature decrease significantly.

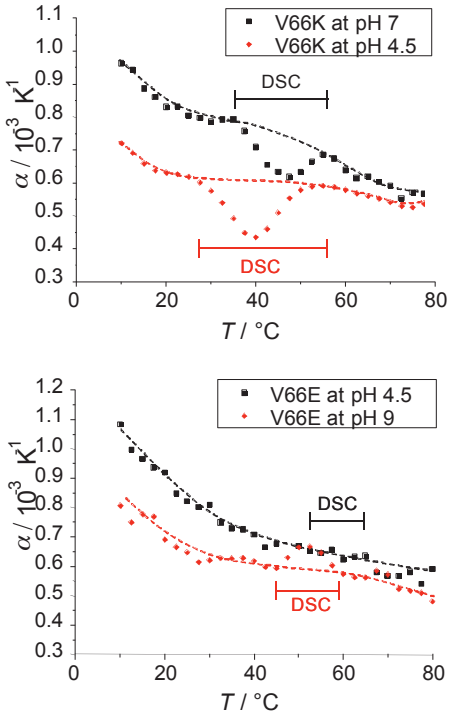

Figure 3. Coefficient of thermal expansion (α) derived from PPC scans of SNase variants. Top: Δ+PHSV66K at pH 7 and pH 4.5. Bottom: Δ+PHS V66E at pH 4.5 and pH 9. The dashed lines correspond to the baselines used for integrating out $\Delta V/V$ of the unfolding transition. The integration limits have also been determined by DSC and are given in brackets (|—|).

The expansivity of the highly stable Δ+PHS protein was also measured (Fig. 4). It can be seen that the expansivity of this hyperstable variant is quite low at low temperature, compared to the WT SNase (Fig. 2) and to variants with internal ionizable groups under conditions of pH where the groups are not charged (Fig. 3). Moreover, this expansivity is essentially constant over the entire temperature range, except at the unfolding transition. We note that at the high temperature of the unfolding transition (~70°C), the cross-over point has been passed and the volume change for unfolding this highly stable variant is positive.

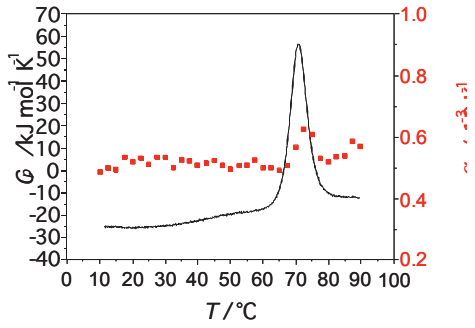

Figure 4. DSC (in black) and PPC (in red) data of the background Δ+PHS variant of SNase at pH 7.

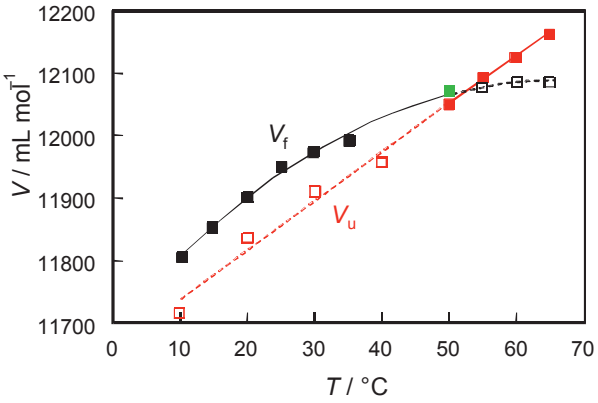

Figure 5. Molar volume *vs.* temperature plots for the folded (V_f) and unfolded (V_u) states of SNase calculated from direct densitometric measurements of WT SNase (■ and ■), obtained from the temperature dependent high pressure unfolding studies (yielding $V(T,p)$) and PPC measurements of the unfolding transition (yielding $\Delta V/V$) (green square, ■), and from the densitometric V_f values and the temperature dependence of ΔV, yielding V_u (squares □), respectively. The dashed line represents a linear fit to these V_u values and is consistent with a linear extrapolation of the V_u values obtained from densitometry at high temperature.

In Fig. 5 we show what is, to our knowledge, the first experimental plot of the volume of both the folded and unfolded states of a protein. Assuming an essentially constant value for the expansivity of the unfolded state (dashed red line in Fig. 5), prior results from densitometric measurements on WT SNase [18] can be used to calculate the complete volume profiles of the folded and unfolded states of SNase. The densitometric studies yielded directly the volume V of SNase as a function of temperature for the folded state (below the transition temperature, black squares (■), connected by full black line) and for the unfolded state (above the transition temperature, red squares (■), connected by full red line). [18,19] It can also be seen from Fig. 5 that the increase in V of the native state of SNase (V_f) with temperature is not linear. Indeed, the value of α for the folded state decreases significantly as the temperature increases, while at high temperature the expansivity of the unfolded state appears to be a constant. Taking the values of ΔV obtained in previous pressure unfolding studies carried out at multiple temperatures [19], we were also able to calculate the volume of the unfolded state (V_u) at low temperature (open red squares (□)), which to a first approximation appears to increase linearly over this temperature range as well, with approximately the same slope as over the high temperature range. Thus, we may conclude that the expansivity of the unfolded state is, to a first approximation, temperature independent, while that of the folded state is not. Hence, $\Delta\alpha$, the difference in expansivity between the two states is most likely not constant. The green square just below 50 °C corresponds to the volume of the folded state calculated from the volume of the unfolded state plus the volume change for unfolding, ΔV, obtained from pressure perturbation calorimetric measurements of the volume change upon unfolding at the transition temperature (Fig. 2). [18] Beyond this point, we do not know the value of the volume and expansivity or the folded state. The black dashed line (and black squares (□)) represents the extrapolation of a polynomial fit of second order to the V_f *vs.* T curve at lower temperature.

The volume diagram presented above clarifies a number of outstanding points of confusion about the effects of temperature and pressure on protein structure and the nature of the volume change upon unfolding. Indeed, while most reported values of ΔV are negative, positive values for ΔV have been reported as well. [20] Figure 5 shows quite clearly that the volume difference between the folded and unfolded states of SNase and its mutants depends strongly on temperature, being relatively large and negative at low temperature, but decreasing in absolute value and eventually changing sign at high temperature. Since most reported ΔV values were measured around room temperature or lower, it is understandable that they are typically negative. Stabilizing mutations (i.e., Δ+PHS *vs.* WT SNase, Table 1) shift the unfolding to higher temperature, and hence the volume change measured by PPC at the transition temperature decreases in absolute value, even becoming positive.

Table 1. Thermodynamic parameters of WT SNase and SNase variants.

Variant	pH	T_m / °C	ΔV / mL·mol^{-1}
SNase WT	5.5	51.5 ± 1	-19 ± 1
	7	50.6 ± 1	-20 ± 1
SNase WT + 1.5 M sorbitol	5.5	58.6 ± 0.5	-5 ± 0.5
SNase WT + 1.5 M urea	5.5	39.1± 0.5	-56 ± 2
Δ+PHS	7	73.8 ± 2	+4 ± 2
Δ+PHS+V66K	7	47.0 ± 1	-22 ± 1
Δ+PHS+V66K	4.5	47.0 ± 1	-16 ± 1
Δ+PHS+V66E	4.5	61.0 ± 1	0 ± 2
Δ+PHS+V66E	9	48.4 ± 2	+15 ± 1

Discussion

Much of our present understanding of the contribution of differential hydration volume has come from recent studies of model compounds and proteins based on pressure perturbation calorimetry. Lin and co-workers [6] measured the coefficient of thermal expansion, α, of individual amino acid side chains (by subtracting the value obtained for glycine) (Fig. 6a). They observed that α the expansivity value for polar amino acids is large and positive at low temperature, and decreases dramatically between 5 and ~50 °C. Quite the opposite is observed for non-polar amino acid side chains, which exhibit large negative α values at low temperature, which increase drastically between 5 and ~50 °C. We have carried out similar studies following a host-guest scheme, in which we subtracted the expansivity measured for a glycine tripeptide, from peptides in which the central glycine residue was substituted with the residue of interest (Fig. 6b). [9] The relative changes of the data from these two studies are similar. We observe that the magnitude of the negative expansivities for the non-polar amino acids is proportional to their hydrophobicity. A qualitatively similar behavior has been observed in densitometric studies of tripeptides. [11]

The temperature dependence of the coefficient of thermal expansion of pure water is also shown in Figure 6a (black line). In particular it is noteworthy that α of pure water is negative below

4 °C. This observation may help to interpret the expansivity data on a molecular level. A negative expansivity means that the density increases upon heating. We may use the same reasoning for the non-polar amino acid residues, assuming that α is largely controlled by the hydrational properties of the amino acid. As their solutions in water are heated, part of the hydrating waters are released to the bulk where they occupy a smaller partial molar volume, resembling ice melting. At these low temperatures, the density of the waters hydrating the non-polar residues is lower than in the bulk, or ice-like. A large positive expansivity for the polar amino acids indicates significant hydration and a degree of "electrostriction" of the hydrating water molecules around charged and polar moieties, leading to a higher density than the bulk. Upon heating, part of these molecules are released gradually into the less dense bulk. With increasing temperature, the differences between bulk and hydrational water properties diminish, and less and less water is released from the hydration shell to the bulk, thus leading to the observed decrease of α. This also means that the direct hydration contribution to the volume and expansivity is most prominent at low temperature and decreases with increasing temperature. Proteins, being composed of a combination of polar and non-polar moieties, more or less exposed to solvent depending upon their conformation, should exhibit expansivities that correspond, in part, to a weighted combination of the expansivities of these moieties. As we have seen above, these contributions tend to compensate for each other, yielding expansivities and a contribution to the volume change which is globally quite small.

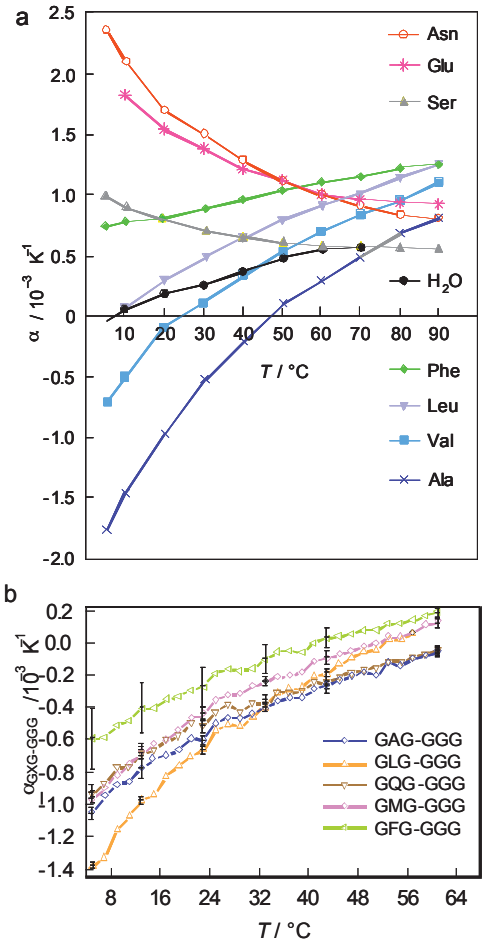

Figure 6. a) PPC data taken from Lin and co-workers [6] for polar and non-polar amino acids (calculated with respect to the signal obtained from glycine). b) Similar studies obtained by us for non-polar amino acids using a host-guest approach. [9]

Based on these model compound studies, the thermodynamic parameters and expansivity profiles of the ionisable variants of Snase can be interpreted at a molecular level in light of the structural differences between the neutral and ionized states of the variants. The smaller α values at low temperature for the V66K and V66E variants of Δ+PHS SNase upon ionization can be explained by the increased hydrophobic character of the hydration (for explanation see also below) upon the local (partial) unfolding that results from ionization. In addition, a decrease of the expansivity of the void volume might play a role. After the dip in α at the unfolding transition due to the volume decrease upon unfolding, the expansivity of the unfolded states is fairly constant. The lower value for the expansivity at low temperature observed for the variants when the internal ionizable groups are charged, along with the nearly constant expansivity for the unfolded states indicates that the volumetric characteristics of hydrophobic and polar hydration compensate each other to a large extent.

The above PPC results for the hyperstable variant of SNase, along with molecular dynamics computer simulations [8] highlight the notion that the stability and dynamics of the various states of the protein, in addition to the degree and type of hydration, may be crucial in determining their expansivities. Indeed, this contribution is responsible for the "thermal volume" increase (i.e., the transient void volume around the solute that results from thermally activated mutual motions of contacting solute and solvent molecules [3]) with increasing temperature. The highly stable Δ+PHS SNase exhibits a very low and constant value of α over the entire temperature range (except for the unfolding transition). This variant, in the absence of any internal ionizable residues, is extremely stable (unfolding occurs above 70 °C), and NMR relaxation experiments (unpublished data) have revealed very limited dynamics on the nanosecond to millisecond timescale. These studies suggest that the value of the expansivity for the folded state results from a combination of surface hydration properties and structural flexibility.

Based on this ensemble of volumetric information, we can reasonably assume two major contributions to the difference in specific volume between the unfolded and folded states of a protein, ΔV. The first contribution arises from the decrease in solvent excluded volume when the tightly, but of course not perfectly, packed folded structure (with packing defects (voids) of the order of 25 % and internal cavities (defined as empty cavities which can accommodate at least one water molecule) of the order of 1 %) is disrupted. Water molecules enter this volume, hence decreasing the overall volume of the protein-solvent system. The magnitude of this contribution depends on the specific properties of the protein, both in its folded and unfolded state. A second contribution, which could not be quantified yet, arises from the change in the volume of the water molecules that hydrate the newly exposed protein surface area, relative to their volume in the bulk liquid.

Taking these contributions into account, the following scenario emerges for $\Delta V(T)$. At low temperature, the differences in expansivity for the polar and non-polar amino acid residues are close to maximal. Since the expansivity (releasing water to the bulk) for charged and polar groups is large and positive, then moving water molecules from the bulk to hydrate newly exposed polar surface area upon unfolding leads to an increase in density or a decrease in volume. Just the opposite is true for non-polar surface area exposed upon unfolding. The ΔV below ~50°C can be considered to be comprised of the sum of negative values for the exposure of each polar moiety, positive values for the exposure of each non-polar moiety and the contribution of the disappearance of solvent excluded volume upon disruption of the tertiary packing. Given that the protein interior contains most of the non-polar amino-acid side chains, and that disruption of the structure would expose this non-polar surface area, one might expect that the result of the contributions from the exposed polar and non-polar surface area would be a *positive* (or more positive) ΔV. This is not the case. Indeed, it is at these low temperatures that ΔV is found to be at its most *negative*. Hence, we propose that the difference in solvent excluded volume is mainly responsible for the decrease in volume upon unfolding of proteins at low temperature, and that this contribution may indeed overcome a positive contribution from differential solvation. Again, we must bear in mind that the magnitude of the difference in solvent excluded volume depends both on the packing density of

the folded state and the degree of disruption of the unfolded state which is rather poorly characterized in most cases.

The folded state presents a relatively more polar surface area than the unfolded state, and it has a specific three dimensional structure that imposes constraints on its expansion. Hence its expansivity decreases drastically with temperature, whereas that of the unfolded state appears to be rather constant (Fig. 5). Thus, as the temperature increases, the unfolded state expands much more than the folded state. This is why the difference in specific volume between the unfolded and folded state of a protein decreases with increasing temperatures and even becomes positive at sufficiently high temperatures. Thus we would suggest that at low temperature the defining contribution to ΔV comes mainly from excluded volume differences and ΔV for unfolding is negative. In contrast, at high temperatures, the larger thermal volume of the unfolded state ensemble, linked to increased conformational dynamics, takes over and ΔV for unfolding eventually becomes positive.

Changes in solution conditions can also lead to modifications in these volume profiles. High concentrations of denaturant or osmolyte can significantly alter the degree of unfolding, and hence hydration and excluded volume, of the unfolded state. This results in shifting the unfolded state volume profile down (denaturant) or up (osmolyte) with respect to the folded state profile, and hence moving the crossover point to higher or lower temperature. Indeed we have observed increased and positive ΔV values in PPC measurements of SNase and RNase A in the presence of high concentrations of osmolyte. [9,17] Previously reported PPC profiles of SNase as a function of the osmolyte, sorbitol, and the denaturant, urea (Table 1) revealed that the transition shifts, as expected, to higher temperature as a function of increasing osmolyte concentration. The ΔV decreases in absolute value from -19 to -5 mL·mol^{-1} by addition of 1.5. M sorbitol (Table 1). The effect of urea on the PPC profiles is just the opposite. The temperature of the transition decreases, and the value of ΔV decreases from -19 to -56 mL·mol^{-1}.

Certain highly destabilizing or stabilizing mutations can lead to partial unfolding (destabilizing mutations) or compaction (stabilizing mutations) of the folded state. For example, ionization of E66 at pH 9 leads to increased hydration of the folded state, shifting its volume profile down with respect to that of the unfolded state, hence lowering the crossover temperature such that ΔV at the transition becomes positive (Table 1). Had the ionization simply destabilized the folded state, without causing a conformational change that increased its hydration, the ΔV at the transition would have been actually more negative than for the neutral form since the transition temperature is lower. The conformational consequences of protonation of the lysine are less dramatic and ΔV only increases slightly.

To conclude, the results presented here provide a quantitative characterization of the volumetric properties of a typical monomeric protein and demonstrate how subtle conformational changes can modulate these properties. We have come to understand, at least qualitatively, the underlying molecular contributions to the volumetric properties of the various states of proteins and how these change with temperature. Moreover, these results underscore the complexity of volumetric properties of proteins and demonstrate that their interpretation in terms of solvation coupled with protein conformational transitions should be interpreted in the context of their complete temperature dependence.

Experimental Section

DSC and PPC. The differential scanning calorimetric traces were measured by means of a high precision VP DSC micro-calorimeter from MicroCal, Northampton, MA, USA. The reference cell was filled with matching buffer/cosolvent solution. Both buffer and protein solutions were degassed before being injected into the respective cells. The instrument is also equipped with a pressuring cap that allows application of 1.8 bar to the cells in order to avoid air bubbles at elevated temperatures. The instrument was operated in the high gain mode at a rate of 40 ^{o}C·h^{-1}. Baseline subtraction (pure buffer) and normalization with respect to the protein concentration were performed by the instrument software, yielding the temperature-dependent apparent molar heat capacity of the protein, C_p, with respect to the buffer/cosolvent solution. The sample concentration of the proteins in 10 mM phosphate buffer was 5 mg·mL^{-1}.

The pressure perturbation (PPC) experiments were performed in the DSC calorimeter using the MicroCal PPC accessory. [7-9] The reference and sample cell volume are identical (0.5 mL) and they open to a common pressure chamber containing a sensor. The protein concentration used for the PPC studies was 5 mg·mL^{-1} as well. An equal pressure of 5 bar was applied to both cells in a programmed manner using nitrogen gas. The software then initiates a pressure release to ambient pressure. The temperature is kept constant by active compensation of the heat change caused by the pressure jump. As the solutions in the sample and reference cell are identical except for the small amount of dissolved solute in the sample cell counterbalanced by the corresponding volume of buffer in the reference cell, the measured differential heats are quite small. The compensation power returns to the baseline typically within one minute and integration of the supplied power *vs*. time yields the heat consumed or released by the sample. After complete equilibration, an opposite pressure jump is applied when the PPC controller reconnects the PPC cells with nitrogen gas. For both compression and decompression experiments, temperature, pressure, and heat flow are recorded as a function of time. The calorimeter is then automatically heated or cooled to the next desired temperature and the next two (compression / decompression) pressure jumps are applied.

Thermodynamics relates the pressure coefficient of the heat reversibly exchanged, $(\partial Q_{rev}/\partial p)_T$, to the coefficient of thermal expansion, $\alpha = (1/V)(\partial V/\partial T)_p$, of the sample volume. For a sufficiently dilute solution containing m_S grams of solute and m_0 grams of solvent, the volume is given by $V = m_0 V_0 + m_S \overline{V}_S$, where V_0 is the specific volume of the solvent and $\overline{V}_S$ the partial specific volume of the solute. One then finds

$$\left(\partial Q_{rev}/\partial p\right)_T = -TV\alpha = -T\left[m_0 V_0 \alpha_0 + m_S \overline{V}_S \overline{\alpha}_S\right]. \quad (1)$$

$\alpha_0 = (1/V_0)\,(\partial V_0/\partial T)_p$ and $\overline{\alpha}_S = (1/\overline{V}_S)(\partial \overline{V}_S/\partial T)_p$ are the coefficients of thermal expansion associated with the solvent volume and the solute partial volume, respectively.

