Molecular dynamics simulations of the native and partially folded states of ubiquitin: Influence of methanol cosolvent, pH, and temperature on the protein structure and dynamics

DAVID B. KONY, PHILIPPE H. HÜNENBERGER, AND WILFRED F. VAN GUNSTEREN

Laboratory of Physical Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland

(RECEIVED June 7, 2006; FINAL REVISION February 1, 2007; ACCEPTED February 7, 2007)

Abstract

A series of explicit-solvent molecular dynamics simulations of the protein ubiquitin are reported, which investigate the effect of environmental factors (presence of methanol cosolvent in the aqueous solution, neutral or low pH value, room or elevated temperature) on the structure, stability, and dynamics of the protein. The simulations are initiated either from the native structure of the protein or from a model of a partially folded state (A-state) that is known to exist at low pH in methanol–water mixtures. The main results of the simulations are: (1) The ubiquitin native structure is remarkably stable at neutral pH in water; (2) the addition of the methanol cosolvent enhances the stability of the secondary structure but weakens tertiary interactions within the protein; (3) this influence of methanol on the protein structure is enhanced at low pH, while the effect of lowering the pH in pure water is limited; and (4) the A-state of ubiquitin can be described as a set of relatively rigid secondary structure elements (a native-like β -sheet and native-like α -helix plus two nonnative α -helices) connected by flexible linkers.

Keywords: molecular dynamics; protein folding; partially folded state; ubiquitin; water–methanol mixture

The factors that determine the conformation adopted by a particular peptide segment in the folded structure of a protein (including intrinsic conformational preferences of the peptide sequence, hydrogen bonding, hydrophobic effect, and packing interactions) are so complex and interdependent that they cannot be elucidated by simple inspection of the native structure alone. Factors promoting protein denaturation (including changes in temperature, pressure, ionic strength, pH, solvent, and concentration of cosolvent) may provide valuable information on protein stability because they alter the balance between the above factors, thereby providing a clue about their relative magnitudes. Under specific environmental

conditions, proteins may present partially folded states, corresponding to structures (or, rather, structural ensembles) intermediate between those characteristic of the native and unfolded states (Ptitsyn 1987, 1995; Kuwajima 1989; Bychkova and Ptitsyn 1993; Arai and Kuwajima 2000; Daggett and Fersht 2003; Sanchez and Kiefhaber 2003; Refield 2004; Dyson and Wright 2005; Jahn and Radford 2005; Wolynes 2005; Cremades et al. 2006). In these partially folded states, tertiary interactions and hydrophobic core size are often reduced, permitting an easier investigation of the local interactions responsible for the formation of specific secondary-structure elements. The investigation of partially folded states is also expected to provide information about the folding pathways of proteins (Dastidar and Mukhopadhyay 2005; Dyson and Wright 1998). This information is extremely valuable because direct probing of transient states along folding pathways using high-resolution structural tools, such as nuclear magnetic resonance (NMR) spectroscopy

Reprint requests to: Wilfred F. van Gunsteren, Laboratory of Physical Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland; e-mail: wfvgn@igc.phys.chem.ethz.ch; fax: 41 1 632 1039.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.062323407>.

or X-ray crystallography, is generally not possible because of the short lifetimes of these states (Kim and Baldwin 1982).

Among partially folded states, particular attention has been paid to the so-called molten globule state of several small proteins, which represents an ensemble of conformations at equilibrium under mildly denaturing conditions (Ptitsyn 1987, 1995; Kuwajima 1989; Bychkova and Ptitsyn 1993; Arai and Kuwajima 2000; Daggett and Fersht 2003; Sanchez and Kiefhaber 2003; Refield 2004; Dyson and Wright 2005; Jahn and Radford 2005; Wolynes 2005; Cremades et al. 2006) and is amenable to detailed experimental studies. Molten globules are typically found at high or low pH values, as well as at high concentrations of alcohols, urea, or guanidinium chloride. This state can be thought of as an ensemble of compact conformations involving both native and non-native secondary structure, and presenting an increased flexibility (compared to the native state) as well as a lack of both specific tertiary interactions and definite side-chain conformations (Ptitsyn and Semisotnov 1991).

The present study is concerned with the partially folded state of a specific protein: ubiquitin. Ubiquitin is a globular 76-residue cytoplasmic protein with a molecular

weight of 8.6 kDa, exempt of prosthetic groups and disulfide bridges (Goldstein et al. 1975). It is found in all eukaryotic cells and is involved in a variety of biochemical functions, including protein degradation and differential gene regulation (Chin et al. 1982; Hershko et al. 1984; Rechsteiner 1988). The native state of ubiquitin has been characterized structurally by X-ray crystallography (Vijay-Kumar et al. 1987) and NMR spectroscopy (Cornilescu et al. 1998) (using NOE intensities, 3J -coupling constants, and dipolar couplings), and dynamically by NMR relaxation measurements (Schneider et al. 1992; Tjandra et al. 1995; Wand et al. 1996). The protein native state involves a high content of secondary-structure elements, comprising a five-stranded β -sheet, an α -helix (fitting into a concavity formed by the β -sheet), a 3_{10} -helix, and seven reverse turns (Fig. 1). Under physiological conditions, ubiquitin is remarkably stable with an unusually rigid backbone structure (as evidenced by NMR relaxation measurements) (Schneider et al. 1992; Tjandra et al. 1995; Wand et al. 1996). In water and at pH 5.8, the protein remains in its native state (N-state) up to 355 K, temperature above which it unfolds and aggregates (precipitation). However, at low pH and in both water and methanol–water mixtures, the protein can

Figure 1. Ribbon diagrams of ubiquitin highlighting the secondary-structure content of the native structure (Vijay-Kumar et al. 1987), the modeled A-state, and the final structures corresponding to selected simulations (see Table 1 and Fig. 2). The structure assumed representative of the A-state (built based on the schematic model proposed by Brutscher et al. 1997; see Fig. 1 therein) and the corresponding structure after 3 ns of MD simulation should be taken with some caution. Owing to the limited amount of experimental data involved in the building of the model (underdetermination), the enhanced segmental mobility in this state (compared to the native one) and the short timescale of the simulation (3 ns), these structures probably represent illustrative conformations in the corresponding conformational ensemble rather than accurate structural models.

Table 1. Systems and conditions corresponding to the 10 MD simulations of ubiquitin

	Neutral pH		Low pH		
	Water	Mixture	Water	Mixture	Mixture
Code	NW	NM	LW	LM	A
Initial structure	Native	Native	Native	Native	Modeled A-state
Protein	761	761	773	773	773
Methanol	0	6048	0	6048	32,562
Water	21,996	9072	22,665	9072	48,804
Ion (Cl ⁻)	1	1	13	13	13
Initial box (nm ³)	232	237	237	237	1203
Time (ns)	10	10	10	10	3

In each case, the simulations were performed at either 300 or 360 K and 1 bar. The simulation codes, associated with the simulation temperature, are used to refer to the different simulations throughout the article: (N) neutral pH; (L) low pH; (W) pure water; (M) 60% methanol–water mixture; (A) simulation starting from the modeled A-state structure (low pH, methanol–water mixture only). Rows 3–6 report the corresponding number of atoms in the system. Row 7 indicates the initial volume of the (truncated-octahedron) computational box. Row 8 indicates the total simulation time excluding equilibration and continuation runs (for LM360) (see Fig. 2).

be reversibly denatured by varying the temperature in a limited range. For example, at a pH lower than 2 and in water, ubiquitin remains soluble up to 368 K with a folding–unfolding transition centered at 350 K (Harding et al. 1991).

Titration of methanol into an aqueous sample of ubiquitin at 298 K and pH 2 results in the gradual appearance of a partially folded state, known as the A-state of the protein. This state becomes detectable above a 30% methanol–water volume ratio (i.e., the solution resulting from mixing three volumes of methanol with seven volumes of water), and is the only populated state above a 60% ratio (Harding et al. 1991). The study of the A-state of ubiquitin has been the scope of several circular dichroism (CD) and NMR experiments over the last two decades (Wilkinson and Mayer 1986; Harding et al. 1991; Pan and Briggs 1992; Stockman et al. 1993; Brutscher et al. 1997). However, the experimental characterization of this partially folded state by NMR is difficult (compared to that of the native state) because of the reduced range of ¹H chemical shifts (only the most recent multi-dimensional heteronuclear experiments could achieve a complete assignment) (Stockman et al. 1993; Brutscher et al. 1997), as well as the sparseness of medium-range and the absence of long-range nuclear Overhauser (NOE) cross-peaks, caused by an increased backbone flexibility and the lack of tertiary interactions. As a consequence, the results of these investigations remain largely qualitative and are sometimes contradictory. The general picture of the A-state emerging from these studies is that of a highly flexible structure, with a slightly more rigid and structured N-terminal half (residues 1–35) compared to the C-terminal half (residues 36–76). All available experimental data (NOEs, chemical shifts, NMR relaxation, and hydrogen exchange) support the presence of native-like secondary structure (β-hairpin and α-helix) in

the N-terminal half. Furthermore, the comparison of NOE cross-peaks for an isolated polypeptide corresponding to residues 1–21 of ubiquitin and for the corresponding segment in the A-state of the protein has shown that the N-terminal β-hairpin may fold independently of any interaction with the rest of the protein (Cox et al. 1993). The situation is less clear for the C-terminal half, where different studies reach diverging conclusions. On the one hand, hydrogen-exchange experiments (Harding et al. 1991; Pan and Briggs 1992) suggest that the C-terminal half remains to a large extent native-like (including the presence of tertiary interaction between the β-sheet and the α-helix). On the other hand, sequential NOEs, chemical shifts, and NMR relaxation parameters

Figure 2. Durations and temperatures of the continuation runs branched from simulation LM360 (see Table 1). The code LM indicates simulations performed at low pH and in a 60% methanol–water mixture. In a first set of continuation runs (subscript 1), three simulations (LM₁300, LM₁360, and LM₁400) are performed during 5 ns using the final configuration of LM360 as a starting structure, and temperatures of 300 K, 360 K, and 400 K, respectively. In a subsequent set (subscript 2), two simulations [LM₂300(360) and LM₂300(400)] are performed during 5 ns using the final configurations of LM₁360 and LM₁400, respectively, as starting structures, and a temperature of 300 K.