In the actual differential PPC experiment, the volume occupied by the solute in the sample cell, $m_S \overline{V}_S$, is replaced by the same volume of solvent in the reference cell. Eq. (1) can be integrated to yield the working equation [7]

$$\Delta Q_{rev} = -T\left[m_s \overline{V}_s \overline{\alpha}_s - m_s \overline{V}_s \alpha_0\right]\Delta p \quad (2)$$

which then rearranges to

$$\overline{\alpha}_s = \alpha_0 - \frac{\Delta Q_{rev}}{T m_s \overline{V}_s \Delta p}. \quad (3)$$

α_0 is determined in a separate PPC experiment, with buffer in the sample cell and pure water in the reference cell. In the paper, the partial specific value of the volume, $\overline{V}_s$, and the coefficient of thermal expansion of the solute, $\overline{\alpha}_s$, will be replaced by their apparent values, denoted V and α, respectively.

The relative volume changes $\Delta V/V$ at the unfolding transition of the proteins, taking place in the temperature interval from T_i to T_f, are obtained by:

$$\frac{\Delta V}{V} = \int_{T_i}^{T_f} \alpha(T)\,\mathrm{d}T. \quad (4)$$

Proteins. The hyperstable Δ+PHS variant of nuclease includes mutations P117G, H124L, S128A, G50F, V51N, and a deletion from residues 44–49. Clones of Δ+PHS nuclease were engineered originally by Prof. David Shortle (Johns Hopkins University). The Δ+PHS/V66K and Δ+PHS/V66E proteins were engineered as described in [22, 23]. Protein was expressed and purified following the method of Shortle and Meeker [21]. The protein was determined to be ~98% pure by SDS-PAGE. The concentration was determined using an extinction coefficient of 15,600 M^{-1}cm^{-1} at 280 nm. [24]

Acknowledgements

The authors would like to acknowledge the COST program n° D30/WG0006/03 to CAR and RW, support from INSERM to CAR Financial support to RW from the DFG, EU and the Bundesland NRW is gratefully acknowledged. BGME gratefully acknowledges grant GM-061597 from NIH.

Keywords: pressure perturbation calorimetry - PPC, protein hydration, volumetric properties, SNase

[1] Y. Levy, J. N. Onuchic, *Annu. Rev. Biophys. Biomol. Struct.* **2006***, 35*, 389-415.
[2] P. L. Privalov, S. J. Gill, *Adv. Protein Chem.* **1988**, *39*, 191-234.
[3] T. V. Chalikian, *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 207-235.
[4] R. Ravindra, R. Winter, *ChemPhysChem* **2003**, *4*, 359-365.
[5] C. A. Royer, *Arch. Biochem. Biophys.* **2008**, *469*, 34-45.
[6] A. Cooper, *Biophys. Chem.* **2005**, *115*, 89-97.
[7] L. N. Lin, J. F. Brandts, J. M. Brandts, V. Plotnikov, *Anal. Biochem.* **2002**, *302*, 144-160.
[8] R. Ravindra, C. Royer, R. Winter, *Phys. Chem. Chem. Phys.* **2004**, *6*, 1952-1961; R. Ravindra, R. Winter, *ChemPhysChem* **2003**, *4*, 359-365.
[9] L. Mitra, N. Smolin, R. Ravindra, C. Royer, R. Winter, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1249-1265.
[10] A. Cooper, D. Cameron, J. Jakus, G. W. Pettigrew, *Biochem. Soc. Trans.* **2007**, *35*, 1547-1550.
[11] M. Häckel, H.-J. Hinz, G. R. Hedwig, *Biophys. Chem.* **1999**, *82*, 35-50.
[12] J. D. Bryngelson, P. G. Wolynes, *Proc. Natl. Acad. Sci. U. S. A* **1987**, *84*, 7524-7528.
[13] J. L. Silva, D. Foguel, C. A. Royer, *Trends Biochem. Sci.* **2001**, *26*, 612-618.
[14] B. Garcia-Moreno, J. J. Dwyer, A. G. Gittis, E. E. Lattman, D. S. Spencer, W. E. Stites, *Biophys. Chem.* **1997**, *64*, 211-224.
[15] V. P. Denisov, J. L. Schlessman, E. B. Garcia-Moreno, B. Halle, *Biophys. J.* **2004**, *87*, 3982-3994.
[16] W. E. Stites, A. G. Gittis, E. E. Lattman, D. Shortle, *J. Mol. Biol.* **1991**, *221*, 7-14.
[17] R. Ravindra, C. Royer, R. Winter, *Phys. Chem. Chem. Phys*. **2004**, *6*, 1952-1961.
[18] H. Seemann, R. Winter, C. A. Royer, *J. Mol. Biol.* **2001**, *307*, 1091-1102.
[19] G. Panick, G. J. Vidugiris, R. Malessa, G. Rapp, R. Winter, C. A. Royer, *Biochemistry* **1999**, *38*, 4157-4164.
[20] C. A. Royer, *Biochim. Biophys. Acta* **2002**, *1595*, 201-209.
[21] D. Shortle, A. Meeker, *Proteins Struct. Funct. Genet.* **1986***, 1*, 81–89.
[22] J. J. Dwyer, A. G. Gittis, D. A. Karp, E. E. Lattman, D. S. Spencer, W. E. Stites, B. Garcia-Moreno E., *Biophys. J.* **2000**, *79*, 1610-1620
[23] D. A. Karp, A. G. Gittis, M. R. Strahley, C. A. Fitch, W. E. Stites, B. Garcia-Moreno E., *Biophys. J.* **2007**, *92*, 2041-2053.
[24] C. A. Fitch, D. A. Karp, K. K. Lee, W. E. Stites; E. A. Lattman, B. Garcia-Moreno, *Biophys. J.* **2002**, *82*, 3289-3304.

Received: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))

Unique Features of the Folding Landscape of a Repeat Protein Revealed by Pressure Perturbation

Jean-Baptiste Rouget,[†] Martin A. Schroer,[‡] Christoph Jeworrek,[§] Matthias Pühse,[§] Jean-Louis Saldana,[†] Yannick Bessin,[†] Metin Tolan,[‡] Doug Barrick,[¶] Roland Winter,[§] and Catherine A. Royer[†]*

[†]Centre de Biochimie Structurale, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Université Montpellier, Montpellier, France; [‡]Faculty of Physics/DELTA, and [§]Faculty of Chemistry, Technische Universität Dortmund, Dortmund, Germany; and [¶]Department of Biophysics, The Johns Hopkins University, Baltimore, Maryland

ABSTRACT The volumetric properties of proteins yield information about the changes in packing and hydration between various states along the folding reaction coordinate and are also intimately linked to the energetics and dynamics of these conformations. These volumetric characteristics can be accessed via pressure perturbation methods. In this work, we report high-pressure unfolding studies of the ankyrin domain of the Notch receptor (Nank1–7) using fluorescence, small-angle x-ray scattering, and Fourier transform infrared spectroscopy. Both equilibrium and pressure-jump kinetic fluorescence experiments were consistent with a simple two-state folding/unfolding transition under pressure, with a rather small volume change for unfolding compared to proteins of similar molecular weight. High-pressure fluorescence, Fourier transform infrared spectroscopy, and small-angle x-ray scattering measurements revealed that increasing urea over a very small range leads to a more expanded pressure unfolded state with a significant decrease in helical content. These observations underscore the conformational diversity of the unfolded-state basin. The temperature dependence of pressure-jump fluorescence relaxation measurements demonstrated that at low temperatures, the folding transition state ensemble (TSE) lies close in volume to the folded state, consistent with significant dehydration at the barrier. In contrast, the thermal expansivity of the TSE was found to be equivalent to that of the unfolded state, indicating that the interactions that constrain the folded-state thermal expansivity have not been established at the folding barrier. This behavior reveals a high degree of plasticity of the TSE of Nank1–7.

INTRODUCTION

One of the major issues remaining to be explained in protein folding involves the characterization of the thermodynamics and dynamics linked to changes in solvation that occur concomitant with protein conformational transitions (1). In particular, the volumetric properties of proteins, which can be accessed via pressure perturbation methods, yield information about the changes in relative solvation and packing between various states along the folding reaction coordinate (2–5). These volumetric properties are also intimately linked to the energetics and dynamics of these conformations. As Kauzmann pointed out several years ago (6), ''… volume and enthalpy changes are equally fundamental properties of the unfolding process, and no model can be considered acceptable unless it accounts for the entire thermodynamic behavior.'' We have studied for a number of years the pressure-induced unfolding of a model globular protein, staphylococcal nuclease (Snase) (7–12). Based on pressure perturbation calorimetric measurements and densitometric and high-pressure studies on Snase, we have proposed a qualitative physical model to explain its pressure-temperature phase diagram (2,8,10,11). According to our model, the magnitude of pressure effects depends primarily upon the packing efficiency of the protein core and the relative thermal expansivity and energetics of the various states on the landscape. Our model further assumes that due to the opposite signs of volume changes caused by polar and hydrophobic hydration, solvent density changes (relative to the bulk) linked to hydration of the newly exposed protein surfaces upon unfolding provide a relatively small contribution to ΔV_u. A number of other studies of pressure-induced unfolding have been reported (see Royer and Winter (3–5) for an overview), but given the relative paucity of groups working in this field, our understanding of pressure effects, despite recent progress (13), lags behind our understanding of protein folding via alternate perturbations (temperature, denaturant).

Here, we report a thorough study of the pressure-induced unfolding of the ankyrin repeat domain of the Notch receptor (14) (Nank1–7). This protein provides an interesting model system to study pressure effects on proteins. Its folding has been thoroughly characterized at atmospheric pressure (14–20). Moreover, it is a rather large (213 residues), all α-helical protein exhibiting the elongated, modular hydrophobic core common to repeat proteins. Using fluorescence-based approaches, we have characterized its equilibrium and kinetic pressure responses as a function of both urea concentration and temperature, allowing direct comparison with previous studies of denaturant-induced unfolding at atmospheric pressure. We also have probed, by high-pressure Fourier transform infrared spectroscopy (FTIR) and small-angle x-ray scattering (SAXS), the effect of

Submitted December 18, 2009, and accepted for publication February 26, 2010.

*Correspondence: catherine.royer@cbs.cnrs.fr

Editor: Patricia L. Clark.

© 2010 by the Biophysical Society

0006-3495/10/06/2712/10 $2.00

doi: 10.1016/j.bpj.2010.02.044

denaturants on the helical content and the degree of chain condensation of the pressure-unfolded states of Nank1–7. Our results reveal that the volume change during unfolding for this protein is rather small in magnitude given the rather large size of the protein. We find that the pressure-unfolded state is more compact and retains considerable helical content compared to the urea-induced unfolded state. Finally, we observe a high degree of plasticity of both the unfolded state ensemble and the transition state ensemble (TSE) of Nank1–7 that contrasts with that of the globular protein model Snase, reported previously (10).

MATERIAL AND METHODS

Protein production and purification

Nank1–7 was produced in *Escherichia coli* transformed via a pET15b (Novagen, Madison, WI) plasmid containing the seven-motif ankyrin repeat sequence as previously described (20). The His-tagged ankyrin peptide was purified using nickel-affinity (HiTrap, GE Healthcare, Chalfont St. Giles, United Kingdom) and size-exclusion (HiLoad 26/60 S75, GE Healthcare) chromatography, also as previously described (20). Protein concentration was estimated by absorbance at 280 nm (extinction coefficient, $\varepsilon = 3550$ L mol^{-1} cm^{-1}).

High-pressure fluorescence spectroscopy

High-pressure fluorescence measurements were carried out as previously described (7). Details are provided in the Supporting Material. Briefly, the protein solution at ~50 μM in Tris buffer is loaded into a 500-μL cell fitted with a DuraSeal cap held in place by an O-ring. The cell is placed in a high-pressure vessel equipped with four sapphire windows and connected to a pump and gauge. The pressure-transducing liquid is 18 MOhm MilliQ water. The excitation light is provided from a xenon lamp and monochromator via an optical fiber. Emission is collected at 90° through another monochromator on a photomultiplier tube. For each fluorescence emission spectrum, the center of spectral mass was calculated as

$$\langle\lambda\rangle_j = \sum_j F_j\lambda_j / \sum_j F_j, \tag{1}$$

where F_j is the fluorescence intensity at wavelength $\lambda_j = 320, 321, 322, \ldots, 450$ nm.

For each pressure-induced unfolding measurement, the center of spectral mass was fitted as a function of pressure for values of $\Delta G^{\circ}{}_{u}$ and $\Delta V^{\circ}{}_{u}$ using the BioEQS program, as previously described (Supporting Material (7,21)), taking into account the correction by the relative quantum yield of unfolded and folded states (0.5) obtained from the ratio of the high- and low-pressure spectra. For pressure-induced equilibrium unfolding analysis, the free energy of unfolding is assumed to evolve linearly with the pressure, p:

$$\Delta G^0_u(p) = \Delta G^0_u(p) + p \times \Delta V^0_u, \tag{2}$$

where

$$\Delta G^{\circ}_u(p) = -RT\ln K_f(p) \tag{3}$$

and

$$K_f(p) = \frac{(\langle\lambda_F\rangle - \langle\lambda(p)\rangle)}{(\langle\lambda(p)\rangle - \langle\lambda_U\rangle)}. \tag{4}$$

The kinetic data were fit for the relaxation time (τ) and the activation volumes and rate constants for folding and unfolding at atmospheric pressure were obtained from the values of the natural logarithm of the relaxation time as a function of pressure, as previously described (7). Briefly, the natural logarithm of the relaxation time can be expressed as

$$\ln(\tau)(p) = \ln\left[\frac{1}{k_f(p) + k_u(p)}\right], \tag{5}$$

where $k_{f(p)}$ and $k_{u(p)}$ are the rate constants of folding and unfolding, respectively, at pressure p, and can be expressed in terms of the apparent equilibrium constant in Eq. 3:

$$K_f(p) = k_f(p)/k_u(p). \tag{6}$$

The variation of the rate constants with pressure is expressed in terms of their values at atmospheric pressure, k^0_f and k^0_u:

$$k_f(p) = k^0_f \times e^{-p\times\Delta V^{\#}_f/RT} \tag{7}$$

and

$$k_u(p) = k^0_u \times e^{-p\times\Delta V^{\#}_u/RT}, \tag{8}$$

where $\Delta V^{\#}_f$ and $\Delta V^{\#}_u$ are the activation volumes for folding and unfolding, respectively, which define the global volume change during folding, ΔV_f, by the equation

$$\Delta V_f = \Delta V^{\#}_f - \Delta V^{\#}_u. \tag{9}$$

Plots of ln(τ) versus pressure were fit for k^0_f and $\Delta V^{\#}_f$ by substituting Eqs. 7–9 into Eq. 3.

SAXS

SAXS (22–24) was carried out on Nank1–7 solutions at several pressures and urea concentrations. Protein solutions were prepared immediately before the measurements. To avoid pressure-induced changes of the pH value, 50 mM TrisHCl buffer was used at a pH value of 7.5 (25). Buffers were prepared using deionized water and adding the appropriate amount of buffer salt and urea (2.0 M and 2.2 M, respectively). The solution concentration range was 7.8–11.0 mg/mL. Under these conditions, the solution is diluted enough so that the single scattering approximation is valid (26) yet provides a scattering signal with a good signal/noise ratio.

The SAXS experiments were performed at beamline BL9 of DELTA (TU Dortmund, Dortmund, Germany) (27) and at beamline BW4 of HASYLAB (DESY, Hamburg, Germany) (28). High-pressure SAXS (up to 3 kbar) was carried out in a special sample cell employing two flat diamond windows (29). At BW4 ($\lambda = 1.3808$ Å), using the MAR CCD detector, a q-range of 0.3–3.0 nm^{-1} could be covered. The exposure time was 20 min. A similar exposure time was used at BL9 by employing a MAR345 image plate detector. A wavelength of $\lambda = 1.239$ Å allowed for coverage of a q-range of 0.3–3.5 nm^{-1}. We note that the loss of scattering intensity with pressure is due to the loss of contrast between protein and water, as the density of the water increases with pressure.

The Guinier approximation (30) was used to derive the radius of gyration, R_g, of the protein in solution. The approximation is valid for q-values (q momentum transfer) only up to $q_{max} \leq 1.3/R_g$ (22). The radius of gyration, R_g, was also derived from the pair-distribution function $p(r)$, which characterizes the shape and size of the particle (23) (Supporting Material). After azimuthal averaging of the SAXS spectra, the scattering signals were normalized and corrected for solvent and background scattering by subtracting the scattering curve of the corresponding pure buffer solution. To obtain the radii of gyration, the low q-range showing a linear curve shape in the log ($I(q)$) vs. q^2 representation was fitted according to Guinier's approximation. The pair-distance distribution functions were calculated by fitting the scattering data using the program GNOM (31). To reveal whether the proteins are folded or denatured, SAXS curves are shown in the so-called Kratky plot, i.e., $q^2I(q)$ vs. q. In the case of

a globular, folded protein, the scattering curve exhibits a distinct maximum, whereas in the case of an unfolded polypetide, the curve displays a flat plateau (22).

High-pressure FTIR spectroscopy

A 2% (w/v) solution of Nank1–7 was prepared in a 50 mM Tris, 150 mM NaCl buffer solution in D_2O (pH 8.0) with or without ^{13}C-urea (Sigma, St. Louis, MO) as denaturant. FTIR spectra were recorded in a diamond anvil cell with a Nicolet MAGNA 550 spectrometer. For each pressure step, 256 scans were summed up. Spectral processing and deconvolution of the amide-I′-band was performed with GRAMS/AI 8.0 software. For comparison, all peak areas were normalized to unity.

RESULTS

Pressure-induced equilibrium unfolding

The pressure-induced equilibrium unfolding of Nank1–7 was characterized as a function of both urea concentration and temperature, by measuring the average wavelength of the fluorescence emission of the single tryptophan residue of the protein (Fig. 1), located in the 5th ankyrin repeat (14). The pressure profiles were collected at urea concentrations ranging from 1.9 to 2.4 M, depending upon the temperature, to observe unfolding over our experimentally accessible pressure range (0.1–3500 bar). We note that at the micromolar protein concentrations used for the fluorescence experiments, these transitions were entirely reversible. At all temperatures, the application of pressure results in a sigmoidal red shift in the fluorescence emission (Fig. 2, *A–C*) that exhibited a distinct high-pressure plateau for the average emission wavelength. However, the magnitude of the pressure-induced red shift was strongly dependent upon the concentration of urea. The data were fit at each temperature either individually or globally, linking the volume change across all urea concentrations, for the volume change upon folding (ΔV°_f) and the free-energy change of unfolding (ΔG°_u) (Table S1 in the Supporting Material). In all cases, the volume change was found to be independent of the urea concentration, and the values reported are those from the global fits.

The increase of the plateau value of the average emission wavelength with urea concentration reveals increased solvent relaxation around the excited state of the tryptophan at the high-pressure plateau, indicative of an increased degree of solvent accessibility. At the highest urea concentrations used, at all temperatures, the high-pressure plateau



FIGURE 1 Ribbon diagram of the 3D structure of the ankyrin repeat domain of the Notch receptor, Nank1–7. Each ankyrin repeat is differently colored. The tryptophan residue is shown in blue.