(Stockman et al. 1993; Brutscher et al. 1997), as well as CD spectra (Cox et al. 1993) seem to support the absence of β -sheet, the predominance of helical character, and a complete lack of tertiary interactions. The latter conclusion suggests a weaker encoding of the secondary structure by the sequence in the C-terminal half (compared to the N-terminal half), where tertiary interactions would be a dominant driving force in promoting the native (β -sheet) structure. As noted by Brutscher et al. (1997), hydrogen-exchange experiments may be less suited (compared to NOE or chemical shift data) to discriminate between different structural models because (1) they involve delicate experimental procedures (refolding experiments) that may perturb the protein structure; and (2) protection factors represent ambiguous structural indicators (the relationship between hydrogen-exchange rates and protein structure may be altered by other factors). The model emerging from Brutscher's work (Brutscher et al. 1997) for the A-state of ubiquitin (in apparent contradiction with hydrogen-exchange data) (Harding et al. 1991; Pan and Briggs 1992) consists of three distinct secondary-structure elements: a β -hairpin and two helices (Fig. 1), forming rather independent motional units.

Alcohols are thought to promote the (partial) denaturation of proteins by disrupting tertiary interactions, presumably through a weakening of the hydrophobic effect relative to hydrogen-bonding interactions, so that conformational propensities tend to be dominated by local (principally hydrogen-bonding) interactions within the polypeptide chain. In water, a key property of the native folded structure adopted by many proteins is that it generates a solvent-excluded environment, in which the stability of hydrogen-bonded structures (i.e., secondary-structure elements) is enhanced compared to the aqueous environment. The addition of methanol to aqueous solutions of polypeptides appears to mimic this effect in a nonspecific manner, promoting the formation of secondary structure without imposing the additional structural constraints resulting from tertiary packing. Experimental evidence, mainly from CD studies, suggests that in the absence of a strong encoding of a secondary-structure element by the peptide sequence, this process predominantly results in helix formation (Brown and Klee 1971; Harding et al. 1991; Buck et al. 1993; Alexandrescu et al. 1994; Shiraki et al. 1995; Hirota et al. 1997). However, these considerations remain largely qualitative, and there is still considerable uncertainty about the mechanisms through which methanol modulates protein structure and stability.

Further insight into the effect of methanol may be provided by NMR (Jourdan and Searle 2001) and calorimetric (Woolfson et al. 1993) measurements characterizing the thermodynamics of the unfolded-to-native ($U \rightarrow N$) and A-state-to-native ($A \rightarrow N$) transitions of ubiquitin in

the presence of the cosolvent. In pure water, the folding of a protein presenting a hydrophobic core in its native state is typically moderately exothermic and affected by a limited adverse (negative) entropy change. This process is also generally accompanied by a large decrease in the heat capacity. In the case of ubiquitin at 298 K and pH 2 in water, the corresponding enthalpy, entropy, and heat capacity changes are $\Delta H_{U \rightarrow N} = -10$ kJ/mol, $\Delta S_{U \rightarrow N} = -34$ J/(K · mol), and $\Delta C_{p,U \rightarrow N} = -4880$ J/(K · mol), respectively (Woolfson et al. 1993). Note that in the presence of denaturant (3.9 M guanidinium chloride), the corresponding values are slightly different, namely, $\Delta H_{U \rightarrow N} = +9.8$ kJ/mol, $\Delta S_{U \rightarrow N} = +31.3$ J/(K · mol), and $\Delta C_{p,U \rightarrow N} = -5180$ J/(K · mol), respectively (Jourdan and Searle 2001) (under these conditions, ubiquitin undergoes cold denaturation below 298 K). The main contribution to the heat-capacity change is probably related to the solvent-ordering effect associated with the exposure of hydrophobic groups in the unfolded state (Woolfson et al. 1993). In methanol–water mixtures, the thermodynamics of the folding processes is very different. In this case, folding is accompanied by much larger negative enthalpy and entropy changes upon folding, together with a negligible change in heat capacity. For ubiquitin at 298 K and pH 4 in a 40% methanol–water mixture, the corresponding changes are $\Delta H_{U \rightarrow N} = -198$ kJ/mol, $\Delta S_{U \rightarrow N} = -660$ J/(K · mol), and $\Delta C_{p,U \rightarrow N} \approx 0$ J/(K · mol), respectively (Woolfson et al. 1993) [in the presence of 2.2 M guanidinium chloride, the enthalpy and entropy changes become $\Delta H_{U \rightarrow N} = -145$ kJ/mol and $\Delta S_{U \rightarrow N} = -487$ J/(K · mol), respectively (Jourdan and Searle 2001); $\Delta C_{p,U \rightarrow N}$ is not accurately known in this case]. The same considerations apply to the $A \rightarrow N$ transition under similar experimental conditions. At 298 K and pH 2 in a 40% methanol–water mixture, the transition (in the presence of 2.2 M guanidinium chloride) is characterized by $\Delta H_{A \rightarrow N} = -216$ kJ/mol, $\Delta S_{A \rightarrow N} = -720$ J/(K · mol), and $\Delta C_{p,A \rightarrow N} \approx 0$ J/(K · mol), respectively (Jourdan and Searle 2001) (the corresponding values in the absence of denaturant have not been measured). The absence of a heat capacity change suggests that the thermodynamics of the $U \rightarrow N$ or $A \rightarrow N$ folding transitions is no longer critically dependent on the solvent-ordering effect associated with the exposure of hydrophobic residues in the unfolded state (Woolfson et al. 1993). Furthermore, because the unfolded and A-states are rather similar regarding the absence of hydrophobic packing but very dissimilar in terms of intramolecular hydrogen bonding, the qualitative similarity between the thermodynamics of the $U \rightarrow N$ and $A \rightarrow N$ transitions suggests that the negative enthalpy change is probably caused by changes in nonpolar (van der Waals and hydrophobic) rather than polar (hydrogen-bonding) interactions (Jourdan and Searle 2001).

A change in the environmental pH affects the stability of a protein by altering the interaction among its ionizable residues, as well as their solvation properties. Ubiquitin possesses 11 acidic residues (five Asp and six Glu residues) and 10 basic residues (seven Lys, four Arg, and one His residues), in addition to the N and C termini. Lowering the pH to 2 will cause the protonation of all acidic residues into their neutral form. This is in particular likely to reduce the strength of the salt bridges Lys11/Glu34, Lys27/Glu51, and Lys27/Asp52 anchoring the α -helix to the five-stranded β -sheet, thereby destabilizing the native state.

To gain insight into the effect of pH and methanol cosolvent on the properties of ubiquitin, theoretical methods have also been applied to this system. Two explicit-solvent molecular dynamics (MD) simulations of ubiquitin in methanol–water mixtures have been reported to date. In the first study (Alonso and Daggett 1995), a series of four 0.5-ns MD simulations was performed at 335 K in pure water or in a 60% methanol–water mixture, starting from two different initial structures for the N-state. In the presence of methanol, the formation of nonnative secondary structure in the C-terminal half was observed (but no increase in helicity), while the N-terminal half of the protein remained native-like. The absence of structural features characteristic of the A-state in the C-terminal half almost certainly results from the very short lengths of these simulations (i.e., insufficient to convert the N-state into the A-state). The second study involved a 33-ns MD simulation of ubiquitin using a variable-temperature scheme to control and speed up the unfolding process under conditions experimentally known to promote the stability of the A-state (Prompers et al. 2001). Despite the considerably longer timescale, the conformations sampled toward the end of this simulation, although more expanded, were similar to those observed in the previous study. However, the simulation length was sufficient to permit a characterization of the protein motion in terms of reorientational quasi-harmonic modes (Prompers and Brueschweiler 2000) as an alternative to the usual interpretation of spin relaxation based on the model-free approach (Clore et al. 1990). Although the simulation did not reach the A-state, the dynamics of the protein (characterized by unusually low-order parameters as compared to typical relaxation results on globular proteins) was in qualitative agreement with experimental data (Brutscher et al. 1997).

In the present study, we report a series of long-timescale explicit-solvent MD simulations of ubiquitin under different conditions (Table 1 and Fig. 2). These new simulations probe the structure, dynamics, and solvation of both the native and partially folded states of ubiquitin, considering separately the effects of pH, of methanol cosolvent, and of the temperature. All simulations were

performed using the GROMOS96 program (van Gunsteren et al. 1996; Scott et al. 1999) together with the recently developed 45A4 force field (Schuler et al. 2001; Soares et al. 2005; Lins and Hünenberger 2005), successor of the GROMOS96 43A1 parameter set (van Gunsteren et al. 1996; Daura et al. 1998; Scott et al. 1999). This new parameter set, reoptimized to reproduce the heats of vaporization and densities of a series of alkanes as well as their hydration free energies, has been subjected to limited validation up to now (Daura et al. 1998; Soares et al. 2004, 2005). Therefore, as a secondary goal, the simulation starting from the folded structure and performed under native conditions also serves the purpose of validating the new parameter set by comparison against high-quality experimental data.

Results and Discussion

Simulations initiated from the native structure

The time series of the C_α root-mean-square atomic positional deviation from the crystallographic structure (RMSD) are displayed in Figure 3 for the simulations initiated from the native structure of ubiquitin. The corresponding averages (beginning and end of the simulations) are reported in Table 2 together with corresponding averages for the total number of intramolecular hydrogen bonds (backbone only) and the radius of gyration (rgyr). The final structures of the high-temperature simulations (NW360, NM360, LW360, and LM360), as well as of simulations NW300 and LM₁300 are also represented in

Figure 3. C_α root-mean-square atomic positional deviation (RMSD) from the crystallographic structure (Vijay-Kumar et al. 1987), displayed as a function of time for the different ubiquitin simulations listed in Table 1 (omitting simulations A300 and A360; four left panels) and Figure 2 (continuations of LM360; right panel).