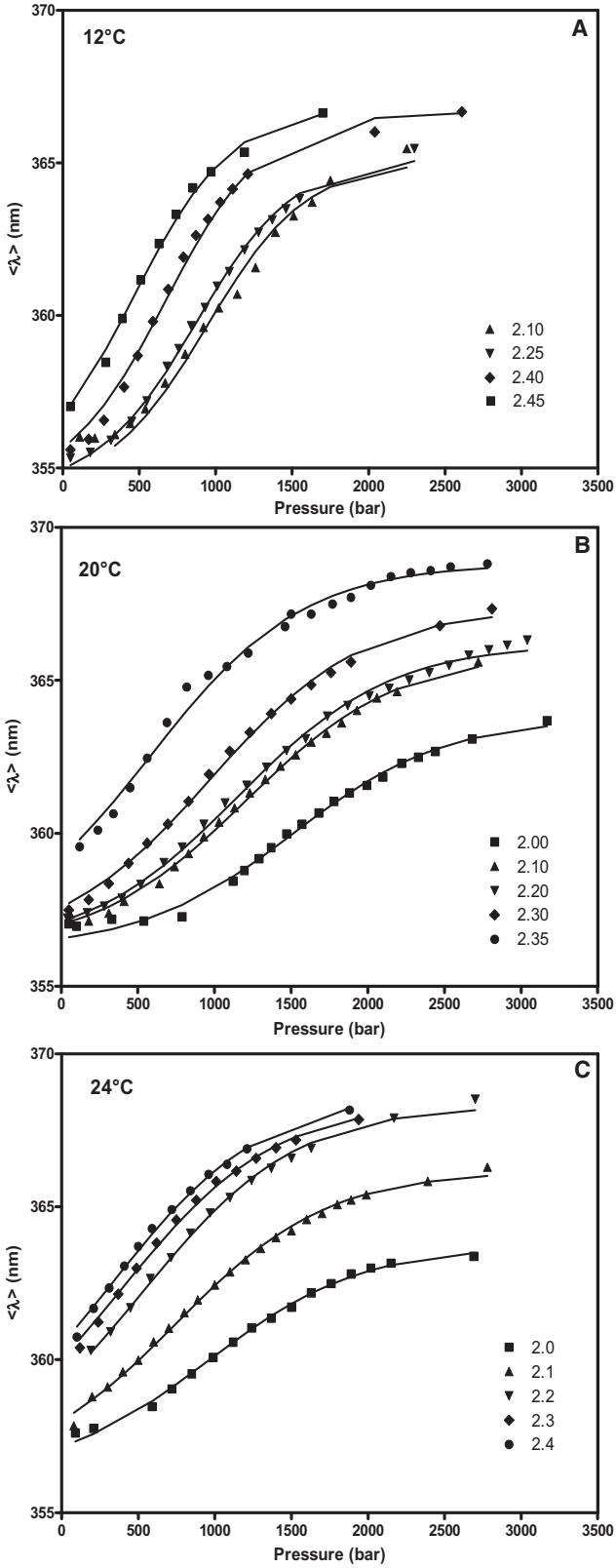

FIGURE 2 Average emission wavelength of the tryptophan fluorescence spectrum of Nank1–7 as a function of pressure and urea concentration (indicated in molar units) at 12°C (*A*), 20°C (*B*), and 24°C (*C*).

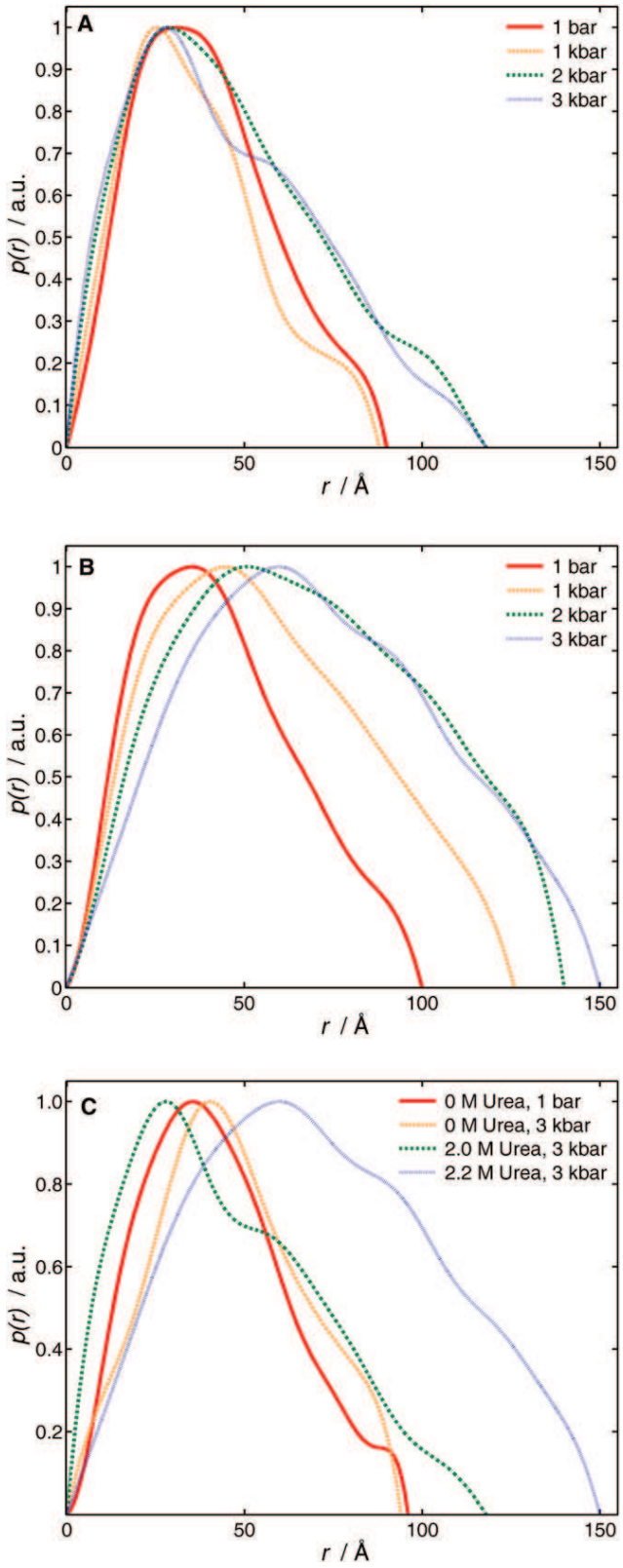

FIGURE 3 (*A* and *B*) Pressure-dependent pair distribution functions obtained from analysis of the SAXS data on Nank1–7 at 24°C at 2 M (*A*) and 2.2 M (*B*) urea. Pressures are as indicated in the figures. (*C*) Comparison of the pair-distribution functions at 1 bar and 0 M urea (*red line*), and at 3 kbar in the presence of 0 M urea (*yellow dotted line*), 2 M urea (*green dotted line*), and 2.2 M urea (*blue line*).

corresponds to the average emission wavelength observed for N-acetyL-tryptophanamide in water, i.e., a completely accessible tryptophan moiety. We hypothesize that the unfolding of Nank1–7 by pressure conforms to a two-state model, with a different degree of tryptophan exposure in the unfolded state depending upon the urea concentration. The values of the ΔG°_u at atmospheric pressure and in the absence of urea are calculated from the linear regression of the plots of $\Delta G^\circ_u(i)$ at atmospheric pressure obtained from the analysis of the pressure unfolding profiles at each urea concentration, i, for a given temperature (Fig. S1 *A*). Although the uncertainty of these extrapolated values is large, given the small number of points (four to five urea concentrations) and the long extrapolation (2.3–0 M urea), the values calculated for ΔG°_u at 16° and 20°C (6.9 ± 1.0 and 7.0 ± 1.25 kcal/mol, respectively) are within the error of those previously reported from urea unfolding studies at atmospheric pressure at 16° and 20°C (8.09 ± 0.07 and 7.52 ± 0.04 kcal/mol, respectively (32)). The nonlinear temperature dependence of ΔG°_u in the absence of urea (Fig. S1 *B*) is consistent with a significant heat capacity change upon unfolding, which is typical of proteins (33). We also find a strong dependence of ΔV°_f upon temperature, ~−1.7 mL/mol K, (Fig. S1 *C*), which arises, as noted above, from the difference in thermal expansivity, $\Delta\alpha_u$, between the unfolded and folded states. The *p-T* phase diagram of Nank1–7 (Fig. S1 *D*) calculated from the values of ΔG°_u and ΔV°_u obtained at 2.1 M urea presents a shape that is similar to that of previously reported *p-T* phase diagrams for proteins (10,34–40).

Chain condensation in the pressure unfolded state

To investigate the structural changes of Nank1–7 associated with changes in pressure and urea concentration, we performed pressure-dependent SAXS experiments at 24°C. We note that, although no aggregation was observed under pressure, these experiments were irreversible upon release of pressure due to aggregation at the high concentrations required for SAXS measurements. In the absence of urea, no appreciable change in the scattering profiles was observed between atmospheric pressure and 3 kbar (Fig. S2 *A* and Fig. 3 *C*). However, in the presence of 2.0 or 2.2 M urea, both the shape and the intensity of the scattering profiles are strongly pressure-dependent (Fig. S2, *B* and *C*), and the pair-distribution functions (Fig. 3, *A* and *B*) reveal a significant increase in their average and largest pairwise distance with increasing pressure and urea concentration. Fig. 3 *C* shows a comparison of the pair-distribution functions for 0, 2, and 2.2 M urea obtained at 3 kbar, corresponding to

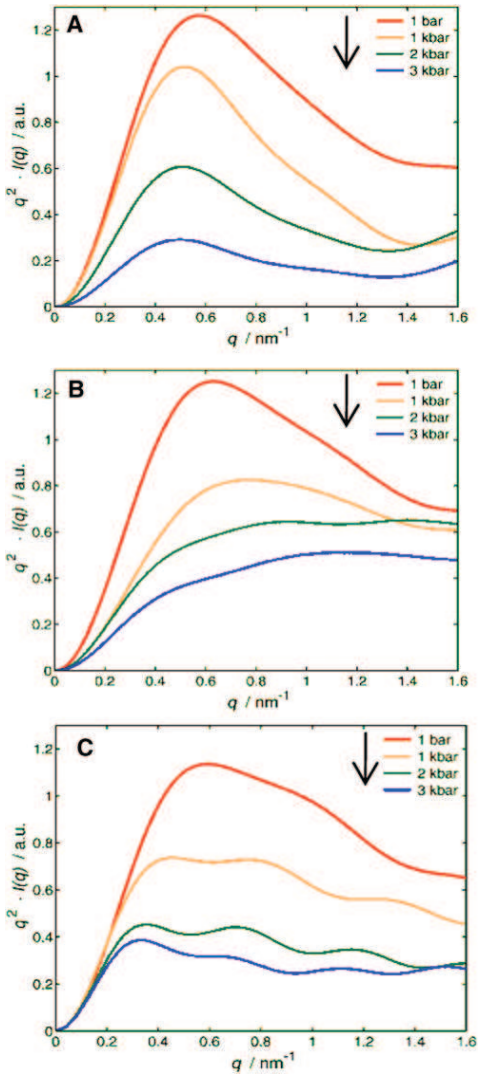

FIGURE 4 Kratky plots obtained from analysis of the SAXS data on Nank1–7 at 24°C for 0, 2, and 2.2 M urea, *A–C* respectively, at the pressures indicated.

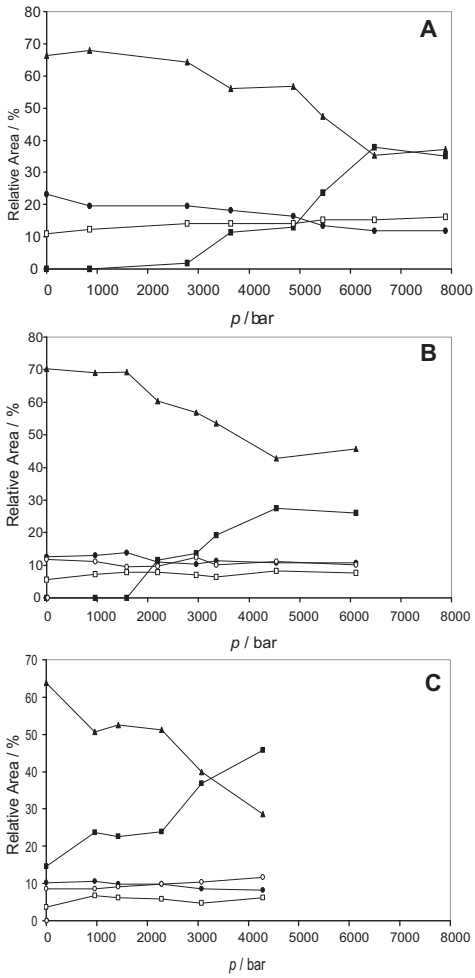

FIGURE 5 Pressure dependence of the secondary structural content for Nank1–7 at 20°C as a function of pressure for 0 (*A*), 2 (*B*), and 2.35 M urea (*C*), where the structural elements are the α-helix (*solid triangles*), the random coil (*solid squares*), the β-sheet (*open squares*), ● and ○ - various turn structures. There is an estimated 5% uncertainty in the calculated secondary structure content.

the high-pressure plateau observed in the fluorescence profiles. The pair-distribution function is much broader at the highest urea concentration, indicating a significantly more expanded chain. Kratky plots at increasing pressure for 0, 2, and 2.2 M urea (Fig. 4, *A–C*) indicate that the chain retains an organized structural character at 3 kbar in the absence of urea but behaves as a Gaussian chain at high pressure for both urea concentrations.

Secondary structure in the pressure-unfolded state

To assess the evolution of secondary structure accompanying the fluorescence- and SAXS-detected unfolding transitions, high-pressure FTIR experiments were carried out in the absence or presence of varying concentrations of urea at 20°C. So that the urea bands would not obscure the protein amide bands, we used ^{13}C-labeled urea as a denaturant (41,42). The application of pressure to Nank1–7, in the absence of urea and in the presence of 2 and 2.35 M ^{13}C-urea, led to a red shift in the amide I absorption band, indicative of a change in secondary structural content (Fig. S3, *A–C*). Analysis of the shift as a function of pressure, as described in the Materials and Methods section, revealed a decrease in helical content (Fig. 4 *A*), although according to the deconvolution software used, residual helical structure remained at the highest pressures, as well as at the 2.35 M urea concentration. Like the SAXS experiments, the FTIR experiments were irreversible upon return to atmospheric pressure due to aggregation at low pressure, but again no aggregation occurred under pressure, because pressure destabilizes protein oligomers and aggregates (43,44). In the absence of urea (Fig. 5 *A*), the pressure-induced loss of ~25% α-helical

content took place over a very broad pressure range (between ~2.5 and 6 kbar), mostly beyond the range of our fluorescence and SAXS high-pressure cells. The lack of any change in secondary structural content below ~3 kbar in the absence of urea is consistent with the insignificant changes in fluorescence and SAXS over the same pressure range. In the presence of 2 M urea (Fig. 5 *B*), a ~25% loss in α-helical content occurred between 1.5 and 4.5 kbar, with about half of the change occurring below 3 kbar, corresponding to the plateau observed in the fluorescence and SAXS experiments. Due to the irreversible nature of the FTIR transition, calculation of an equilibrium volume change is not possible, although we note that the sensitivity of the FTIR transition to pressure is substantially less (by ~50%) than that observed in the reversible fluorescence-monitored transitions. At 2.35 M urea (Fig. 5 *C*), we observed a lower helical content already at atmospheric pressure, and the ~35% pressure-induced loss of α-helix occurred between 0 and 4 kbar. At 2.2 kbar and 2.35 M urea, the fluorescence spectrum of Nank1–7 is as red-shifted as possible (equivalent to N-acetyl L-tryptophanamide in water), and the chain expansion, as observed by SAXS, has also reached a maximum. Therefore, the FTIR results indicate that the Nank1–7 α-helices resist despite considerable disruption of the tertiary structure by pressure. Indeed, it has been shown previously that pressure has very little effect on isolated α-helices (45,46).

Pressure-jump kinetics

Pressure-jump relaxation profiles were obtained as a function of temperature and urea concentration by monitoring the fluorescence intensity at the maximum of the folded-state fluorescence emission. The relaxation time under all these conditions was consistent with a single exponential decay of the profiles of intensity versus time after the pressure jump. Examination of the plots of the natural logarithm of the relaxation times versus pressure (Fig. 6, *A–C*) reveals that, as in the case of the equilibrium unfolding, urea has no effect on the volume change of unfolding. However, it can be seen that the high-pressure slopes of the profiles are strongly dependent on temperature, with a distinct downward curvature (Chevron-like behavior) observed at low temperature and high pressure but fairly linear profiles over the entire unfolding transition at higher temperatures. The data obtained at all temperatures were analyzed for the activation volume and atmospheric rate constant for folding, $\Delta V^{\#}_{f}$ and k°_{f} (Table S2 and Materials and Methods). The corresponding parameters for unfolding ($\Delta V^{\#}_{u}$ and k°_{u}) were derived from the equilibrium free-energy and volume changes. At low temperature, the specific volume of the TSE was found to be significantly closer to that of the folded state than to that of the unfolded state ($\Delta V^{*}_{f} = +50$ mL/mol, $\Delta V^{\#}_{u} = -18$ mL/mol), indicating that at the barrier, a significant degree of dehydration occurs. According to our model, at the barrier, the protein is collapsed, forming a significant

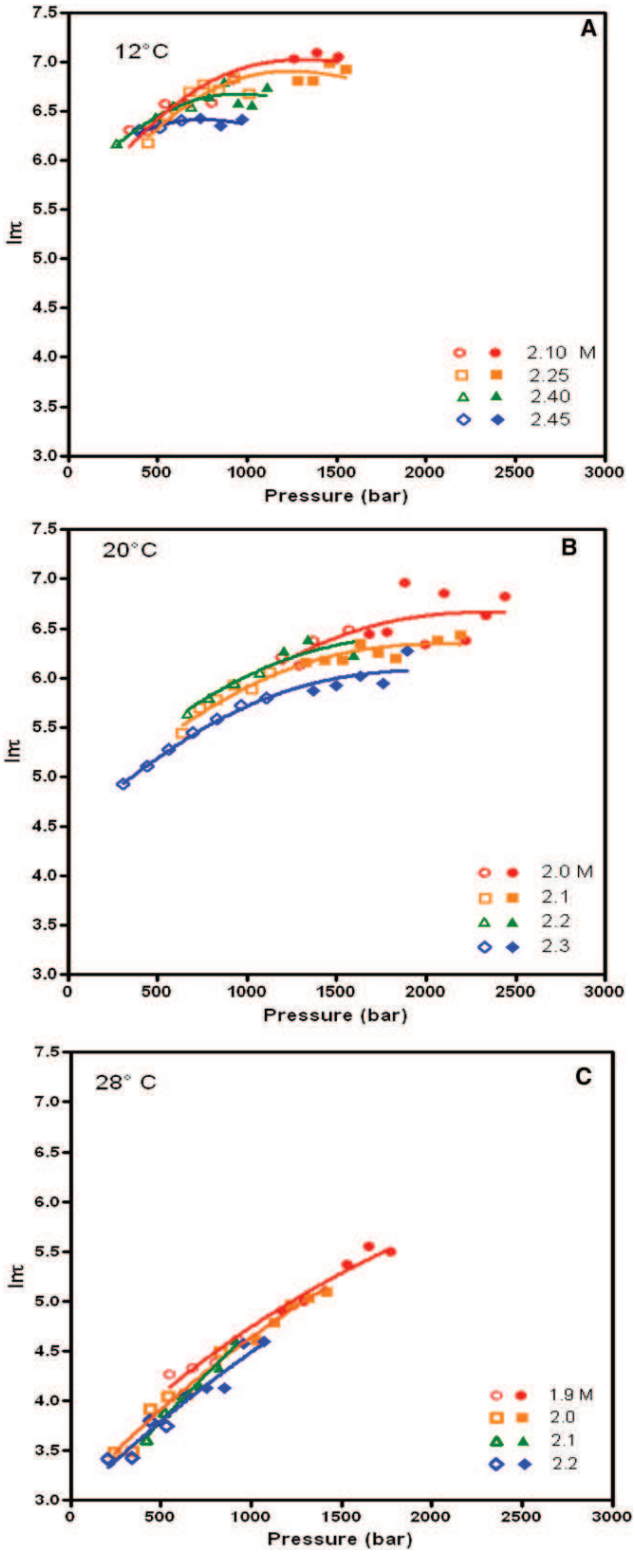

FIGURE 6 Natural logarithm of the pressure-jump fluorescence-detected relaxation time for Nank1–7 as a function of pressure for the urea concentrations indicated at 12°C (*A*), 20°C (*B*), and 28°C (*C*). Lines through the points correspond to fits of the data, as described in Materials and Methods. Open symbols represent data points below the transition midpoint and solid symbols those above the transition midpoint.

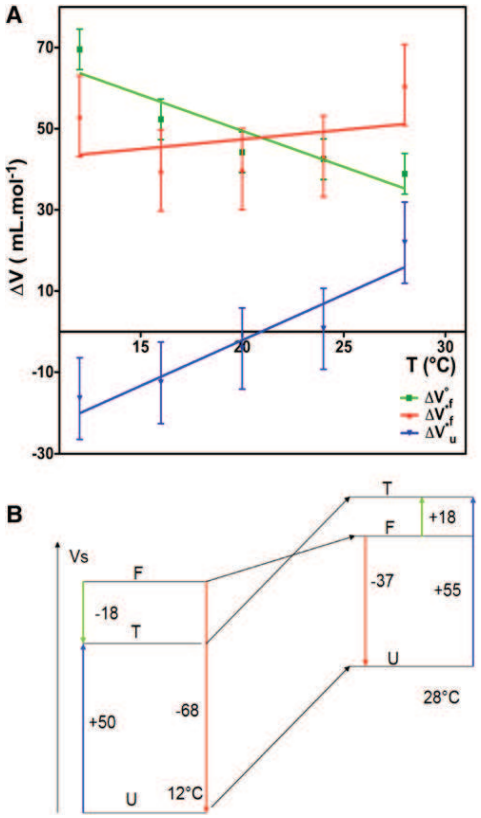

FIGURE 7 Volumetric properties of Nank1–7. (*A*) Temperature dependence of the equilibrium volume change for unfolding, ΔV°_{u} (■, *green*), and the activation volumes for folding, $\Delta V^{\#}_{f}$ (Δ, *red*) and unfolding, $\Delta V^{\#}_{u}$ (▽, *blue*). Points correspond to the values obtained from fits of the data, and lines represent the linear regressions. (*B*) Volumetric diagram (V_M is molar volume) for Nank1–7 at 12° and 28°C, showing the relationships between the specific molar volumes as well as the thermal expansivity of the folded state, F, the unfolded state, U, and the TSE.

proportion of the native-state solvent-excluded volume, and is thus relatively ''dry'' (Fig. 7 *B*).

Comparison of the temperature dependence of the activation volumes for folding and unfolding with that of the equilibrium volume change (Fig. 7 *A*) reveals that the entire temperature dependence of the ΔV°_{f} resides in the activation volume for unfolding, $\Delta V^{\#}_{u}$. Moreover, the value of ΔV^{*}_{u} is negative at low temperature and changes sign to become positive at high temperature. Indeed, unlike denaturant-based unfolding, pressure-induced unfolding need not display Chevron-like behavior, as the sign of the activation volumes can be either positive or negative. The temperature dependence of the activation volumes, like that of the equilibrium volume change, is equivalent to the difference in thermal expansivity between the two states implicated in the reaction. We observe no dependence of the $\Delta V^{\#}_{f}$ on temperature (0.12 ± 0.54 mL/mol K). In contrast, we observe a strong dependence of $\Delta V^{\#}_{u}$ on temperature (2.08 ± 0.46 mL/mol K), highlighting a significant difference in thermal expansivity between the TSE and the folded state. This value is equivalent, within error, to that (1.72 ± 0.25 mL/mol K) observed for the ΔV°_{u} (the difference in thermal expansivity between the unfolded and folded states). Hence, we deduce that the thermal expansivity of the TSE is equivalent to that of the unfolded state (Fig. 7 *B*).

DISCUSSION

The folding energetics and kinetics of Nank1–7 have been characterized and linked to structural folding pathways (17,18,47), and prior results provide a solid basis for comparison and interpretation of the high-pressure studies presented here. These SAXS and fluorescence studies reveal pressure-induced unfolding with a plateau near 3 kbar. Nank1–7, with a value of ~45 ml/mol, exhibits a rather small pressure sensitivity (3) given its size (213 residues).

Comparison of the fluorescence, SAXS, and FTIR data reveals significant effects of urea on the physical properties of pressure-unfolded Nank1–7. In Fig. 8, the structural properties of the pressure-unfolded state of Nank1–7 are presented under different conditions of temperature and urea concentration. At moderate urea concentration, the pressure-unfolded ensemble, although a Gaussian chain, is not completely expanded and retains a major fraction of the native secondary structure. Indeed, the chain becomes much more expanded and the tryptophan significantly more exposed to solvent, and the secondary structure content decreases to ~50% that of the native protein at the high-pressure plateau reached in the presence of slightly higher urea concentrations. It is not surprising that the secondary structure persists at high pressure. It has been shown that pressure has little effect on α-helices (45,46,48), and hence, the loss of secondary structure under pressure typically occurs only if the helices are not intrinsically stable in the absence of tertiary interactions. Estimation of helix stability for the Nank helices using the program Agadir (49), revealed that

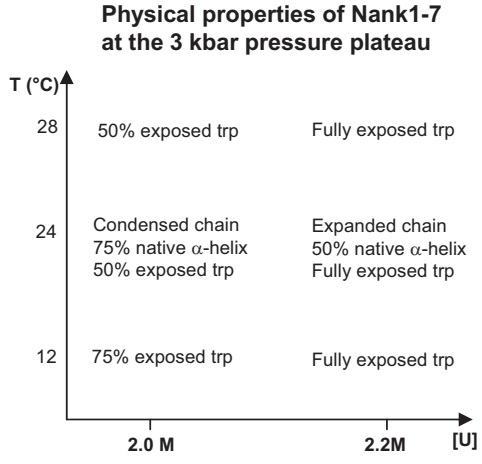

FIGURE 8 Schematic of the temperature and urea concentration dependence of the physical properties of Nank1–7 at the 3-kbar pressure plateau.

in comparison to the ~60–65% helicity of the peptide described by Baldwin and co-workers (AAAAKAAAA-KAAAAKA*)* (50), the Nank1–7 helices were highly unstable. However, the residual interactions responsible for the relative chain compaction and blue tryptophan emission at low urea likely stabilize the helices. Results of recent experimental and theoretical studies on volume changes suggest that pressure-induced unfolding results from internal voids in the native state. Since these voids are not always homogeneously distributed throughout the structure, certain regions can be less pressure-sensitive than others.

It is interesting that fits of the fluorescence equilibrium profiles to a two-state unfolding model yielded equivalent volume changes for unfolding, despite the larger plateau values for the average emission wavelength of the tryptophan with increasing urea. We and others have shown that denaturant itself does not influence the value of the volume change upon unfolding (51,52). However, we observe here that increasing urea leads to greater solvent exposure of the tryptophan residue and a more expanded Gaussian chain at high pressure, without any change in molar volume. Such an observation is consistent with our working model, in which disappearance of internal voids upon unfolding, rather than density changes due to differential (polar versus hydrophobic) solvation, provide the bulk of the volume change upon unfolding. Once the protein has expanded sufficiently to allow occupation of these internal voids by water, pressure has no further effect. Hence, we suggest that the pressure-unfolded state, which is not maximally expanded or devoid of α-helices, may more closely resemble the unfolded state that is sampled by this protein under native conditions than those unfolded by denaturant or temperature perturbation.

Consistent with the lack of energetic barriers between these unfolded conformations, extrapolation to atmospheric pressure and 0 M urea of the urea- and pressure-dependent fluorescence free energies of unfolding yields values that are in reasonably good agreement with the global stability of the protein previously determined from urea denaturation profiles under similar conditions (32). For this transition from the folded state to the compact denatured state, the *p-T* phase behavior of Nank1–7 displays the shape observed for the few globular proteins that have been studied to date (10,34–40). Indeed, it has already been shown that Nank1–7 exhibits a large heat capacity upon unfolding (19), and we show here that it also exhibits a significant change in thermal expansivity, the second most important parameter giving rise to the curvature in the *p-T* plane.

Our pressure-jump kinetics measurements on Nank1–7 fit well to a two-state unfolding/refolding model. In contrast, thorough studies of Nank1–7 folding and unfolding kinetics using denaturant revealed multiple kinetic phases, resulting from denatured-state proline isomerization and an on-pathway unfolding intermediate (18). This intermediate displayed tryptophan emission and circular dichroism similar to the folded state, a modest free-energy gain, and ~40% loss of solvent-accessible surface area compared to the unfolded state. Proline isomerization reactions, albeit quite slow, are not rate-limiting at high pressure (53), due to the large positive activation volume for folding, which slows the main folding transition even further (12,53). Hence, it is not surprising that the proline-based kinetic complexity was not observed in our pressure-jump studies. Likewise, the on-pathway intermediate was not detected in our pressure-jump studies, either because it was not highly populated enough to be observable in our experiments or because of the pressure-dependent increase in the main folding barrier.

The TSE for the single transition that we observe exhibits a molar volume at low temperature quite close to that of the folded state. According to our working model of pressure effects, this indicates that the bulk of the solvent-excluded packing defects have formed in this TSE. In contrast, the thermal expansivity of the TSE is equivalent to that of the unfolded state, indicating that most of the native contacts that limit the thermal expansion of the folded state are not formed at the barrier. Previous studies on Nank1–7 folding indicated that it folds inside-out, with the main barrier corresponding to the formation of the central repeats (3–5,54). Our expansivity data suggest that this TSE remains quite malleable. An expansivity for the TSE of Nank1–7 equivalent to that of the unfolded state is in contrast with that found for the TSE of Snase (the only other protein for which this quantity has been measured) (10), which was nearly equivalent to that of the folded state, consistent with a substantial degree of structure at the barrier for this globular protein. That the tertiary interactions in Nank1–7 are not locked in at the folding barrier may derive from the elongated and modular nature of the hydrophobic core of Nank1–7, which imposes fewer spatial constraints on the structure.

In conclusion, these pressure perturbation studies on the ankyrin repeat domain of the Notch receptor reveal a high degree of plasticity in both the unfolded-state ensemble and the TSE. The physical properties of these states, such as chain condensation and thermal expansion, are strongly dependent upon conditions. Overall, the observation of a less drastic effect of pressure, as compared to chemical denaturants, can be understood in light of the different physical basis for their perturbations. Urea interacts with the protein, leading to more expanded chains in the unfolded state via an increase in the persistence length. Pressure effects arise simply due to the smaller system volume when the protein is in the unfolded state, and therefore do not necessarily result in maximal expansion of the chain, or maximal loss of helical content. Our understanding of protein folding mechanisms is dependent upon our ability to characterize the conformational states on the protein folding landscape. Here, we show that pressure can reveal a unique set of conformational states with distinct structural and thermodynamic properties, thus increasing our knowledge and broadening our perspective of this complex process.

SUPPORTING MATERIAL

Two tables and three figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00324-3.

The BL9 and BW4 beamline crews and the machine groups of DELTA and HASYLAB are acknowledged. For their excellent help during the SAXS beamtimes, we thank M. Paulus, C. Sternemann, C. Krywka, D. C. F. Wieland, A. Steffen, and C. J. Sahle. C.A.R. also thanks Reinhard Lange for the loan of his optical high-pressure cell for part of this work.

This work was supported by funds from Institut National de la Santé et de la Recherche Médicale and the Agence National pour la Recherche (Prirbio-09 455024) to C.A.R. and grants from the Deutsche Forschungsgemeinschaft (to R.W.) and the National Institutes of Health (GM060842, to D.B.). J.B. is a recipient of a graduate fellowship from the French Ministry of Research and Higher Education. M.A.S. acknowledges the Deutsche Forschungsgemeinschaft (DFG TO169/14-1) for financial support.

REFERENCES

1. Levy, Y., and J. N. Onuchic. 2006. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* 35:389–415.
2. Mitra, L., J. B. Rouget, …, R. Winter. 2008. Towards a quantitative understanding of protein hydration and volumetric properties. *ChemPhysChem.* 9:2715–2721.
3. Royer, C. A. 2002. Revisiting volume changes in pressure-induced protein unfolding. *Biochim. Biophys. Acta.* 1595:201–209.
4. Royer, C. A., and R. Winter. 2008. Volumetric properties of proteins and the role of solvent in conformational dynamics. *In* Water and Biomolecules: Physical Chemistry of Life Phenomena. K. Kuwajima, editor. Springer, New York. 173–186.
5. Royer, C. A. 2008. The nature of the transition state ensemble and the mechanisms of protein folding: a review. *Arch. Biochem. Biophys.* 469:34–45.
6. Kauzmann, W. 1987. Thermodynamics of unfolding. *Nature.* 325:763–764.
7. Brun, L., D. G. Isom, …, C. A. Royer. 2006. Hydration of the folding transition state ensemble of a protein. *Biochemistry.* 45:3473–3480.
8. Mitra, L., N. Smolin, …, R. Winter. 2006. Pressure perturbation calorimetric studies of the solvation properties and the thermal unfolding of proteins in solution—experiments and theoretical interpretation. *Phys. Chem. Chem. Phys.* 8:1249–1265.
9. Mitra, L., K. Hata, …, C. A. Royer. 2007. V(i)-value analysis: a pressure-based method for mapping the folding transition state ensemble of proteins. *J. Am. Chem. Soc.* 129:14108–14109.
10. Panick, G., G. J. Vidugiris, …, C. A. Royer. 1999. Exploring the temperature-pressure phase diagram of staphylococcal nuclease. *Biochemistry.* 38:4157–4164.
11. Seemann, H., R. Winter, and C. A. Royer. 2001. Volume, expansivity and isothermal compressibility changes associated with temperature and pressure unfolding of Staphylococcal nuclease. *J. Mol. Biol.* 307:1091–1102.
12. Vidugiris, G. J., J. L. Markley, and C. A. Royer. 1995. Evidence for a molten globule-like transition state in protein folding from determination of activation volumes. *Biochemistry.* 34:4909–4912.
13. Schweiker, K. L., V. W. Fitz, and G. I. Makhatadze. 2009. Universal convergence of the specific volume changes of globular proteins upon unfolding. *Biochemistry.* 48:10846–10851.
14. Zweifel, M. E., D. J. Leahy, …, D. Barrick. 2003. Structure and stability of the ankyrin domain of the *Drosophila* Notch receptor. *Protein Sci.* 12:2622–2632.
15. Barrick, D., D. U. Ferreiro, and E. A. Komives. 2008. Folding landscapes of ankyrin repeat proteins: experiments meet theory. *Curr. Opin. Struct. Biol.* 18:27–34.
16. Bradley, C. M., and D. Barrick. 2006. The notch ankyrin domain folds via a discrete, centralized pathway. *Structure.* 14:1303–1312.
17. Mello, C. C., and D. Barrick. 2004. An experimentally determined protein folding energy landscape. *Proc. Natl. Acad. Sci. USA.* 101: 14102–14107.
18. Mello, C. C., C. M. Bradley, …, D. Barrick. 2005. Experimental characterization of the folding kinetics of the notch ankyrin domain. *J. Mol. Biol.* 352:266–281.
19. Zweifel, M. E., and D. Barrick. 2001. Studies of the ankyrin repeats of the *Drosophila melanogaster* Notch receptor. 2. Solution stability and cooperativity of unfolding. *Biochemistry.* 40:14357–14367.
20. Zweifel, M. E., and D. Barrick. 2001. Studies of the ankyrin repeats of the *Drosophila melanogaster* Notch receptor. 1. Solution conformational and hydrodynamic properties. *Biochemistry.* 40:14344–14356.
21. Royer, C. A., W. R. Smith, and J. M. Beechem. 1990. Analysis of binding in macromolecular complexes: a generalized numerical approach. *Anal. Biochem.* 191:287–294.
22. Kataoka, M., I. Nishii, …, Y. Goto. 1995. Structural characterization of the molten globule and native states of apomyoglobin by solution x-ray scattering. *J. Mol. Biol.* 249:215–228.
23. Lindner, P., and T. Zemb. 2002. Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter. Elsevier Science, Amsterdam.
24. Svergun, D. I., and M. H. J. Koch. 2003. Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* 66: 1735–1782.
25. Neumann, Jr., R. C., W. Kauzmann, and A. Zipp. 1973. Pressure dependence of weak acid ionization in aqueous buffers. *J. Phys. Chem.* 77:2687–2693.
26. Javid, N., K. Vogtt, …, R. Winter. 2007. Protein-protein interactions in complex cosolvent solutions. *ChemPhysChem.* 8:679–689.
27. Krywka, C., C. Sternemann, …, M. Tolan. 2007. The small-angle and wide-angle x-ray scattering set-up at beamline BL9 of DELTA. *J. Synchrotron Radiat.* 14:244–251.
28. Roth, S. V., R. Döhrmann, …, P. Muller-Buschbaum. 2006. Small-angle options of the upgraded ultrasmall-angle x-ray scattering beamline BW4 at HASYLAB. *Rev. Sci. Instrum.* 77:085106.
29. Krywka, C., C. Sternemann, …, R. Winter. 2008. Effect of osmolytes on pressure-induced unfolding of proteins: a high-pressure SAXS study. *ChemPhysChem.* 9:2809–2815.
30. Guinier, A., and G. Fournet. 1955. Small Angle Scattering of X-rays. Wiley, New York.
31. Svergun, D. I. 1992. Determination of regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.
32. Street, T. O., C. M. Bradley, and D. Barrick. 2007. Predicting coupling limits from an experimentally determined energy landscape. *Proc. Natl. Acad. Sci. USA.* 104:4907–4912.
33. Zweifel, M. E., and D. Barrick. 2002. Relationships between the temperature dependence of solvent denaturation and the denaturant dependence of protein stability curves. *Biophys. Chem.* 101–102:221–237.
34. Hawley, S. A. 1971. Reversible pressure-temperature denaturation of chymotrypsinogen. *Biochemistry.* 10:2436–2442.
35. Herberhold, H., and R. Winter. 2002. Temperature- and pressure-induced unfolding and refolding of ubiquitin: a static and kinetic Fourier transform infrared spectroscopy study. *Biochemistry.* 41:2396–2401.
36. Lassalle, M. W., H. Yamada, and K. Akasaka. 2000. The pressure-temperature free energy-landscape of staphylococcal nuclease monitored by (1)H NMR. *J. Mol. Biol.* 298:293–302.
37. Meersman, F., L. Smeller, and K. Heremans. 2006. Protein stability and dynamics in the pressure-temperature plane. *Biochim. Biophys. Acta.* 1764:346–354.
38. Smeller, L., F. Meersman, and K. Heremans. 2006. Refolding studies using pressure: the folding landscape of lysozyme in the pressure-temperature plane. *Biochim. Biophys. Acta.* 1764:497–505.

39. Wiedersich, J., S. Köhler, …, J. Friedrich. 2008. Temperature and pressure dependence of protein stability: the engineered fluorescein-binding lipocalin FluA shows an elliptic phase diagram. *Proc. Natl. Acad. Sci. USA.* 105:5756–5761.

40. Zipp, A., and W. Kauzmann. 1973. Pressure denaturation of metmyoglobin. *Biochemistry.* 12:4217–4228.

41. From, N. B., and B. E. Bowler. 1998. Urea denaturation of staphylococcal nuclease monitored by Fourier transform infrared spectroscopy. *Biochemistry.* 37:1623–1631.

42. Reinstädler, D., H. Fabian, …, D. Naumann. 1996. Refolding of thermally and urea-denatured ribonuclease A monitored by time-resolved FTIR spectroscopy. *Biochemistry.* 35:15822–15830.

43. Silva, J. L., and G. Weber. 1993. Pressure stability of proteins. *Annu. Rev. Phys. Chem.* 44:89–113.

44. Weber, G., and H. G. Drickamer. 1983. The effect of high pressure upon proteins and other biomolecules. *Q. Rev. Biophys.* 16:89–112.

45. Imamura, H., and M. Kato. 2009. Effect of pressure on helix-coil transition of an alanine-based peptide: an FTIR study. *Proteins.* 75:911–918.

46. Paschek, D., S. Gnanakaran, and A. E. Garcia. 2005. Simulations of the pressure and temperature unfolding of an α-helical peptide. *Proc. Natl. Acad. Sci. USA.* 102:6765–6770.

47. Kloss, E., N. Courtemanche, and D. Barrick. 2008. Repeat-protein folding: new insights into origins of cooperativity, stability, and topology. *Arch. Biochem. Biophys.* 469:83–99.

48. Takekiyo, T., N. Takeda, …, Y. Taniguchi. 2007. Pressure stability of the α-helix structure in a de novo designed protein (α-l-α)(2) studied by FTIR spectroscopy. *Biopolymers.* 85:185–188.

49. Muñoz, V., and L. Serrano. 1997. Development of the multiple sequence approximation within the AGADIR model of α-helix formation: comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers.* 41:495–509.

50. Marqusee, S., V. H. Robbins, and R. L. Baldwin. 1989. Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. USA.* 86:5286–5290.

51. Kitahara, R., C. Royer, …, C. Roumestand. 2002. Equilibrium and pressure-jump relaxation studies of the conformational transitions of P13MTCP1. *J. Mol. Biol.* 320:609–628.

52. Pappenberger, G., C. Saudan, …, T. Kiefhaber. 2000. Denaturant-induced movement of the transition state of protein folding revealed by high-pressure stopped-flow measurements. *Proc. Natl. Acad. Sci. USA.* 97:17–22.

53. Vidugiris, G. J., D. M. Truckses, …, C. A. Royer. 1996. High-pressure denaturation of staphylococcal nuclease proline-to-glycine substitution mutants. *Biochemistry.* 35:3857–3864.

54. Tripp, K. W., and D. Barrick. 2008. Rerouting the folding pathway of the Notch ankyrin domain by reshaping the energy landscape. *J. Am. Chem. Soc.* 130:5681–5688.