Table 2. Average structural properties calculated for the X-ray structure (Vijay-Kumar et al. 1987), the set of 10 NMR model structures (Cornilescu et al. 1998), and for the different simulations

	$\langle \text{RMSD} \rangle_{\text{beg}}$ (nm)	$\langle \text{RMSD} \rangle_{\text{end}}$ (nm)	$\langle \text{hb} \rangle_{\text{beg}}$	$\langle \text{hb} \rangle_{\text{end}}$	$\langle \text{rgyr} \rangle_{\text{beg}}$ (nm)	$\langle \text{rgyr} \rangle_{\text{end}}$ (nm)
X-ray	0		39		1.19	
NMR	0.04		40		1.26	
NW300	0.12	0.14	34	30	1.20	1.18
NM300	0.11	0.15	33	32	1.21	1.23
LW300	0.10	0.12	34	34	1.19	1.19
LM300	0.10	0.17	33	32	1.22	1.22
NW360	0.14	0.19	30	29	1.21	1.18
NM360	0.19	0.30	26	24	1.25	1.23
LW360	0.14	0.64	30	20	1.22	1.20
LM360	0.16	0.33	29	23	1.24	1.28
LM ₁ 300	0.34	0.33	23	25	1.29	1.29
LM ₁ 360	0.49	0.83	19	19	1.35	1.54
LM ₁ 400	0.63	1.13	15	13	1.39	1.62
LM ₂ 300(360)	0.71	0.76	18	17	1.42	1.45
LM ₂ 300(400)	1.29	1.26	17	19	1.68	1.66
A300	1.03	1.27	34	31	3.45	3.47
A360	1.10	2.05	21	16	3.34	4.48

See Table 1 and Figure 2. The C_{α} root-mean-square atomic positional deviation from the X-ray structure (RMSD), the total number of backbone–backbone hydrogen bonds (hb), and the average radius of gyration (rgyr) are reported. Averages are performed over either the first (beg) or the last (end) 2 ns of the trajectories.

Figure 1. The discussion first focuses on the eight 10-ns simulations initiated from the protein native structure (continuation runs and A-state simulations being discussed subsequently).

For the simulations performed at 300 K (except simulation LM300), the RMSD remains low, with an average value between 0.12 and 0.15 nm. The value for simulation LM300 is somewhat higher, with a sharp increase around 4 ns (to a maximum of ~ 0.3 nm) followed by a stabilization at a plateau value of ~ 0.17 nm. The RMSD value for simulation NW360 remains very close to the one for NW300, an indication of the particularly high stability of the ubiquitin native structure under native conditions (Schneider et al. 1992; Tjandra et al. 1995; Wand et al. 1996). In contrast, for all other simulations performed at 360 K, the RMSD value increases continuously along the trajectory. The average RMSD values over the last 2 ns of the simulations LW360, LM360, and NM360 are 0.64, 0.33, and 0.30 nm, respectively. This suggests that (1) a change of pH (from neutral to low pH) in water leads to a more significant destabilization of the ubiquitin native structure compared to a change in solvent (from water to methanol–water mixture) at neutral pH; and (2) the change of pH has a much more limited effect in the methanol–water mixture than in water. The large increase in the RMSD value along the LW360 simulation occurs in three steps: A first plateau value of ~ 0.15 – 0.25 nm is maintained during the first 6 ns, followed by a second plateau value of ~ 0.45 nm between 6 and 7 ns, and by a continuous increase up to a value of 0.7 nm by the end of the 10-ns

simulation. The final structure (Fig. 1) is largely distorted, probably in part owing to the lack of the salt bridges anchoring the α -helix to the β -sheet under low pH conditions. The protein region that corresponds to the α -helix in the native structure presents an orientation approximately parallel to the β -hairpin, orientation that largely differs from that observed in the native structure and in the structures reached at the end of all other simulations (Fig. 1). Note also that, in spite of similar RMSD values at the end of the simulation, the simulation LM360 leads to a considerably more distorted structure (in particular, in terms of secondary structure elements) compared to the NM360 simulation. The observation that the simultaneous change of pH and solvent (NW360 $\rightarrow$ LM360) appears to have less impact than the change of pH alone in water (NW360 $\rightarrow$ LW360) suggests that the influence of pH and methanol cosolvent on the protein structure are not simply additive, and that a more complex interplay is involved between these two environmental factors.

The trends observed in the RMSD values are in good qualitative agreement with those observed in the average number of intramolecular (backbone only) hydrogen bonds (hb) (Table 2). Here, a higher RMSD value is generally correlated with a lower number of hydrogen bonds and conversely. The average numbers of hydrogen bonds (calculated over the last 2 ns of the simulations) are between 30 and 34 for the simulations performed at 300 K. For comparison, this number is 39 for the X-ray structure and 40 for the set of NMR model structures. Consistent with the low RMSD value observed in simulation NW360, the final part of the simulation is characterized

by a high number of hydrogen bonds (29). Similarly, the very large RMSD value reached at the end of simulation LW360 coincides with the lowest number of hydrogen bonds in the final part of the simulation (20). The final parts of the two other simulations at 360 K (LM360 and NM360) are characterized by intermediate numbers of hydrogen bonds (23–24).