Résumé

Les protéines sont des macromolécules complexes consistant en une chaine linéaire d'acides aminés. Cette chaine adopte une structure tridimensionnelle requise généralement pour son activité par un processus complexe de repliement. Le repliement protéique est impliqué dans de nombreux processus biologiques et maladies, telles les maladies à prions et les fibres amyloïdes qui leur sont associées. Ainsi, un examen approfondi des mécanismes de repliement/dépliement est nécessaire à une meilleure compréhension de ces maladies, ainsi qu'à la découverte de nouveaux moyens d'inhibition. Un second point d'intérêt de l'étude du repliement est l'existence de protéines nativement dépliées. En effet, de nombreuses protéines humaines et virales sont nativement dépliées et ne se replient qu'au moment de leur activité, rendant la caractérisation de leurs structures très difficiles. Malheureusement, ces protéines peuvent être en principe d'excellentes cibles pour le design de molécules thérapeutiques. Une connaissance complète du processus de repliement pourrait donc amener à l'émergence de meilleurs modèles ou structures in-silico, qui pourraient éventuellement être à l'origine d'études dans le but de les stabiliser ou de créer des modèles de pharmacophores d'intérêt biologique. Enfin, une compréhension fondamentale du repliement protéique est un réel challenge méritant toute notre attention. L'histoire de la science a prouvé à de nombreuses reprises que nous ne pouvons jamais savoir à l'avance ce que de nouvelles découvertes fondamentales peuvent apporter en termes d'opportunités, de biotechnologies ou de nouvelles voies de recherches, mais il est évident que plus grand est le challenge, plus nombreux sont ses débouchés potentiels.

Afin de mieux comprendre le processus de repliement protéique, nous cherchons dans ce travail à améliorer notre compréhension de la thermodynamique du repliement, celle-ci étant déjà avancée mais incomplète. En effet, grâce à la dénaturation thermique et chimique via l'utilisation de dénaturants tels l'urée ou la guanidine hydro chlorée, il a été beaucoup appris sur comment contribuent à la stabilité protéique les interactions intramoléculaires comme les interactions charge-charge, les ponts hydrogènes, les interactions de Van Der Waals ou les effets hydrophobes. Toutefois, notre compréhension du repliement protéique reste incomplète, en particulier car les effets déstabilisants de la pression hydrostatique, l'un des paramètres fondamentaux de la thermodynamique, ne sont pas compris.

Les protéines sont naturellement marginalement stables, avec une différence d'énergie

libre de Gibbs entre les états repliés et dépliés de la protéine (ΔG) de l'ordre de 20 à 40 KJ/mol. Cette faible stabilité est en général imputée au fait que les protéines ont besoin d'une grande flexibilité conformationnelle pour permettre leur fonction, résultant de la compensation entre entropie et enthalpie, mais aussi du fait des fluctuations statistiques des variables thermodynamiques qui peuvent être importantes selon leur hydratation. D'après le modèle à deux états, une barrière énergétique existe entre ces deux populations énergétiquement différentes. D'un point de vue cinétique, les protéines subissent des mouvements thermiques, induisant des changements énergétiques qui impliquent des changements structuraux.

Nous savons que la Terre est un environnement où les hautes pressions sont prédominantes, et 62% de la biosphère totale de la planète est caractérisée par des pressions supérieures à 1 kbar. Grâce au travail de Hawley [51], Brandts [52] et Kauzmann [53], nous avons depuis le début des années 70 une bonne idée du comportement de protéines à deux états dans le plan température-pression. Ce comportement est représenté par une ellipse où ΔG est nul en fonction de la température et de la pression. A l'intérieur de l'ellipse, les protéines sont repliées, et à l'extérieur elles sont dépliées.

Cette ellipse est définie par l'équation suivante (1): $\Delta G(p,T) = \Delta G_0 - \Delta S_0(T - T_0) - \Delta Cp[T(ln(T/T_0) - 1) + T_0] + \Delta V_0(p - p_0) + (\Delta^{1/2}\beta)(p - p_0)2 + \Delta\alpha(T - T_0)(p - p_0)$

L'équation 1 est une série de Taylor du second ordre de l'équation suivante (2): $\delta\Delta G = \Delta V\delta p - \Delta S\delta T$, autour d'un point de référence choisi (p_0, T_0) [55,56]. En analysant ce diagramme de stabilité sur le plan température-pression, nous voyons que ΔG varie avec la température à pression constante. En effet, la température est connue pour affecter la stabilité protéique, ce qui est reflété par des valeurs positives ou négatives du ΔG, selon que la température stabilise ou déstabilise la protéine. La variation du ΔG n'est pas linéaire et provient de la compensation entre l'enthalpie et l'entropie selon la température. En effet, le ΔH, qui dépend du ΔCp, lui-même étant corrélé à la différence de surface accessible au solvant entre les deux états, augmente avec la température car l'hydrophobicité est différente entre deux états de la protéine. L'évolution du ΔH avec la température est linéaire, et sa pente correspond au ΔCp. Le ΔS, la différence d'entropie entre les deux états de la protéine, augmente aussi avec la température, mais d'une manière non linéaire. En effet, si nous considérons l'entropie de l'état replié, en augmentant la température, celle-ci augmente car le désordre augmente, mais moins que l'entropie de l'état déplié (entropie de la chaine) qui augmente beaucoup plus, sa structure n'étant pas restreinte. Ainsi, selon la température et le degré de désordre de l'état replié, ΔS peut être négatif ou positif. La compensation d'entropie et d'enthalpie est donc bien responsable des valeurs de ΔG avec la température. Les effets de la température sont donc bien définis et compris sur les protéines.

Nous constatons de même que le ΔG varie aussi à température constante en fonc-

tion de la pression appliquée. Ce dernier aspect provient de l'existence d'un ΔV, car ce seul paramètre dépend uniquement de la pression dans ce cas. En effet, à température constante, l'équation (1) perd ΔS , ΔCp et $\Delta\alpha$, comme $T = T_0$. Les seuls paramètres restants sont ΔV et $\Delta\beta$. Or, $\Delta\beta$ est connu pour être peu important car les deux états protéiques sont comprimés de la même manière. Donc le seul paramètre expliquant l'effet de la pression est ΔV. Nous constatons de plus que la droite des ΔV=0 n'est pas parallèle à l'axe X, et coupe celui-ci, indiquant la capacité du ΔV à changer de signe, alors que celui-ci est connu pour être généralement négatif. Ce possible changement de signe du ΔV a été nommé « paradoxe du ΔV » par Chalikian et Breslauer, et n'est encore pas bien compris. Il semblerait que l'amplitude de cet effet provienne de la différence d'expansivité thermique entre les états repliés et dépliés de la protéine ($\Delta\alpha$), mais nous ne comprenons pas encore quels sont les paramètres définissant cette différence d'expansivité.

Comme nous avons pu le constater, les effets de la température sur la stabilité protéique sont bien compris, mais pas les effets de la pression, du fait du manque de connaissances que nous avons concernant le définition des ΔV, et le peu de connaissances sur le $\Delta\alpha$ et sur l'expansivité de l'état replié. Pour mieux comprendre les effets de la pression, il nous faut donc mieux comprendre le ΔV, et en particulier d'où proviennent les volumes des deux états. Le volume molaire protéique est constitué du volume de la protéine et de celui de l'eau. Le volume de la protéine correspond à celui des atomes ($V_{prot,atomes}$) et des vides entre les atomes dans la structure ($V_{prot,vide}$) provenant des cavités et mésappariement.

Le volume de l'eau autour de la protéine est composé de l'eau du « bulk » ($V_{eau,bulk}$) et de l'eau d'hydratation ($V_{eau,hyd}$). Donc nous pouvons définir le volume molaire de la protéine comme étant : $Vpr = V_{prot,atomes} + V_{prot,vides} + V_{eau,bulk} + V_{eau,hyd}$. Comme il existe un ΔV entre les deux états, une partie de ces volumes doit en être la cause. Le volume des atomes ne peut changer, donc $\Delta V = \Delta V_{vides} + \Delta V_{hydratation} = \Delta V_{vides} + \Delta V_{polaires} + \Delta V_{apolaires}$, où ΔV_{vides} est défini comme la différence de volume entre les deux états, provenant du remplissage des vides par l'eau et le $\Delta V_{hydratation}$ par le changement du volume du solvant induit par l'exposition de la surface accessible au solvant, terme qui contient des contributions d'hydratation hydrophobe et polaire. Nous notons que ce dernier terme doit donc être proportionnel au ΔASA, comme le ΔCp. Mais nous savons que le ΔV, contrairement au ΔCp ne montre pas de corrélation avec le ΔASA. Nous devons donc apprendre à paramétrer le ΔV pour mieux comprendre son implication dans les effets de la pression.

Comme l'a clairement expliqué Kauzmann : « Une compréhension fondamentale du comportement et des propriétés de la matière requiert une caractérisation et une compréhension fondamentale des réponses de la matière en considération des deux variables thermodynamiques fondamentales, la température et la pression. » L'importance du ΔV pour le repliement protéique a été mise en lumière en 1914 quand Bridgman a

déplié l'ovalbumine sous pression. Mais étonnamment nous ne comprenons pas encore exactement comment la pression agit sur les protéines. Un paramétrage quantitatif des ΔV à la base des effets de la pression accompagnant les transitions conformationnelles des protéines sera nécessaire à l'avenir pour cette compréhension. De nombreuses études ont été faites sur des protéines globulaires, mais la complexité de ces structures rend difficile aujourd'hui le dégagement de règles fondamentales aptes à être appliquées à l'ensemble des protéines. Nous utilisons donc une protéine à répétition de motifs Ankyrine, pour laquelle la structure de chaque motif est identique et connue, et pour laquelle les interactions ne sont qu'à courtes distances. Bien que l'action de la pression sur les protéines n'ait jamais été clairement expliquée, l'hypothèse majoritaire pour expliquer le dépliement est que la pression pousse les molécules d'eau du solvant vers l'intérieur de la protéine, favorisant ainsi le plus faible volume : celui de l'état déplié. L'eau, entrant ainsi dans les cœurs hydrophobes, pourrait ne pas directement agir sur les interactions hydrophobes mais plutôt remplir les cavités, conduisant à une diminution du volume molaire protéique et à la disruption des interactions accompagnant le dépliement. Pour mieux comprendre le ΔV, nous tentons donc de disséquer ses contributions et de comprendre comment les cavités et vides conduisant à l'existence du ΔV sont distribués dans la structure. Nous souhaitons de plus comprendre si tous les vides sont importants dans la définition du ΔV ou au contraire seulement une petite partie d'entre eux. Enfin, en nous basant sur ce modèle simple et sur les mutants dont nous disposons, nous souhaitons comprendre si la taille, la séquence et la forme des protéines affectent l'hydratation différentielle qui contribue au ΔV.

Afin de répondre à cette question thermodynamique spécifique, nous nous intéressons dans ce travail à un modèle protéique simple, pour lequel le repliement est réversible, à deux états, sans intermédiaire visible : les protéines à répétitions de motifs ankyrine. Nank 1-7* est le plus grand des mutants en notre possession et est composé de 7 motifs répétés de séquences différentes mais de structures identiques. Par souci de nomenclature, les chiffres apparaissant dans le nom des protéines correspondent aux motifs présents : 1-7 pour les motifs 1 à 7, 2-7 pour les motifs 2 à 7, et ainsi de suite. Tous les mutants en taille et en séquence que nous avons utilisés ont été crées à partir de celui-ci. Nous avons étudié son dépliement à l'équilibre et en cinétiques à température ambiante (20°C) en fonction de la pression appliquée mais aussi à diverses températures (8, 12, 16, 20, 24 et 28°C) par fluorescence sous pression, grâce à son unique tryptophane au cœur du 5ème motif. Le dépliement est alors apparu comme étant à deux états, très coopératif, dans chacun des cas. Le ΔV à l'équilibre de Nank1-7* diminue en valeur absolue avec la température, avec une pente de -1.7 mL/mol/K, correspondant au $\Delta\alpha_u$. Par calorimétrie par perturbation par pression (PPC), nous avons déterminé l'expansivité de la protéine repliée, ainsi que son ΔV à 42°C, qui confirme le précédent résultat. Par l'étude des cinétiques de cette protéine, nous avons pu déterminer la différence de volume entre l'état déplié et l'ensemble d'états de

transition (TSE), c'est-à-dire son volume d'activation de repliement $\Delta V_f^{\#}$, mais aussi la différence de volume entre le TSE et l'état déplié, le $\Delta V_u^{\#}$. Grâce à ces données, nous avons pu apprendre que l'ensemble d'états de transition de cette protéine possède un volume très proche de celui de l'état replié, et que ce volume augmente avec la température, alors que le ΔV_{eq}^{f} diminue. Ainsi, à plus de 20°C, le TSE devient nettement plus volumineux que la protéine repliée. A plus haute température, le ΔV_{eq}^{f} de cette protéine change alors de signe et devient négatif, confirmant l'existence du « paradoxe du ΔV ». Par SAXS, nous avons déterminé que la pression agissait en dépliant la protéine, et en augmentant son rayon de giration, et ces données ont été complétées par du FTIR sous pression afin d'avoir une idée de l'évolution de la structuration de la protéine à 20°C à différentes pressions et concentrations d'urée. Il nous a ainsi été possible de déterminer la présence dans l'état déplié de la protéine sous pression de résidus d'hélices α, et de confirmer qu'à 3000 bars, avec peu d'urée, nous arrivons bien à la déplier. L'ensemble de ces données nous a amené à la formulation de l'hypothèse suivante : il semble que les changements d'hydratation apparaissant avec la formation de structures secondaires ne sont pas un paramètre à prendre en compte dans le ΔV_{eq}^{f}. Comme les hélices ne contiennent pas de vides, nous pouvons de plus être sûrs que le ΔV_{eq}^{f} ne peut pas changer du fait du mésappariement dans la structure secondaire. Ainsi, la seule possibilité est que le ΔV_{eq}^{f} apparaisse uniquement du fait de la structuration tridimensionnelle de la protéine. Nous avons finalement pu avoir une vue d'ensemble du comportement du ΔV_{eq}^{f}, $\Delta V_f^{\#}$ et $\Delta V_u^{\#}$ en fonction de la température. De ces résultats, nous avons appris que l'ensemble d'états de transition était energétiquement déplié, mais se comportait comme la forme repliée de la protéine sous pression, et que tout l'effet de la pression semblait s'exercer sur le TSE. Nous avons pu de même définir que l'effet de la pression sur les protéines n'est pas linéaire mais varie selon la pression pour une température donnée. Finalement, nous avons pu proposer une explication du paradoxe du ΔV grâce à l'ensemble de nos résultats et à des expériences complémentaires de densitométrie. Ainsi, il nous apparaît que le ΔV évolue car les expansivités des deux états sont différentes. Le ΔV peut ainsi changer de signe à une température donnée car le volume de la forme repliée de la protéine devient plus petit que celui de la forme dépliée à haute température, alors qu'elle est plus grande à basse température. Nous avons continué cette étude par l'analyse de différents variants de cette protéine : des mutants à délétions N- et C-terminales de 1, 2 ou 3 motifs, deux mutants à mutation ponctuelle d'une alanine changée en glycine dans deux motifs différents de Nank 1-7*, et deux mutants à 2 et 3 motifs respectivement.

Ces mutants nous ont appris de nombreuses choses quant aux effets de la pression sur ces protéines. En effet, nous avons défini le ΔV d'équilibre à 20°C de chacun de ces mutants (Nank 2-7*, Nank3-7*, Nank4-7*, Nank1-6*, Nank1-5*, NWC, NW, NAG5 et NAG2), et avons analysé la corrélation entre le nombre de motifs et le ΔV. De manière

intéressante, cette étude nous a indiqué que pour l'ensemble de ces mutants, le ΔV ne semble pas varier avec la taille des protéines, mis à part une légère augmentation du ΔV de Nank2-7* par rapport à Nank1-7*, et une légère diminution du ΔV de Nank4-7* par rapport à Nank3-7*. En revanche, les mutants de 2 et 3 motifs (respectivement NW et NWC), donc les séquences diffèrent largement, ont des ΔV nettement inférieurs. Il est important de noter ici que Nank4-7* possède toutefois un ΔV très similaire à celui de Nank1-7*. Nous pouvons comprendre ce ΔV équivalent entre Nank1-6* et Nank1-5* car le 6ème motif est connu pour être déplié en l'absence du 7ème. En revanche, il semble clair que pour l'ensemble des mutants étudiés, le ΔV ne peut être similaire que car l'hydratation des liaisons peptidiques n'est pas un facteur important dans la détermination du ΔV. Afin de mieux comprendre pourquoi les mutants par délétion N et C terminales montraient une similarité importante des ΔV, nous avons procédé à une caractérisation des cavités dans la structure protéique de Nank1-7*, dans le but de déterminer quelles sont les cavités définissant précisément le ΔV, à l'échelle atomique.

De manière intéressante, il nous et apparu que l'ensemble des cavités n'était pas à prendre en compte dans le paramétrage du ΔV à 20°C, étant donné que leur volume total approche 160 mL/mol, contrairement au volume expérimental de l'ordre de 70 mL/mol à cette même température. Une analyse fine de la proximité entre les premières molécules d'eau et les cavités dans la structure nous ont permis de définir certaines cavités très hautement déshydratées, à 6 Angströms ou plus de la première molécule d'eau et qui restent stable durant les simulations, alors que la majorité des cavités apparentes est transitoirement hydratée par des molécules d'eau qui y entrent et en ressortent. Ainsi, ces cavités hautement déshydratées à l'interface entre motifs, d'un volume moyen d'environ 10 mL/mol semblent correspondre aux cavités rentrant réellement en jeu dans la définition du ΔV. Ces cavités importantes pour le ΔV ne se trouvent en revanche pas à toutes les interfaces entre les motifs, mais uniquement entre les motifs 2-3, 3-4, 4-5, expliquant désormais clairement pourquoi les protéines Nank1-7*, Nank1-6* et Nank1-7* ont un volume identique, et pourquoi les ΔV ne varient pas entre Nank1-7*, Nank1-6*, Nank1-5*, Nank2-7*, et peu entre Nank2-7*, Nank3-7* et Nank 4-7*. De plus, des données bibliographiques indiquent que le premier motif, lorsqu'il est présent, servirait à « capper » la protéine, conduisant à une compaction de celle-ci, et donc à un ΔV plus petit pour Nank1-7* que pour Nank2-7*.

Le TSE de l'ankyrine n'est ressemblant à la forme native que dans les motifs centraux de la protéine, et le chemin de repliement connu correspond à cette séquence de motifs qui se replient : 4 – 4,5 – 3,4,5 – 2,3,4,5 – 1,2,3,4,5 – 1,2,3,4,5,6,7. Nous pouvons donc constater que les premiers motifs à se replier correspondent aux motifs entre lesquels sont présentes ces cavités hautement déshydratées. Il semble donc que ces cavités soient reliées à la mise en place de l'ensemble d'états de transition.

En étudiant les diagrammes volumétriques de l'ensemble de ces protéines à 20°C,

nous avons remarqué qu'effectivement, pour Nank1-7*, Nank1-6* et Nank1-5*, le volume de l'état replié reste relativement stable, alors que dès que nous enlevons le premier motif, les volumes de l'état replié et du TSE augmentent sensiblement, et que lorsque l'on enlève le septième motif, seul celui du TSE augmente, alors qu'aucune cavité n'est retirée, indiquant un relâchement de l'ensemble de la protéine lorsque le capping du 1er motif est annulé, et du TSE uniquement lorsque le capping du 7ème motif est annulé. Cette hypothèse étant renforcée par la baisse importante de stabilité de Nank1-6* par rapport à Nank1-7*, alors que les stabilités de Nank1-6* et Nank1-5* sont très similaires. Il semble donc encore une fois que les variations de volumes soient intimement reliées au TSE. A l'instar de cela, diminuer le nombre de motifs après clivage des premiers motifs N et C terminaux n'amène pas de variation du ΔV car les interfaces présentant les cavités ne sont pas supprimées, sauf dans le cas de Nank4-7*. Enfin, en analysant les $\Delta V_f^{\#}$ nous pouvons constater que Nank2-7*, Nank3-7*, Nank4-7* et Nank1-5* possèdent le même écart de volume entre le TSE et l'état déplié, indiquant qu'effectivement, les cavités donnant le ΔV apparaissent clairement dans le TSE.

Nous avons poursuivi ce travail par l'analyse des cinétiques de dépliement sur l'ensemble de notre gamme de température, afin de mieux comprendre l'évolution du ΔV, $\Delta V_f^{\#}$ et $\Delta V_u^{\#}$ avec la température. Nous avons constaté pour chacun de nos mutants que le TSE est systématiquement proche en terme de volume à l'état replié, pour l'ensemble des températures testées, mais que celui-ci devient de plus en plus volumineux avec l'augmentation de la température, jusqu'à devenir plus grand que l'état replié. Nous pouvons remarquer que le TSE semble définitivement être le point de départ de la mise en place du ΔV, alors que celui-ci est énergétiquement proche de l'état déplié.

Une étude de la proportion des acides aminés hydrophobiques et chargés dans chaque motif de la structure repliée de Nank1-7* nous a de plus confirmé que la présence de cavités prises en compte dans la définition du ΔV correspondait à un grand nombre d'acides aminés hydrophobes. Il semble donc clair que les cavités à l'origine de ΔV sont majoritairement dues à la force hydrophobe, et comme celle-ci semble apparaître au niveau du TSE, nous pouvons raisonnablement affirmer qu'elle est une force majeure du repliement protéique.