In contrast to the RMSD and number of intramolecular hydrogen bonds, the radius of gyration (*rgyr*) is rather insensitive to the simulation conditions (Table 2). Irrespective of the system and temperature (even in the simulation LW360), the values for all simulations started from the native state of ubiquitin remain in the range 1.18–1.28 nm for the final parts of the simulations, similar to those corresponding to the X-ray (1.19 nm) and NMR model (1.26 nm) structures. This suggests that the size of the protein hydrophobic core is relatively well preserved during the 10-ns simulations of Table 1 (see also Fig. 1). However, the simulations performed in the methanol–water mixture are characterized by slightly larger values for the radius of gyration. For example, the value reached toward the end of simulation LM360 (1.28 nm) is significantly larger than the one observed for the simulation LW360 (1.20 nm), in spite of a much higher RMSD value at the end of the latter simulation. This observation is consistent with the ability of the methanol cosolvent to mitigate hydrophobic interactions (Buck et al. 1993; Alexandrescu et al. 1994; Shiraki et al. 1995; Hirota et al. 1997), thereby allowing for the deployment of the protein structure.

The results for the continuation runs branched from simulation LM360 (Fig. 2) are also displayed in Figures 1 and 3, the corresponding averages of structural properties being reported in Table 2. For these simulations, the RMSD value steadily increases during the simulation segments performed at high temperatures (360 K and, to a larger extent, 400 K), increased thermal motion leading to the expected enhancement of the conformational search. Simultaneously, the number of intramolecular hydrogen bonds decreases, and the radius of gyration increases. During the simulation segments performed at 300 K, the RMSD value levels off (i.e., fluctuates around its initial value), while the number of hydrogen bonds and radius of gyration remain essentially unaltered, indicating a failure of the attempt to quench the system toward the A-state structure. In particular, these simulations are characterized by a progressive loss of the secondary and tertiary structure content (during the high-temperature segments), without the expected trend of β -sheet to α -helix conversion. However, it is interesting to note that the C-terminal β -hairpin of the protein is preserved during all simulations, in agreement with the high stability of this secondary-structure element (even in the absence of tertiary interactions) suggested by several independent

experiments (Wilkinson and Mayer 1986; Harding et al. 1991; Pan and Briggs 1992; Cox et al. 1993; Stockman et al. 1993; Brutscher et al. 1997; Cordier and Grzesiek 2002). Besides these observations, the continuation runs failed to achieve a significant extent of conversion to the A-state and will not be considered further here.

The time evolution of the secondary-structure content is displayed in Figure 4 together with the C_{α} root-mean-square atomic positional fluctuations (RMSF) as a function of the residue sequence number, for most of the simulations initiated from the ubiquitin native structure (Table 1; not shown for NM300 and LW300). The native secondary structure is well preserved along both simulations NW300 and LM300. However, it appears that the presence of methanol at low pH suppresses most of the transient disruptions of the secondary structure elements that occur in water at neutral pH. Inspection of the corresponding graphs for simulations NM300 and LW300 (data not shown), which show similar stability of the secondary structure, suggest that the presence of methanol (rather than the lower pH) is the key factor for this suppression. Thus, in spite of the larger RMSD value at the end of the LM300 simulation compared to the NW300 simulation (0.17 vs. 0.14 nm, respectively), probably indicative of a larger disruption of the protein tertiary structure, methanol appears to have a stabilizing effect on individual secondary-structure elements. Just as the RMSD values, the RMSF values are larger in the presence of methanol at low pH, especially in the loop regions connecting secondary-structure elements.

At 360 K, the secondary-structure elements remain remarkably stable under native environmental conditions (NW360), and the corresponding RMSF values remain very low. In comparison, all other simulations performed at 360 K are characterized by (1) a loss of the α -helical content (especially in LW360 and LM360); (2) a partial loss of the β -sheet at the level of the β_3 -, β_4 -, and β_5 -strands (especially in NM360 and LW360); (3) a loss of the 3_{10} -helix (except in simulation LM360, where it is replaced by a short α -helix); and (4) increased RMSF values. At low pH in pure water (LW360), the α -helix simultaneously shrinks from both ends and disappears almost entirely after 5.5 ns. In contrast, in the presence of methanol at both neutral and low pH (NM360 and LM360), the α -helix is more labile on its C-terminal side, whereas the N terminus remains intact for a significantly longer time. At low pH and in the methanol–water mixture (LM360), the α -helix disappears (from its C-terminal end) after ~ 9 ns. In agreement with the previous analyses, simulation NW360 (LW360) is characterized by the smallest (largest) disruption of the secondary structure and the lowest (highest) RMSF values. However, as has been noted previously, the effects of changes in the solution pH and in the solvent

Figure 4. Secondary-structure content as a function of time (horizontal graphs) and C_{α} root-mean-square atomic positional fluctuations (RMSF) as a function of the residue sequence number (vertical graphs) corresponding to selected ubiquitin simulations (Table 1). The secondary-structure assignment is performed according to the criteria proposed by Kabsch and Sander (1983): (black) 3_{10} -helix, (dark gray) α -helix, and (light gray) β -strand. Regions corresponding to β -bridges, turns, bends, or undefined structure are not displayed. The RMSF values are calculated over the entire 10-ns simulations.

composition on the RMSF are not simply additive. The addition of methanol increases the RMSF values for the simulation at neutral pH, but has the opposite effect at low pH. Similarly, the effect of the pH on the RMSF is larger when the protein is in water than when it is in the methanol–water mixture.