La première conclusion claire de ce travail est que les premiers motifs N et C-terminaux de Nank1-7* n'ajoutent aucune valeur au ΔV_{eq}^f. En effet, ajouter le motif 6 à Nank1-5* n'a aucun effet sur la valeur du ΔV_{eq}^f, alors qu'ajouter le motif 3 à Nank4-7* ou le motif 2 à Nank3-7* augmente ΔV_{eq}^f. De manière intéressante, ajouter le premier motif à Nank2-7* diminue la valeur de ΔV_{eq}^f de manière significative. Ce comportement inattendu des motifs terminaux est encore plus prononcé concernant le volume d'activation du repliement. Ainsi, retirer ces motifs induit un volume du TSE nettement plus grand.

Nous savons que retirer un motif à une de ces protéines va diminuer de manière équivalente le volume dû aux atomes pour tous les états de la protéine. L'hydratation des liaisons peptidiques exposées dans l'état déplié ne peut pas compter dans les différentes contributions de chaque motif au ΔV_{eq}^{f} car leurs structures et tailles sont identiques, et donc impliquent l'exposition du même nombre de liaisons peptidiques. L'expansivité thermique met en jeu une libération de l'eau d'hydratation vers le solvant lorsque l'on chauffe. Ainsi, l'expansivité thermique des composés polaires dans l'eau est large et positive à faible et moyenne température. Hydrater les résidus polaires nouvellement exposés est relié à une augmentation de la densité pour les nouvelles molécules d'eau interagissant, et donc à une diminution du volume. De là, l'hydratation de résidus polaires nouvellement exposés devrait contribuer négativement au ΔV_{eq}^{u}, de sorte que leurs suppressions devraient induire un ΔV_{eq}^{u} plus faible. L'inverse est vrai pour l'hydratation des cavités hydrophobes, leur suppression devant en principe conduire à un volume plus petit de l'état déplié et donc un plus grand ΔV_{eq}^{u}. Nous avons remarqué que les motifs entourant les cavités sont les motifs ayant le plus de résidus hydrophobes, ce qui implique un plus grand $\Delta\alpha$ lors de délétion de motifs, ce qui a été confirmé pat PPC pour plusieurs mutants. En nous basant sur ces résultats, nous pouvons donc émettre l'hypothèse que dans le cas de l'ankyrine, les propriétés volumétriques ne peuvent pas provenir d'effets dus à l'hydratation différentielle. Nos résultats indiquent que les cavités présentes dans la structure sont de l'ordre de 10 mL/mol/interface pour le ΔV_{eq}^{f}. Si la délétion des motifs 1 et 7 induit une diminution des interactions stabilisantes de l'état replié et du TSE, cela induirait donc une augmentation de leur volume. Les motifs centraux exhibent les plus grandes cavités et la plus faible hydratation, et les cavités les moins hydratées sont rencontrées entre les motifs 2-3, 3-4 et 4-5. Nous interprétons donc l'augmentation de volume du TSE et de l'état replié de Nank2-7* comme la résultante de la déstabilisation de l'interface entre les motifs 2 et 3, induisant une expansion de la structure. Notre investigation systématique de l'amplitude des effets de la pression sur les protéines nous indique que la contribution majeure des effets de la pression est l'empaquetage imparfait de la structure.

La première conclusion claire de ce travail est que l'hydratation des ponts hydrogènes et l'hydratation différentielle des acides aminés ne contribuent pas de manière significative au ΔV_{eq}^{f} pour les protéines à répétitions de motifs ankyrine. Au contraire, la contribution principale au ΔV_{eq}^{f} semble être l'existence de vides internes exclus du solvant qui ont une origine structurale et dynamique. Nous avons montré que les cavités les plus larges et sèches dans Nank1-7* sont trouvées à l'interface entre les motifs 2-3, 3-4 et 4-5, ce qui est totalement consistant avec la contribution prépondérante des motifs centraux de la protéine à la magnitude du ΔV_{eq}^{u}. Nous avons interprété la large augmentation du ΔV_{eq}^{u} lors de la délétion du premier motif comme étant due à une déstabilisation des interactions entre les motifs 2 et 3 et/ou une plus grande

expansivité de l'état replié de la protéine sans le premier motif.

Une expansion de la structure protéique, avec une diminution des interactions tertiaires, a été proposée comme constituant la première étape du processus de repliement [235]. L'effet des délétions de motifs ankyrine sur l'état replié est limité aux motifs 1 et 2, et tandis que l'effet volumétrique sur le TSE est important, celui-ci est faible à l'équilibre. Dans tous les cas, les conséquences énergétiques semblent assez limitées, ce qui est consistant avec la possible application de modèle Ising pour expliquer le paysage énergétique de la protéine complète, en se basant sur les stabilités de l'ensemble des mutants.

La sensibilité accrue des propriétés volumétriques du TSE à la délétion de motifs est cohérent avec notre hypothèse générale (basée sur l'observation d'un $\Delta V_f^{\#}$ généralement positif et large pour le repliement) que le TSE contient des régions pour lesquelles des interactions clé sont manquantes, et que dans certains cas son volume global peut être plus important que celui de l'état replié. Ceci a été observé pour Nank1-7* à 28°C ainsi que pour la plupart des mutants. Une investigation systématique du rôle de la taille et de la séquence dans la détermination de l'amplitude des effets de la pression sur les structures protéiques et leurs stabilités indique clairement la principale contribution au ΔV_{eq}^{u} provient de l'empaquetage imparfait des protéines.

Ces problèmes d'empaquetage apparaissent localisés très spécifiquement dans la structure 3D, conduisant à des cavités précises, et rendant ainsi les effets de la pression idiosyncrasiques, alors que la température et les dénaturants agissent uniformément sur les propriétés de la structure protéique.

Ainsi, le degré avec lequel la pression casse les interactions à travers la structure protéique nous donne des informations concernant la coopérativité intrinsèque du repliement. De manière intéressante, nous avons aussi trouvé que les propriétés énergétiques et dynamiques des conformations protéiques peuvent moduler de manière importante les propriétés volumétriques. Nous avons montré de plus que la perte d'interactions peut conduire à l'expansion des vides internes, tandis que l'augmentation en dynamique qui en résulte peut minimiser les vides par occlusion des chaines latérales ou par inclusion de molécules du solvant.

Tandis que la dynamique moléculaire peut être utilisée pour localiser les vides exclus du solvant avec une relative efficacité, il reste toutefois nettement plus difficile de déterminer leurs tailles, ainsi que la magnitude absolue du ΔV_{eq}^{f} car cela se trouve être fortement dépendant de l'expansivité, de la dynamique et de l'énergétique. Ainsi, nous avons déterminé, contrairement au point de vue habituel concernant le ΔV_{eq}^{f}, que nous devons prendre en compte le ΔV_{eq}^{f} comme étant une propriété hautement dynamique des protéines.

Dans cette étude, grâce à la fluorescence sous haute pression et à la densitométrie, nous avons de plus pu expliquer le "paradoxe du ΔV_{eq}^{f}"; en effet, le changement de signe du ΔV_{eq}^{f} provient directement des différences en expansivité thermique entre

l'état replié et déplié des protéines. Ainsi, les expansivités peuvent se croiser selon la température, et dès lors l'expansivité de la forme dépliée devient plus faible que celle de la forme repliée, impliquant que le ΔV_{eq}^{f} change de signe.

Même si la compréhension totale des effets de la pression sur les protéines est loin d'être établie, ce travail a pu y contribuer de manière importante.

Bibliography

[1] "So much more to know," *Science*, no. 309, pp. 78–102.

[2] H. Mu, Z.-R. Lü, D. Park, B.-C. Kim, J. Bhak, F. Zou, J.-M. Yang, S. Li, Y.-D. Park, H.-C. Zou, and H.-M. Zhou, "Kinetics of zn(2+)-induced brain type creatine kinase unfolding and aggregation," *Appl Biochem Biotechnol*, vol. 160, pp. 1309–20, Mar 2010.

[3] J. L. Silva, T. C. R. G. Vieira, M. P. B. Gomes, A. P. A. Bom, L. M. T. R. Lima, M. S. Freitas, D. Ishimaru, Y. Cordeiro, and D. Foguel, "Ligand binding and hydration in protein misfolding: Insights from studies of prion and p53 tumor suppressor proteins (dagger)," *Acc Chem Res*, vol. 43, pp. 271–9, Feb 2010.

[4] A. Surguchev and A. Surguchov, "Conformational diseases: looking into the eyes," *Brain Res Bull*, vol. 81, pp. 12–24, Jan 2010.

[5] L. Kupfer, W. Hinrichs, and M. H. Groschup, "Prion protein misfolding," *Curr Mol Med*, vol. 9, pp. 826–35, Sep 2009.

[6] N. J. Cobb and W. K. Surewicz, "Prion diseases and their biochemical mechanisms," *Biochemistry*, vol. 48, pp. 2574–85, Mar 2009.

[7] R. A. Moore, L. M. Taubner, and S. A. Priola, "Prion protein misfolding and disease," *Curr Opin Struct Biol*, vol. 19, pp. 14–22, Feb 2009.

[8] F. E. Cohen and S. B. Prusiner, "Pathologic conformations of prion proteins," *Annu Rev Biochem*, vol. 67, pp. 793–819, Jan 1998.

[9] J. Bieschke, P. Weber, N. Sarafoff, M. Beekes, A. Giese, and H. Kretzschmar, "Autocatalytic self-propagation of misfolded prion protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, p. 12207, Aug 2004.

[10] L. J. M. van Keulen, A. Bossers, and F. van Zijderveld, "Tse pathogenesis in cattle and sheep," *Vet Res*, vol. 39, p. 24, Jan 2008.

[11] S. B. Prusiner, "Prions," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 13363–83, Nov 1998.

[12] S. Balaji, R. Kalpana, and P. Shapshak, "Paradigm development: comparative and predictive 3d modeling of hiv-1 virion infectivity factor (vif)," *Bioinformation*, vol. 1, pp. 290–309, Jan 2006.

[13] B. R. Cullen, "Role and mechanism of action of the apobec3 family of antiretroviral resistance factors," *J Virol*, vol. 80, pp. 1067–76, Feb 2006.

[14] X. Yu, Y. Yu, B. Liu, K. Luo, W. Kong, P. Mao, and X.-F. Yu, "Induction of apobec3g ubiquitination and degradation by an hiv-1 vif-cul5-scf complex," *Science*, vol. 302, pp. 1056–60, Nov 2003.

[15] K. Shirakawa, A. Takaori-Kondo, M. Kobayashi, M. Tomonaga, T. Izumi, K. Fukunaga, A. Sasada, A. Abudu, Y. Miyauchi, H. Akari, K. Iwai, and T. Uchiyama, "Ubiquitination of apobec3 proteins by the vif-cullin5-elonginb-elonginc complex," *Virology*, vol. 344, pp. 263–6, Jan 2006.

[16] J. R. Banavar and A. Maritan, "Physics of proteins," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 36, pp. 261–80, Jan 2007.

[17] R. L. Baldwin, "Protein folding from 1961 to 1982," *Nat Struct Biol*, vol. 6, pp. 814–7, Sep 1999.

[18] C. B. Anfinsen, "Principles that govern the folding of protein chains," *Science*, vol. 181, pp. 223–30, Jul 1973.

[19] C. B. Anfinsen and H. A. SCHERAGA, "Experimental and theoretical aspects of protein folding," *Adv Protein Chem*, vol. 29, pp. 205–300, Jan 1975.

[20] R. L. Baldwin, "Intermediates in protein folding reactions and the mechanism of protein folding," *Annu Rev Biochem*, vol. 44, pp. 453–75, Jan 1975.

[21] R. L. Baldwin, "How does protein folding get started?," *Trends Biochem Sci*, vol. 14, pp. 291–4, Jul 1989.

[22] P. S. Kim and R. L. Baldwin, "Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding," *Annu Rev Biochem*, vol. 51, pp. 459–89, Jan 1982.

[23] R. Jaenicke, "Folding and association of proteins," *Prog Biophys Mol Biol*, vol. 49, pp. 117–237, Jan 1987.

[24] D. Baker, "A surprising simplicity to protein folding," *Nature*, vol. 405, pp. 39–42, May 2000.

[25] R. L. Baldwin and B. H. Zimm, "Are denatured proteins ever random coils?," *Proc Natl Acad Sci USA*, vol. 97, pp. 12391–2, Nov 2000.

[26] T. E. Creighton, "Protein folding," *Biochem J*, vol. 270, pp. 1–16, Aug 1990.

[27] V. Daggett and A. R. Fersht, "Is there a unifying mechanism for protein folding?," *Trends Biochem Sci*, vol. 28, pp. 18–25, Jan 2003.

[28] R. L. Baldwin, "How hofmeister ion interactions affect protein stability," *Biophys J*, vol. 71, pp. 2056–63, Oct 1996.

[29] P. Luo and R. L. Baldwin, "Interaction between water and polar groups of the helix backbone: an important determinant of helix propensities," *Proc Natl Acad Sci USA*, vol. 96, pp. 4930–5, Apr 1999.

[30] H. A. SCHERAGA, G. NEMETHY, and I. Z. STEINBERG, "The contribution of hydrophobic bonds to the thermal stability of protein conformations," *J Biol Chem*, vol. 237, pp. 2506–8, Aug 1962.

[31] K. A. Dill, "Dominant forces in protein folding," *Biochemistry*, vol. 29, pp. 7133–55, Aug 1990.

[32] H. Eyring, "The activated complex in chemical reactions," *The Journal of chemical physics*, vol. 3, p. 107, Jan 1935.

[33] C. B. Iii, J. Onuchic, and D. Wales, "Statistical thermodynamics: Taking a walk on a landscape," *Science*, vol. 293, p. 612, Jul 2001.

[34] J. N. Onuchic, P. G. Wolynes, Z. Luthey-Schulten, and N. D. Socci, "Toward an outline of the topography of a realistic protein-folding funnel," *Proc Natl Acad Sci USA*, vol. 92, pp. 3626–30, Apr 1995.

[35] R. M. Daniel, M. Dines, and H. H. Petach, "The denaturation and degradation of stable enzymes at high temperatures," *Biochem J*, vol. 317 (Pt 1), pp. 1–11, Jul 1996.

[36] A. A. Nickson, K. E. Stoll, and J. Clarke, "Folding of a lysm domain: entropy-enthalpy compensation in the transition state of an ideal two-state folder," *Journal of Molecular Biology*, vol. 380, pp. 557–69, Jul 2008.

[37] E. Fisicaro, C. Compari, and A. Braibanti, "Entropy/enthalpy compensation: hydrophobic effect, micelles and protein complexes," *Phys. Chem. Chem. Phys.*, no. 6, pp. 4156–4166, 2004.

[38] K. Sharp, "Entropy-enthalpy compensation: fact or artifact?," *Protein Sci*, vol. 10, pp. 661–7, Mar 2001.

[39] A. Cooper, "Thermodynamic fluctuations in protein molecules," *Proc Natl Acad Sci USA*, vol. 73, pp. 2740–1, Aug 1976.

[40] A. L. Tournier, V. Réat, R. Dunn, R. Daniel, J. C. Smith, and J. Finney, "Temperature and timescale dependence of protein dynamics in methanol : water mixtures," *Phys Chem Chem Phys*, vol. 7, pp. 1388–93, Apr 2005.

[41] T. V. Chalikian, "Volumetric properties of proteins," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 32, pp. 207–35, Jan 2003.

[42] O. Bieri and T. Kiefhaber, "Elementary steps in protein folding," *Biol Chem*, vol. 380, pp. 923–9, Jan 1999.

[43] S. Williams, T. P. Causgrove, R. Gilmanshin, K. S. Fang, R. H. Callender, W. H. Woodruff, and R. B. Dyer, "Fast events in protein folding: Helix melting and formation in a small peptide ‚Ä†," *Biochemistry*, vol. 35, pp. 691–697, Jan 1996.

[44] E. O. Armah, "Mathematics of protein pathological misfolding," *Math Biosci*, vol. 208, pp. 1–25, Jul 2007.

[45] C. Clementi, H. Nymeyer, and J. N. Onuchic, "Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins," *J Mol Biol*, vol. 298, pp. 937–53, May 2000.

[46] P. Wolynes, Z. Luthey-Schulten, and J. Onuchic, "Fast-folding experiments and the topography of protein folding energy landscapes," *Chem Biol*, vol. 3, pp. 425–32, Jun 1996.

[47] L. Qiu, S. A. Pabit, s. Adrian E Roitberg, ‚Ä°, , and S. J. Hagen, "Smaller and faster: The 20-residue trp-cage protein folds in 4 [mu]s,"

[48] P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, "Navigating the folding routes," *Science*, vol. 267, pp. 1619–20, Mar 1995.

[49] B. B. Boonyaratanakornkit, C. B. Park, and D. S. Clark, "Pressure effects on intra- and intermolecular interactions within proteins," *Biochim Biophys Acta*, vol. 1595, pp. 235–49, Mar 2002.

[50] V. V. Mozhaev, K. Heremans, J. Frank, P. Masson, and C. Balny, "High pressure effects on protein structure and function," *Proteins*, vol. 24, pp. 81–91, Jan 1996.

[51] S. A. Hawley, "Reversible pressure-temperature denaturation of chymotrypsinogen," *Biochemistry*, vol. 10, pp. 2436–2442, Jun 1971.

[52] J. F. Brandts, R. J. Oliveira, and C. Westort, "Thermodynamics of protein denaturation. effect of pressure on the denaturation of ribonuclease a," *Biochemistry*, vol. 9, pp. 1038–47, Feb 1970.

[53] A. Zipp and W. KAUZMANN, "Pressure denaturation of metmyoglobin.," *Biochemistry*, vol. 12, no. 21, pp. 4217–28, 1973.

[54] R. Winter and W. Dzwolak, "Exploring the temperature-pressure configurational landscape of biomolecules: from lipid membranes to proteins," *Philos Transact A Math Phys Eng Sci*, vol. 363, pp. 537–62; discussion 562–3, Feb 2005.

[55] K. Heremans and L. Smeller, "Protein structure and dynamics at high pressure," *Biochim Biophys Acta*, vol. 1386, pp. 353–70, Aug 1998.

[56] R. Ravindra and R. Winter, "On the temperature–pressure free-energy landscape of proteins," *Chemphyschem : a European journal of chemical physics and physical chemistry*, vol. 4, pp. 359–65, Apr 2003.

[57] F. Meersman, L. Smeller, and K. Heremans, "Protein stability and dynamics in the pressure-temperature plane," *Biochim Biophys Acta*, vol. 1764, pp. 346–54, Mar 2006.

[58] L. Smeller, F. Meersman, and K. Heremans, "Refolding studies using pressure: the folding landscape of lysozyme in the pressure-temperature plane," *Biochim Biophys Acta*, vol. 1764, pp. 497–505, Mar 2006.

[59] N. W. A. van Uden, H. Hubel, D. A. Faux, D. J. Dunstan, and C. A. Royer, "Negative effective pressures in liquid mixtures," *International Journal of High Pressure Research*, vol. 23, no. 3, pp. 205–209, 2003.

[60] A. Kato, M. Yamada, S. Nakamura, S.-I. Kidokoro, and Y. Kuroda, "Thermodynamic properties of bpti variants with highly simplified amino acid sequences," *J Mol Biol*, vol. 372, pp. 737–46, Sep 2007.

[61] G. I. Makhatadze and P. L. Privalov, "Contribution of hydration to protein folding thermodynamics. i. the enthalpy of hydration," *J Mol Biol*, vol. 232, pp. 639–59, Jul 1993.

[62] P. L. Privalov, "Stability of proteins: small globular proteins," *Adv Protein Chem*, vol. 33, pp. 167–241, Jan 1979.

[63] V. J. Hilser, J. Gómez, and E. Freire, "The enthalpy change in protein folding and binding: refinement of parameters for structure-based calculations," *Proteins*, vol. 26, pp. 123–33, Oct 1996.

[64] P. J. Fleming and G. D. Rose, "Do all backbone polar groups in proteins form hydrogen bonds?," *Protein Sci*, vol. 14, pp. 1911–7, Jul 2005.

[65] M. Blaber, J. D. Lindstrom, N. Gassner, J. Xu, D. W. Heinz, and B. W. Matthews, "Energetic cost and structural consequences of burying a hydroxyl group within

the core of a protein determined from ala–>ser and val–>thr substitutions in t4 lysozyme," *Biochemistry*, vol. 32, pp. 11363–73, Oct 1993.