The violations with respect to experimental NOE interproton distance bounds (averaged over each residue and over the last 2 ns of the simulations) are displayed in Figures 5 and 6 for most of the simulations initiated from the native state of ubiquitin (Table 1; not shown for NM300 and LW300). For comparison, the corresponding violations are also displayed in Figure 5 for the experimentally derived X-ray structure and the set of 10 NMR model structures. In addition, the corresponding numbers of proton pairs satisfying or violating the bounds, in the latter case sorted by classes of violation magnitude, are reported in Table 3.

For the simulations performed at 300 K (NW300, LW300, NM300, and LM300), the number of violated NOE bounds and their distribution into magnitude classes are similar to the values reported for the experimental structures, which indicates that the native structure is overall well preserved during these simulations. However, significant differences can still be observed. Simulation LW300 presents the closest agreement with experimental data. The three other simulations (including simulation

NW300 performed under native conditions) present a somewhat higher number of violations, including more violations of large magnitude (especially for simulation LM300). The large-magnitude violations in simulation LM300 are primarily associated with long-range distance bounds between the N-terminal residues of β -sheet strands β_1 and β_5 (Fig. 5). In contrast, short- and intermediate-range distance bounds are generally satisfied. This is consistent with the picture that the secondary-structure elements are well preserved in the condition stabilizing the A-state, while the opposite is true for the tertiary structure.

Among the simulations performed at 360 K (NW360, LW360, NM360, and LM360), simulation NW360 performed under native conditions stands out as presenting the best agreement with experiment. The number of large-magnitude violations (0.2–0.4 nm and >0.4 nm) is even lower than in the corresponding simulation at 300 K. This indicates that under native conditions, the protein preserves not only its secondary structure but also most of its tertiary interactions even at 360 K (on the 10-ns timescale), in good qualitative agreement with the high experimental melting temperature (355 K) of ubiquitin in these environmental conditions (Lenkinski et al. 1977; Cary et al. 1980). In all other simulations at 360 K, the significant increase in the number of violations (in particular, the large-magnitude ones) and the distribution

Figure 5. Average violations of the calculated interproton distances with respect to the corresponding experimentally derived (NOE) upper bounds as a function of the residue sequence number, for the crystallographic (X-ray) structure (Vijay-Kumar et al. 1987), the set of 10 NMR model structures (Cornilescu et al. 1998), and the simulations NW300 and LM300 (see Table 1). The differences are calculated based on r^{-3} -averaging over the structures or trajectory configurations (last 2 ns), whereby only positive values are considered as violations, and the possible violations are averaged over all proton pairs belonging to a given residue pair. The black squares in the *upper-left* triangles indicate residue pairs involving one or more experimental NOE-derived distance bounds, none of them being violated. The squares in the *bottom-right* triangles indicate residue pairs with positive average violations of (green) 0.0–0.2 nm, (blue) 0.2–0.4 nm, and (red) >0.4 nm. The secondary-structure elements present in the native state of ubiquitin are indicated along with the residue numbers on each axis. Note that the number of squares of a given kind is not directly comparable to the data in Table 2 (given on a per-proton-pair basis) because of averaging over the residues.

of the corresponding pairs along the sequence (Fig. 6) indicate a decrease of the secondary-structure stability as well as a disruption of the protein tertiary interactions. In simulation NM360, long-range distance violations appear between the C-terminal part (starting with the 3_{10} -helix) and the rest of the protein, indicating a loss of tertiary packing between these two regions induced by the presence of the methanol cosolvent. In addition, short- and medium-range distance violations appear at the level of the α -helix and of the β -sheet strands β_3 and β_4 , revealing a fraying effect at the C terminus of the α -helix as well as a separation between the two β -sheet strands and between the α -helix and the strand β_3 in agreement with the previous analysis of the secondary structure (see Figs. 1 and 4). The C-terminal part of the α -helix includes

the hydrophobic contact between residue Ile30 and residue Ile36 that has been recently shown by mutational studies to significantly contribute to the thermodynamic stability of the ubiquitin native structure, probably by providing a C-terminal capping to this helix (Thomas and Makhatadze 2000). The average distance of interproton violations between these two residues is 0.29 nm in simulation NM360 (vs. 0.13 nm for simulation NW360). This observation is consistent with a reduction in the magnitude of the hydrophobic effect induced by the presence of methanol. In the simulations LW360 and LM360 performed at low pH, long-range distance violations appear primarily between the α -helix and the rest of the protein. Simulation LW360 presents additional violations between the β -sheet strand β_3 and the rest of the

Figure 6. Average violations of the calculated interproton distances with respect to the experimentally derived (NOE) upper bounds as a function of the residue sequence number, for the simulations NW360, NM360, LW360, and LM360 (see Table 1). See the legend for Figure 5.