[66] W.-G. Han, K. J. Jalkanen, , M. Elstner, and S. Suhai, "Theoretical study of aqueous n-acetyl-l-alanine n'-methylamide: Structures and raman, vcd, and roa spectra," *J. Phys. Chem. B*, vol. 102, no. 14, pp. 2587–2602, 1998.

[67] J. J. Dannenberg, "Enthalpies of hydration of n-methylacetamide by one, two, and three waters and the effect upon the c=o stretching frequency. an ab initio dft study," *J Phys Chem A*, vol. 110, pp. 5798–802, May 2006.

[68] P. Salvador, A. Asensio, and J. J. Dannenberg, "The effect of aqueous solvation upon alpha-helix formation for polyalanines," *The journal of physical chemistry B*, vol. 111, pp. 7462–6, Jun 2007.

[69] D. Sitkoff, K. A. Sharp, and B. Honig, "Accurate calculation of hydration free energies using macroscopic solvent models," *J. Phys. Chem.*, vol. 98, no. 7, pp. 1978–1988, 1994.

[70] G. I. Makhatadze and P. L. Privalov, "Energetics of protein structure," *Adv Protein Chem*, vol. 47, pp. 307–425, Jan 1995.

[71] T. Lazaridis, G. Archontis, and M. Karplus, "Enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics," *Adv Protein Chem*, vol. 47, pp. 231–306, Jan 1995.

[72] C. Tanford, "Protein denaturation," *Adv Protein Chem*, vol. 23, pp. 121–282, Jan 1968.

[73] C. Tanford, "Protein denaturation. c. theoretical models for the mechanism of denaturation," *Adv Protein Chem*, vol. 24, pp. 1–95, Jan 1970.

[74] R. L. Baldwin, "Energetics of protein folding," *J Mol Biol*, vol. 371, pp. 283–301, Aug 2007.

[75] J. M. Sturtevant, "Heat capacity and entropy changes in processes involving proteins," *Proc Natl Acad Sci USA*, vol. 74, pp. 2236–40, Jun 1977.

[76] R. L. Baldwin, "Temperature dependence of the hydrophobic interaction in protein folding.," *Proc Natl Acad Sci USA*, vol. 83, p. 8069, Nov 1986.

[77] G. I. Makhatadze and P. L. Privalov, "On the entropy of protein folding," *Protein Sci*, vol. 5, pp. 507–10, Mar 1996.

[78] G. Némethy, S. J. Leach, and H. A. Scheraga, "The influence of amino acid side chains on the free energy of helix-coil transitions," *J Phys Chem*, vol. 70, no. 4, pp. 998–1004, 1966.

[79] J. D. Bryngelson and P. G. Wolynes, "Spin glasses and the statistical mechanics of protein folding," *Proc Natl Acad Sci USA*, vol. 84, pp. 7524–8, Nov 1987.

[80] W. KAUZMANN, "Some factors in the interpretation of protein denaturation," *Adv Protein Chem*, vol. 14, pp. 1–63, Jan 1959.

[81] L. Fu and E. Freire, "On the origin of the enthalpy and entropy convergence temperatures in protein folding," *Proc Natl Acad Sci USA*, vol. 89, pp. 9335–8, Oct 1992.

[82] B. Lee, "Isoenthalpic and isoentropic temperatures and the thermodynamics of protein denaturation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, pp. 5154–8, Jun 1991.

[83] K. L. Schweiker, V. W. Fitz, and G. I. Makhatadze, "Universal convergence of the specific volume changes of globular proteins upon unfolding," *Biochemistry*, vol. 48, pp. 10846–51, Nov 2009.

[84] Y. V. Griko, E. Freire, G. Privalov, H. van Dael, and P. L. Privalov, "The unfolding thermodynamics of c-type lysozymes: a calorimetric study of the heat denaturation of equine lysozyme," *Journal of Molecular Biology*, vol. 252, pp. 447–59, Sep 1995.

[85] H. Seemann, R. Winter, and C. A. Royer, "Volume, expansivity and isothermal compressibility changes associated with temperature and pressure unfolding of staphylococcal nuclease," *Journal of Molecular Biology*, vol. 307, pp. 1091–102, Apr 2001.

[86] J. K. Myers, C. N. Pace, and J. M. Scholtz, "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding," *Protein Sci*, vol. 4, pp. 2138–48, Oct 1995.

[87] K. P. Murphy and E. Freire, "Thermodynamics of structural stability and cooperative folding behavior in proteins," *Adv Protein Chem*, vol. 43, pp. 313–61, Jan 1992.

[88] J. Gómez, V. J. Hilser, D. Xie, and E. Freire, "The heat capacity of proteins," *Proteins*, vol. 22, pp. 404–12, Aug 1995.

[89] N. V. Prabhu and K. A. Sharp, "Heat capacity in proteins," *Annu Rev Phys Chem*, vol. 56, pp. 521–48, Jan 2005.

[90] R. Hawkes, M. G. Grutter, and J. Schellman, "Thermodynamic stability and point mutations of bacteriophage t4 lysozyme," *J Mol Biol*, vol. 175, pp. 195–212, May 1984.

[91] G. Desai, G. Panick, M. Zein, R. Winter, and C. A. Royer, "Pressure-jump studies of the folding/unfolding of trp repressor," *J Mol Biol*, vol. 288, pp. 461–75, May 1999.

[92] L. Mitra, N. Smolin, R. Ravindra, C. Royer, and R. Winter, "Pressure perturbation calorimetric studies of the solvation properties and the thermal unfolding of proteins in solution–experiments and theoretical interpretation," *Physical chemistry chemical physics : PCCP*, vol. 8, pp. 1249–65, Mar 2006.

[93] C. A. Royer, "Revisiting volume changes in pressure-induced protein unfolding," *Biochim Biophys Acta*, vol. 1595, pp. 201–9, Mar 2002.

[94] M. et Tayauchi M et Sakurai M et Nitta K Dewa, "Compression refolding of cyrochrome c," *Prot. Pept. Lett.*, vol. 5, no. 5, pp. 265–268, 1998.

[95] R. M. Chura-Chambi, L. A. Genova, R. Affonso, and L. Morganti, "Refolding of endostatin from inclusion bodies using high hydrostatic pressure," *Anal Biochem*, vol. 379, pp. 32–9, Aug 2008.

[96] L.-N. Lin, J. F. Brandts, J. M. Brandts, and V. Plotnikov, "Determination of the volumetric properties of proteins and other solutes using pressure perturbation calorimetry," *Anal Biochem*, vol. 302, pp. 144–60, Mar 2002.

[97] L. Mitra, J.-B. Rouget, B. Garcia-Moreno, C. A. Royer, and R. Winter, "Towards a quantitative understanding of protein hydration and volumetric properties," *Chemphyschem : a European journal of chemical physics and physical chemistry*, vol. 9, pp. 2715–21, Dec 2008.

[98] A. A. Paladini and G. Weber, "Pressure-induced reversible dissociation of enolase," *Biochemistry*, vol. 20, pp. 2587–2593, Apr 1981.

[99] S. J. Perkins, "Protein volumes and hydration effects. the calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences," *Eur J Biochem*, vol. 157, pp. 169–80, May 1986.

[100] C. A. Royer, "Insights into the role of hydration in protein structure and stability obtained through hydrostatic pressure studies," *Braz J Med Biol Res*, vol. 38, pp. 1167–73, Aug 2005.

[101] T. V. Chalikian and K. J. Bresiauer, "On volume changes accompanying conformational transitions of biopolymers," *Biopolymers*, vol. 39, pp. 619–26, Nov 1996.

[102] R. Mohana-Borges, J. L. Silva, and G. de Prat-Gay, "Protein folding in the absence of chemical denaturants. reversible pressure denaturation of the noncovalent complex formed by the association of two protein fragments," *J Biol Chem*, vol. 274, pp. 7732–40, Mar 1999.

[103] M. H. Jacob, C. Saudan, G. Holtermann, A. Martin, D. Perl, A. E. Merbach, and F. X. Schmid, "Water contributes actively to the rapid crossing of a protein unfolding barrier," *Journal of Molecular Biology*, vol. 318, pp. 837–45, May 2002.

[104] R. Day and A. E. García, "Water penetration in the low and high pressure native states of ubiquitin," *Proteins*, vol. 70, pp. 1175–84, Mar 2008.

[105] R. Kitahara, C. Royer, H. Yamada, M. Boyer, J.-L. Saldana, K. Akasaka, and C. Roumestand, "Equilibrium and pressure-jump relaxation studies of the conformational transitions of p13mtcp1," *Journal of Molecular Biology*, vol. 320, no. 3, pp. 609–628, 2002.

[106] K. Sasahara and K. Nitta, "Pressure-induced unfolding of lysozyme in aqueous guanidinium chloride solution," *Protein Sci*, vol. 8, pp. 1469–74, Jul 1999.

[107] L. Brun, D. G. Isom, P. Velu, B. García-Moreno, and C. A. Royer, "Hydration of the folding transition state ensemble of a protein," *Biochemistry*, vol. 45, pp. 3473–80, Mar 2006.

[108] K. J. Frye and C. A. Royer, "Probing the contribution of internal cavities to the volume change of protein unfolding under pressure," *Protein Sci*, vol. 7, pp. 2217–22, Oct 1998.

[109] K. L. Schweiker and G. I. Makhatadze, "Methods in enzymology," vol. 466, pp. 527–547, Jan 2009.

[110] A. D. Tsamaloukas, N. K. Pyzocha, and G. I. Makhatadze, "Pressure perturbation calorimetry of unfolded proteins (‚Äì†)," *The journal of physical chemistry B*, Sep 2010.

[111] G. Panick, G. J. Vidugiris, R. Malessa, G. Rapp, R. Winter, and C. A. Royer, "Exploring the temperature-pressure phase diagram of staphylococcal nuclease," *Biochemistry*, vol. 38, pp. 4157–64, Mar 1999.

[112] R. Ravindra, C. Royer, and R. Winter, "Pressure perturbation calorimetic studies of the solvation properties and the thermal unfolding of staphylococcal nuclease," *Physical chemistry chemical physics : PCCP*, vol. 6, pp. 1952–1961, Jan 2004.

[113] J.-B. Rouget, M. A. Schroer, C. Jeworrek, M. Puhse, J.-L. Saldana, Y. Bessin, M. Tolan, D. Barrick, R. Winter, and C. A. Royer, "Unique features of the folding landscape of a repeat protein revealed by pressure perturbation," *Biophysj*, vol. 98, pp. 2712–2721, Feb 2010.

[114] W. Dzwolak, R. Ravindra, J. Lendermann, and R. Winter, "Aggregation of bovine insulin probed by dsc/ppc calorimetry and ftir spectroscopy," *Biochemistry*, vol. 42, pp. 11347–55, Sep 2003.

[115] R. Ravindra and R. Winter, "Pressure perturbation calorimetry: a new technique provides surprising results on the effects of co-solvents on protein solvation and unfolding behaviour," *Chemphyschem : a European journal of chemical physics and physical chemistry*, vol. 5, pp. 566–71, Apr 2004.

[116] H. Oshima, T. Yoshidome, K. ichi Amano, and M. Kinoshita, "A theoretical analysis on characteristics of protein structures induced by cold denaturation," *The Journal of chemical physics*, vol. 131, p. 205102, Nov 2009.

[117] W. KAUZMANN, "Thermodynamics of unfolding," *Nature*, no. 325, pp. 763–764, 1987.

[118] J. Cheftel, "Applications des hautes pressions en technologie alimentaire," *Industries alimentaires et agricoles*, vol. 108, pp. 141–153, Jan 1991.

[119] P. Masson, C. Tonello, and C. Balny, "High-pressure biotechnology in medicine and pharmaceutical science," *J Biomed Biotechnol*, vol. 1, pp. 85–88, Jan 2001.

[120] P. Bridgman, "The coagulation of albumen by pressure," *Journal of Biological Chemistry*, vol. 19, no. 4, pp. 511–512, 1914.

[121] K. Akasaka and H. Li, "Low-lying excited states of proteins revealed from non-linear pressure shifts in 1h and 15n nmr," *Biochemistry*, vol. 40, pp. 8665–71, Jul 2001.

[122] E. Bismuto, G. Irace, I. Sirangelo, and E. Gratton, "Pressure-induced perturbation of ans-apomyoglobin complex: frequency domain fluorescence studies on native and acidic compact states," *Protein Sci*, vol. 5, pp. 121–6, Jan 1996.

[123] R. Kitahara, H. Yamada, K. Akasaka, and P. E. Wright, "High pressure nmr reveals that apomyoglobin is an equilibrium mixture from the native to the unfolded," *J Mol Biol*, vol. 320, pp. 311–9, Jul 2002.

[124] O. Sire, B. Alpert, and C. A. Royer, "Probing ph and pressure effects on the apomyoglobin heme pocket with the 2'-(n,n-dimethylamino)-6-naphthoyl-4-trans-cyclohexanoic acid fluorophore," *Biophys J*, vol. 70, pp. 2903–14, Jun 1996.

[125] R. Kitahara, C. Royer, H. Yamada, M. Boyer, J.-L. Saldana, K. Akasaka, and C. Roumestand, "Equilibrium and pressure-jump relaxation studies of the conformational transitions of p13mtcp1," *J Mol Biol*, vol. 320, pp. 609–28, Jul 2002.

[126] Q. Ruan, K. Ruan, C. Balny, M. Glaser, and W. W. Mantulin, "Protein folding pathways of adenylate kinase from e. coli: hydrostatic pressure and stopped-flow studies," *Biochemistry*, vol. 40, pp. 14706–14, Dec 2001.

[127] K. Ruan, C. Xu, Y. Yu, J. Li, R. Lange, N. Bec, and C. Balny, "Pressure-exploration of the 33-kda protein from the spinach photosystem ii particle," *Eur J Biochem*, vol. 268, pp. 2742–50, May 2001.

[128] C. A. Royer, A. P. Hinck, S. N. Loh, K. E. Prehoda, X. Peng, J. Jonas, and J. L. Markley, "Effects of amino acid substitutions on the pressure denaturation of staphylococcal nuclease as monitored by fluorescence and nuclear magnetic resonance spectroscopy," *Biochemistry*, vol. 32, pp. 5222–32, May 1993.

[129] G. Panick and R. Winter, "Pressure-induced unfolding/refolding of ribonuclease a: static and kinetic fourier transform infrared spectroscopy study," *Biochemistry*, vol. 39, pp. 1862–9, Feb 2000.

[130] A. C. Oliveira, L. P. Gaspar, A. T. D. Poian, and J. L. Silva, "Arc repressor will not denature under pressure in the absence of water," *J Mol Biol*, vol. 240, pp. 184–7, Jul 1994.

[131] G. Mei, A. D. Venere, F. M. Campeggi, G. Gilardi, N. Rosato, F. D. Matteis, and A. Finazzi-Agrò, "The effect of pressure and guanidine hydrochloride on azurins mutated in the hydrophobic core," *Eur J Biochem*, vol. 265, pp. 619–26, Oct 1999.

[132] M. W. Lassalle, H. Yamada, H. Morii, K. Ogata, A. Sarai, and K. Akasaka, "Filling a cavity dramatically increases pressure stability of the c-myb r2 subdomain," *Proteins*, vol. 45, pp. 96–101, Oct 2001.

[133] M. H. Jacob, C. Saudan, G. Holtermann, A. Martin, D. Perl, A. E. Merbach, and F. X. Schmid, "Water contributes actively to the rapid crossing of a protein unfolding barrier," *J Mol Biol*, vol. 318, pp. 837–45, May 2002.

[134] H. Herberhold, S. Marchal, R. Lange, C. H. Scheyhing, R. F. Vogel, and R. Winter, "Characterization of the pressure-induced intermediate and unfolded state of red-shifted green fluorescent protein–a static and kinetic ftir, uv/vis and fluorescence spectroscopy study," *J Mol Biol*, vol. 330, pp. 1153–64, Jul 2003.

[135] A. L. CURL and E. F. JANSEN, "Effect of high pressures on trypsin and chymotrypsin," *J Biol Chem*, vol. 184, pp. 45–54, May 1950.

[136] K. Akasaka, H. Li, H. Yamada, R. Li, T. Thoresen, and C. K. Woodward, "Pressure response of protein backbone structure. pressure-induced amide 15n chemical shifts in bpti," *Protein Sci*, vol. 8, pp. 1946–53, Oct 1999.

[137] N. Smolin and R. Winter, "A molecular dynamics simulation of snase and its hydration shell at high temperature and high pressure," *Biochim Biophys Acta*, vol. 1764, pp. 522–34, Mar 2006.

[138] G. I. Makhatadze and P. L. Privalov, "Protein interactions with urea and guanidinium chloride. a calorimetric study," *J Mol Biol*, vol. 226, pp. 491–505, Jul 1992.

[139] N. N. Khechinashvili, "Thermodynamic properties of globular proteins and the principle of stabilization of their native structure," *Biochim Biophys Acta*, vol. 1040, pp. 346–54, Sep 1990.

[140] J. R. Livingstone, R. S. Spolar, and M. T. Record, "Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area," *Biochemistry*, vol. 30, pp. 4237–44, Apr 1991.

[141] G. Hummer, S. Garde, A. E. García, M. E. Paulaitis, and L. R. Pratt, "The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins," *Proc Natl Acad Sci USA*, vol. 95, pp. 1552–5, Feb 1998.

[142] Y. Levy and J. N. Onuchic, "Water mediation in protein folding and molecular recognition," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 35, pp. 389–415, Jan 2006.

[143] V. P. Grantcharova and D. Baker, "Circularization changes the folding transition state of the src sh3 domain," *J Mol Biol*, vol. 306, pp. 555–63, Feb 2001.

[144] J. L. Neira, L. S. Itzhaki, A. G. Ladurner, B. Davis, G. de Prat Gay, and A. R. Fersht, "Following co-operative formation of secondary and tertiary structure in a single protein module," *J Mol Biol*, vol. 268, pp. 185–97, Apr 1997.

[145] J. Woenckhaus, R. Köhling, P. Thiyagarajan, K. C. Littrell, S. Seifert, C. A. Royer, and R. Winter, "Pressure-jump small-angle x-ray scattering detected kinetics of staphylococcal nuclease folding," *Biophys J*, vol. 80, pp. 1518–23, Mar 2001.

[146] E. Kloss, N. Courtemanche, and D. Barrick, "Repeat-protein folding: new insights into origins of cooperativity, stability, and topology," *Arch Biochem Biophys*, vol. 469, pp. 83–99, Jan 2008.

[147] K. W. Tripp and D. Barrick, "The tolerance of a modular protein to duplication and deletion of internal repeats," *J Mol Biol*, vol. 344, pp. 169–78, Nov 2004.

[148] M. E. Zweifel and D. Barrick, "Studies of the ankyrin repeats of the drosophila melanogaster notch receptor. 1. solution conformational and hydrodynamic properties," *Biochemistry*, vol. 40, pp. 14344–56, Dec 2001.

[149] S. Artavanis-Tsakonas, M. D. Rand, and R. J. Lake, "Notch signaling: cell fate control and signal integration in development," *Science*, vol. 284, pp. 770–6, Apr 1999.

[150] S. Bray, "A notch affair," *Cell*, vol. 93, pp. 499–503, May 1998.

[151] M. A. Muskavitch, "Delta-notch signaling and drosophila cell fate choice," *Dev Biol*, vol. 166, pp. 415–30, Dec 1994.

[152] M. E. Zweifel and D. Barrick, "Studies of the ankyrin repeats of the drosophila melanogaster notch receptor. 2. solution stability and cooperativity of unfolding," *Biochemistry*, vol. 40, pp. 14357–67, Dec 2001.

[153] K. W. Tripp and D. Barrick, "Enhancing the stability and folding rate of a repeat protein through the addition of consensus repeats," *J Mol Biol*, vol. 365, pp. 1187–200, Jan 2007.

[154] O. Y. Lubman, M. X. G. Ilagan, R. Kopan, and D. Barrick, "Quantitative dissection of the notch:csl interaction: insights into the notch-mediated transcriptional switch," *J Mol Biol*, vol. 365, pp. 577–89, Jan 2007.

[155] D. C. Desrosiers and Z. yu Peng, "A binding free energy hot spot in the ankyrin repeat protein gabpbeta mediated protein-protein interaction," *J Mol Biol*, vol. 354, pp. 375–84, Nov 2005.

[156] A. H. Batchelor, D. E. Piper, F. C. de la Brousse, S. L. McKnight, and C. Wolberger, "The structure of gabpalpha/beta: an ets domain- ankyrin repeat heterodimer bound to dna," *Science*, vol. 279, pp. 1037–41, Feb 1998.

[157] K. S. Tang, A. R. Fersht, and L. S. Itzhaki, "Sequential unfolding of ankyrin repeats in tumor suppressor p16," *Structure*, vol. 11, pp. 67–73, Jan 2003.

[158] R. Venkataramani, K. Swaminathan, and R. Marmorstein, "Crystal structure of the cdk4/6 inhibitory protein p18ink4c provides insights into ankyrin-like repeat structure/function and tumor-derived p16ink4 mutations," *Nat. Struct Biol.*, vol. 5, pp. 74–81, Jan 1998.

[159] T. Sakai, L. Liu, Y. Shishido, and K. Fukui, "Identification of a novel, embryonal carcinoma cell-associated molecule, nucling, that is up-regulated during cardiac muscle differentiation," *J Biochem*, vol. 133, pp. 429–36, Apr 2003.

[160] D. Barrick, "Biological regulation via ankyrin repeat folding," *ACS Chem Biol*, vol. 4, pp. 19–22, Jan 2009.

[161] L. Breeden and K. Nasmyth, "Similarity between cell-cycle genes of budding yeast and fission yeast and the notch gene of drosophila," *Nature*, vol. 329, pp. 651–4, Jan 1987.

[162] K. A. Wharton, K. M. Johansen, T. Xu, and S. Artavanis-Tsakonas, "Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing egf-like repeats," *Cell*, vol. 43, pp. 567–81, Dec 1985.

[163] P. Bork, "Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally?," *Proteins*, vol. 17, pp. 363–74, Dec 1993.

[164] S. G. Sedgwick and S. J. Smerdon, "The ankyrin repeat: a diversity of interactions on a common structural framework," *Trends Biochem Sci*, vol. 24, pp. 311–6, Aug 1999.

[165] R. Foord, I. A. Taylor, S. G. Sedgwick, and S. J. Smerdon, "X-ray structural analysis of the yeast cell cycle regulator swi6 reveals variations of the ankyrin fold and has implications for swi6 function," *Nat Struct Biol*, vol. 6, pp. 157–65, Feb 1999.

[166] S. Gorina and N. P. Pavletich, "Structure of the p53 tumor suppressor bound to the ankyrin and sh3 domains of 53bp2," *Science*, vol. 274, pp. 1001–5, Nov 1996.

[167] T. Huxford, D. B. Huang, S. Malek, and G. Ghosh, "The crystal structure of the ikappabalpha/nf-kappab complex reveals mechanisms of nf-kappab inactivation," *Cell*, vol. 95, pp. 759–70, Dec 1998.

[168] M. D. Jacobs and S. C. Harrison, "Structure of an ikappabalpha/nf-kappab complex," *Cell*, vol. 95, pp. 749–58, Dec 1998.

[169] M. E. Zweifel, D. J. Leahy, F. M. Hughson, and D. Barrick, "Structure and stability of the ankyrin domain of the drosophila notch receptor," *Protein Sci*, vol. 12, pp. 2622–32, Nov 2003.

[170] S. Stifani, C. M. Blaumueller, N. J. Redhead, R. E. Hill, and S. Artavanis-Tsakonas, "Human homologs of a drosophila enhancer of split gene product define a novel family of nuclear proteins," *Nat Genet*, vol. 2, pp. 119–27, Oct 1992.

[171] O. Y. Lubman, R. Kopan, G. Waksman, and S. Korolev, "The crystal structure of a partial mouse notch-1 ankyrin domain: Repeats 4 through 7 preserve an ankyrin fold," *Protein Sci.*, vol. 14, pp. 1274–1281, May 2005.

[172] C. M. Bradley and D. Barrick, "Limits of cooperativity in a structurally modular protein: response of the notch ankyrin domain to analogous alanine substitutions in each repeat," *J Mol Biol*, vol. 324, pp. 373–86, Nov 2002.

[173] C. M. Bradley and D. Barrick, "Effect of multiple prolyl isomerization reactions on the stability and folding kinetics of the notch ankyrin domain: experiment and theory," *J Mol Biol*, vol. 352, pp. 253–65, Sep 2005.

[174] C. M. Bradley and D. Barrick, "The notch ankyrin domain folds via a discrete, centralized pathway," *Structure*, vol. 14, pp. 1303–12, Aug 2006.

[175] C. C. Mello and D. Barrick, "Measuring the stability of partly folded proteins using tmao," *Protein Sci.*, vol. 12, pp. 1522–1529, Jul 2003.

[176] C. C. Mello and D. Barrick, "An experimentally determined protein folding energy landscape," *Proc Natl Acad Sci USA*, vol. 101, pp. 14102–7, Sep 2004.

[177] J. D. Robertson, "Osmotic constitutents of the blood plasma and parietal muscle of squalus acanthias l," *Biol Bull*, vol. 148, pp. 303–19, Apr 1975.

[178] P. H. Yancey and G. N. Somero, "Counteraction of urea destabilization of protein structure by methylamine osmoregulatory compounds of elasmobranch fishes," *Biochem J*, vol. 183, pp. 317–23, Nov 1979.

[179] T. Y. Lin and S. N. Timasheff, "Why do some organisms use a urea-methylamine mixture as osmolyte? thermodynamic compensation of urea and trimethylamine n-oxide interactions with protein," *Biochemistry*, vol. 33, pp. 12695–701, Oct 1994.

[180] P. H. Yancey, M. E. Clark, S. C. Hand, R. D. Bowlus, and G. N. Somero, "Living with water stress: evolution of osmolyte systems," *Science*, vol. 217, pp. 1214–22, Sep 1982.

[181] A. Wang and D. W. Bolen, "A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation," *Biochemistry*, vol. 36, pp. 9101–8, Jul 1997.

[182] D. Shortle and A. K. Meeker, "Mutant forms of staphylococcal nuclease with altered patterns of guanidine hydrochloride and urea denaturation," *Proteins*, vol. 1, pp. 81–9, Sep 1986.

[183] Y. Yang, N. S. Rao, E. Walker, S. Sen, and J. Qin, "Nuclear magnetic resonance assignment and secondary structure of an ankyrin-like repeat-bearing protein: myotrophin," *Protein Sci*, vol. 6, pp. 1347–51, Jun 1997.

[184] C. C. Mello, C. M. Bradley, K. W. Tripp, and D. Barrick, "Experimental characterization of the folding kinetics of the notch ankyrin domain," *J Mol Biol*, vol. 352, pp. 266–81, Sep 2005.

[185] K. S. Tang, B. J. Guralnick, W. K. Wang, A. R. Fersht, and L. S. Itzhaki, "Stability and folding of the tumour suppressor protein p16," *Journal of Molecular Biology*, vol. 285, pp. 1869–86, Jan 1999.

[186] J. F. Brandts, H. R. Halvorson, and M. Brennan, "Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues," *Biochemistry*, vol. 14, pp. 4953–63, Nov 1975.

[187] P. A. Evans, C. M. Dobson, R. A. Kautz, G. Hatfull, and R. O. Fox, "Proline isomerism in staphylococcal nuclease characterized by nmr and site-directed mutagenesis," *Nature*, vol. 329, pp. 266–8, Jan 1987.

[188] T. O. Street, C. M. Bradley, and D. Barrick, "Predicting coupling limits from an experimentally determined energy landscape," *Proc Natl Acad Sci USA*, vol. 104, pp. 4907–12, Mar 2007.

[189] T. Aksel and D. Barrick, "Analysis of repeat-protein folding using nearest-neighbor statistical mechanical models," *Meth Enzymol*, vol. 455, pp. 95–125, Jan 2009.

[190] A. T. Iavarone and J. H. Parks, "Conformational change in unsolvated trp-cage protein probed by fluorescence," *J. Am. Chem. Soc.*, vol. 127, pp. 8606–7, Jun 2005.

[191] P. L. Muiño and P. R. Callis, "Solvent effects on the fluorescence quenching of tryptophan by amides via electron transfer. experimental and computational studies (dagger)," *The journal of physical chemistry B*, vol. 113, pp. 2572–7, Mar 2009.

[192] C. A. Royer, "Probing protein folding and conformational transitions with fluorescence," *Chem Rev*, vol. 106, pp. 1769–84, May 2006.

[193] P. R. Callis, A. Petrenko, P. L. Muiño, and J. R. Tusell, "Ab initio prediction of tryptophan fluorescence quenching by protein electric field enabled electron transfer," *The journal of physical chemistry B*, vol. 111, pp. 10335–9, Sep 2007.

[194] J. T. Vivian and P. R. Callis, "Mechanisms of tryptophan fluorescence shifts in proteins," *Biophys J*, vol. 80, pp. 2093–109, May 2001.

[195] C. A. Royer, "Improvements in the numerical analysis of thermodynamic data from biomolecular complexes," *Anal Biochem*, vol. 210, pp. 91–7, Apr 1993.

[196] T. Rosales and C. A. Royer, "A graphical user interface for bioeqs: a program for simulating and analyzing complex biomolecular interactions," *Anal Biochem*, vol. 381, pp. 270–2, Oct 2008.

[197] K. J. Frye and C. A. Royer, "The kinetic basis for the stabilization of staphylococcal nuclease by xylose," *Protein Sci*, vol. 6, pp. 789–93, Apr 1997.

[198] V. Pogenberg, J.-F. Guichou, V. Vivat-Hannah, S. Kammerer, E. Pérez, P. Germain, A. R. de Lera, H. Gronemeyer, C. A. Royer, and W. Bourguet, "Characterization of the interaction between retinoic acid receptor/retinoid x receptor (rar/rxr) heterodimers and transcriptional coactivators through structural and fluorescence anisotropy studies," *J Biol Chem*, vol. 280, pp. 1625–33, Jan 2005.

[199] A. Bourdoncle, G. Labesse, R. Margueron, A. Castet, V. Cavaillès, and C. A. Royer, "The nuclear receptor coactivator pgc-1alpha exhibits modes of interaction with the estrogen receptor distinct from those of src-1," *J Mol Biol*, vol. 347, pp. 921–34, Apr 2005.

[200] G. J. Vidugiris, D. M. Truckses, J. L. Markley, and C. A. Royer, "High-pressure denaturation of staphylococcal nuclease proline-to-glycine substitution mutants," *Biochemistry*, vol. 35, pp. 3857–64, Mar 1996.

[201] S. J. Maleki, C. A. Royer, and B. K. Hurlburt, "Analysis of the dna-binding properties of myod, myogenin, and e12 by fluorescence anisotropy," *Biochemistry*, vol. 41, pp. 10888–94, Sep 2002.

[202] A. Zipp and W. KAUZMANN, "Pressure denaturation of metmyoglobin," *Biochemistry*, vol. 12, pp. 4217–28, Oct 1973.

[203] W. Dzwolak, M. Kato, and Y. Taniguchi, "Fourier transform infrared spectroscopy in high-pressure studies on proteins," *Biochim Biophys Acta*, vol. 1595, pp. 131–44, Mar 2002.

[204] N. B. From and B. E. Bowler, "Urea denaturation of staphylococcal nuclease monitored by fourier transform infrared spectroscopy ‚Ä†," *Biochemistry*, vol. 37, pp. 1623–1631, Feb 1998.

[205] N. G. Sgourakis, R. Day, S. A. McCallum, and A. E. Garcia, "Pressure effects on the ensemble dynamics of ubiquitin inspected with molecular dynamics simulations and isotropic reorientational eigenmode dynamics," *Biophys J*, vol. 95, pp. 3943–55, Oct 2008.

[206] B. H. G. G. A. E. M. H. J. C. B. David Van Der Spoel, Erik Lindahl, "Gromacs: Fast, flexible, and free," *Journal of Computational Chemistry*, vol. 26, pp. 1701–1718, Jan 2005.

[207] G. A. Kaminski, , R. A. Friesner, J. Tirado-Rives, , and W. L. Jorgensen, "Evaluation and reparametrization of the opls-aa force field for proteins via comparison with accurate quantum chemical calculations on peptides,Ä†,"

[208] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, "The missing term in effective pair potentials,"

[209] N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M.-Y. Shen, U. Pieper, and A. Sali, "Comparative protein structure modeling using modeller," *Curr Protoc Bioinformatics*, vol. Chapter 5, p. Unit 5.6, Oct 2006.

[210] W. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Phys Rev, A*, vol. 31, pp. 1695–1697, Mar 1985.

[211] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *J. Appl. Phys.*, vol. 52, no. 12, pp. 7182–7190, 1981.

[212] T. Darden, D. York, and L. Pedersen, "Particle mesh ewald : an n.log(n) method for ewald sums in large systems," *The Journal of chemical physics*, vol. 98, pp. 10089–10092, Jan 1993.

[213] G. Pappenberger, C. Saudan, M. Becker, A. E. Merbach, and T. Kiefhaber, "Denaturant-induced movement of the transition state of protein folding revealed by high-pressure stopped-flow measurements," *Proc Natl Acad Sci USA*, vol. 97, pp. 17–22, Jan 2000.

[214] C. Krywka, C. Sternemann, M. Paulus, M. Tolan, C. Royer, and R. Winter, "Effect of osmolytes on pressure-induced unfolding of proteins: a high-pressure saxs study," *Chemphyschem : a European journal of chemical physics and physical chemistry*, vol. 9, pp. 2809–15, Dec 2008.

[215] S. Paul and G. N. Patey, "Hydrophobic interactions in urea-trimethylamine-n-oxide solutions," *The journal of physical chemistry B*, vol. 112, pp. 11106–11, Sep 2008.

[216] T. O. Street, N. Courtemanche, and D. Barrick, "Protein folding and stability using denaturants," *Methods Cell Biol*, vol. 84, pp. 295–325, Jan 2008.

[217] V. K. Dubey and M. V. Jagannadham, "Differences in the unfolding of procerain induced by ph, guanidine hydrochloride, urea, and temperature," *Biochemistry*, vol. 42, pp. 12287–97, Oct 2003.

[218] M. E. Zweifel and D. Barrick, "Relationships between the temperature dependence of solvent denaturation and the denaturant dependence of protein stability curves," *Biophys Chem*, vol. 101-102, pp. 221–37, Dec 2002.

[219] A. N. Naganathan and V. Munoz, "Insights into protein folding mechanisms from large scale analysis of mutational effects," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 8611–8616, May 2010.

[220] H. Herberhold and R. Winter, "Temperature- and pressure-induced unfolding and refolding of ubiquitin: a static and kinetic fourier transform infrared spectroscopy study.," *Biochemistry*, vol. 41, no. 7, pp. 2396–401, 2002.

[221] M. W. Lassalle, H. Yamada, and K. Akasaka, "The pressure-temperature free energy-landscape of staphylococcal nuclease monitored by (1)h nmr," *Journal of Molecular Biology*, vol. 298, pp. 293–302, Apr 2000.

[222] J. Wiedersich, S. Köhler, A. Skerra, and J. Friedrich, "Temperature and pressure dependence of protein stability: the engineered fluorescein-binding lipocalin flua shows an elliptic phase diagram," *Proc Natl Acad Sci USA*, vol. 105, pp. 5756–61, Apr 2008.

[223] D. Reinstädler, H. Fabian, J. Backmann, and D. Naumann, "Refolding of thermally and urea-denatured ribonuclease a monitored by time-resolved ftir spectroscopy," *Biochemistry*, vol. 35, pp. 15822–30, Dec 1996.

[224] J. L. Silva and G. Weber, "Pressure stability of proteins," *Annu Rev Phys Chem*, vol. 44, pp. 89–113, Jan 1993.

[225] G. Weber and H. G. Drickamer, "The effect of high pressure upon proteins and other biomolecules," *Q Rev Biophys*, vol. 16, pp. 89–112, Feb 1983.

[226] H. Imamura and M. Kato, "Effect of pressure on helix-coil transition of an alanine-based peptide: an ftir study," *Proteins*, vol. 75, pp. 911–8, Jun 2009.

[227] D. Paschek, S. Gnanakaran, and A. E. Garcia, "Simulations of the pressure and temperature unfolding of an alpha-helical peptide," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 6765–70, May 2005.

[228] S. Marqusee, V. H. Robbins, and R. L. Baldwin, "Unusually stable helix formation in short alanine-based peptides," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, pp. 5286–90, Jul 1989.

[229] R. Mohana-Borges, J. L. Silva, J. Ruiz-Sanz, and G. de Prat-Gay, "Folding of a pressure-denatured model protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, pp. 7888–93, Jul 1999.

[230] P. Venkatesu, M.-J. Lee, and H. mu Lin, "Trimethylamine n-oxide counteracts the denaturing effects of urea or gdnhcl on protein denatured state," *Arch Biochem Biophys*, vol. 466, pp. 106–15, Oct 2007.

[231] M. S. Till and G. M. Ullmann, "Mcvol - a program for calculating protein volumes and identifying cavities by a monte carlo algorithm," *J Mol Model*, vol. 16, pp. 419–29, Mar 2010.

[232] T. V. Chalikian and R. B. Macgregor, "Origins of pressure-induced protein transitions," *Journal of Molecular Biology*, vol. 394, pp. 834–42, Dec 2009.

[233] T. Takekiyo, N. Takeda, Y. Isogai, M. Kato, and Y. Taniguchi, "Pressure stability of the alpha-helix structure in a de novo designed protein (alpha-l-alpha)(2) studied by ftir spectroscopy," *Biopolymers*, vol. 85, pp. 185–8, Feb 2007.

[234] M. Schätzle and T. Kiefhaber, "Shape of the free energy barriers for protein folding probed by multiple perturbation analysis," *Journal of Molecular Biology*, vol. 357, pp. 655–64, Mar 2006.

[235] R. L. Baldwin, C. Frieden, and G. D. Rose, "Dry molten globule intermediates and the mechanism of protein unfolding," *Proteins*, vol. 78, pp. 2725–37, Oct 2010.

[236] A. D. Tsamaloukas, N. K. Pyzocha, and G. I. Makhatadze, "Pressure perturbation calorimetry of unfolded proteins," *Journal of physical chemistry*.

STUDY OF THE PHYSICAL BASIS OF PRESSURE EFFECTS ON PROTEINS USING MODEL ANKYRIN REPEAT CONSTRUCTS

Le dépliement thermique et chimique est raisonnablement compris, mais pas les effets déstabilisants de la pression. Dans une tentative de caractérisation des facteurs à la base des effets de la pression sur les protéines, nous avons étudié le dépliement sous pression d'une protéine modulaire, le domaine ankyrine du récepteur Notch (Nank1-7*), ainsi que plusieurs mutants en nombre de motifs. Nos expériences montrent un dépliement à deux états sous pression. La dépendance à la température des sauts de pression a montré qu'à faibles températures l'ensemble d'état de transition (TSE) est proche en volume de l'état replié alors que son énergie est proche de celle de l'état déplié, ce qui est significatif d'une déshydratation importante de la barrière énergétique, et a montré que l'augmentation de la température conduit à un TSE plus grand en volume que l'état replié. Ce comportement révèle une grande plasticité du TSE de Nank1-7*. L'étude des mutants (délétion de motifs) nous a montré que le ΔV n'est déterminé ni par l'hydratation des liaisons peptidiques ni par l'hydratation différentielle des acides aminés, mais par l'existence de vides internes exclus du solvant. Seule une petite fraction des cavités est déterminante pour le ΔV, celles qui ne sont pas significativement solvatées dans les cœurs hydrophobes. Finalement, nous avons déterminé que le TSE possède même des cavités plus grandes, et nous émettons l'hypothèse que l'énergétique et la dynamique contribuent aux effets de la pression sur les protéines.

Thermal and chemical unfolding of proteins are reasonably well understood, but the destabilizing effects of pressure are not. In an attempt to characterize the factors at the basis of pressure effects, we investigated the pressure unfolding of a modular protein, the ankyrin domain of the Notch receptor (Nank1–7*) and several of its deletion constructs. All our experiments were consistent with a simple two-state folding/unfolding transition under pressure. The temperature dependence of pressure-jumps demonstrated that at low temperature, the transition state ensemble (TSE) lies close in volume to the folded state despite its unfolded-like energetics, consistent with significant dehydration at the barrier, and that increasing temperature leads to a volume of the TSE larger than that of the folded state. This behavior reveals a high degree of plasticity of the TSE of Nank1–7*. Studies of the deletion mutants showed us that ΔV is determined not by hydration of peptide bonds or by differential hydration of residues, but by the existence of internal solvent-excluded void volumes. Only a small fraction of the internal cavities are relevant of the ΔV, those that are not significantly solvated in the hydrophobic core. Finally, we determined that the transition state ensemble show even larger cavities, and we hypothesize both energetics and dynamics contribute to pressure effects on proteins.

DISCIPLINE : Biophysique

MOTS-CLES : Repliement, Volume, Pression, Ankyrine, Biophysique, Fluorescence

KEYWORDS : Folding, Volume, Pressure, Ankyrin, Biophysics, Fluorescence

LABORATOIRE :
Centre biochimie structurale
CNRS UMR 5048 – UM 1 – UM 2 – INSERM UMR 554
29, rue de Navacelles
34090 MONTPELLIER Cedex - France