protein, and between the region extending from the β -sheet strand β_4 to the 3_{10} -helix and the N terminus of the α -helix. For the simulation LW360, short- and intermediate-range violations primarily involve the β -sheet strands β_3 and β_4 (as in NM360), as well as the 3_{10} -helix and the α -helix. In comparison, short- and intermediate-range violations are more limited for simulation LM360, except for those involving the C-terminal half of the α -helix, and reflect a relative preservation of the secondary-structure elements other than the α -helix. Under these low pH conditions, the α -helix is pulled away from the rest of the protein and progressively disappears (see Figs. 1 and 4). The destabilization of the α -helix at low pH is thus probably primarily caused by the disruption of the salt bridges Lys111/Glu34, Lys27/Glu51, and Lys27/Asp52 (between the two termini of the α -helix and the β -sheet), while the reduction of the capping interaction (Ile30/Ile36) in the presence of methanol appears to play a less important role at this pH. Nevertheless, this hydrophobic contact is more altered in simulation

LM360, where the corresponding average of interproton distance violations is 0.24 nm vs. 0.17 nm for simulation LW360.

The comparison between experimental and calculated 3J -coupling constants (amide to C_α proton) for the simulations performed at 360 K is displayed in Figure 7, together with the corresponding data for the experimentally derived X-ray structure and the set of 10 NMR model structures. The corresponding properties of the correlation between simulated and experimental values are also reported in Table 3. As expected, the agreement between experimental and calculated 3J -coupling constants for the X-ray and NMR structures is very high. Owing to the structure determination procedure (partly based on the 3J -coupling constant data), the NMR structures stand out as presenting the best agreement with experiment. For the simulations performed at 300 K, the deviations are (expectedly) somewhat larger, but still the agreement with experimental data remains reasonable. Interestingly, the correlation is best (correlation

Table 3. Average NMR-related properties calculated for the X-ray structure (Vijay-Kumar et al. 1987), the set of 10 NMR model structures (Cornilescu et al. 1998), and the different simulations

	$\langle N_{\text{viol}} \rangle_{\text{end}}$				$\langle \Delta J \rangle_{\text{end}}$		
	0.0 nm	0.0–0.2 nm	0.2–0.4 nm	>0.4 nm	RMSD (Hz)	Corr	Slope
X-ray	1098	515	47	1	0.10	0.94	0.89
NMR	1090	539	31	1	0.07	0.97	0.95
NW300	1003	571	80	7	0.28	0.61	0.60
NM300	1020	558	76	7	0.21	0.86	0.82
LW300	1040	570	49	2	0.25	0.67	0.65
LM300	1016	563	62	20	0.23	0.72	0.54
NW360	989	597	70	4	0.22	0.78	0.62
NM360	874	540	162	85	0.23	0.73	0.63
LW360	783	465	171	242	0.30	0.31	0.20
LM360	824	504	186	147	0.23	0.65	0.42
LM ₁ 300	827	467	145	222	0.27	0.46	0.35
LM ₁ 360	725	403	141	392	0.27	0.39	0.18
LM ₁ 400	614	287	147	613	0.28	0.24	0.08
LM ₂ 300(360)	708	367	120	466	0.33	0.08	0.05
LM ₂ 300(400)	584	283	116	678	0.33	0.13	0.07
A300	642	323	102	594	0.41	0.23	0.19
A360	608	262	93	698	0.36	0.08	0.06

See Table 1 and Figure 2. The number of interproton distance violations with respect to the corresponding experimentally derived (NOE) distance upper bounds (N_{viol} , sorted by classes of violation magnitudes) (see also Figs. 5 and 6) and the properties of the correlation between experiment and simulation for the 3J -coupling constant between amide and C_{α} protons (ΔJ) (see Fig. 7) are reported. Averaging is performed over the last (end) 2 ns of the trajectories. The results for ΔJ correspond to the RMSD between simulated and experimental values, together with the correlation coefficient (Corr) and slope of a regression line fitting simulation data as a function of experimental data.

coefficient and slope closest to one, smallest RMSD) for simulations NM300 and, to a lesser extent, LM300. Indeed, this is consistent with the ability of methanol to better preserve the secondary-structure elements and thus, the local conformational properties of the polypeptide. For the simulations performed at 360 K, the simulations characterized by the largest structural deviations from the native structure (Fig. 2; Table 2) consistently present the largest discrepancy with experimental coupling constants. The discrepancy systematically increases upon raising the temperature, except for the simulation in water at neutral pH (NW300 vs. NW360). In all cases, discrepancies appear in the form of a larger RMSD between simulated and experimental coupling constants, a lower correlation coefficient, and a smaller slope of the corresponding regression line (simulation data as a function of experimental data). The latter observation is at least in part a consequence of the functional form of the Karplus equation, which leads to an average value of ~ 8 Hz when the corresponding backbone dihedral angle undergoes (quasi-) unhindered rotation. It also probably reflects the reduction of the helical content, for which 3J -coupling constants are typically low, ~ 3.9 Hz (α -helix) or 4.2 Hz (3_{10} -helix). These two effects are particularly evident for simulation LW360 and, to a lesser extent, simulations NM360 and LM360 (Fig. 7).

Secondary chemical shifts are qualitative indicators of the secondary-structure content of proteins (Richarz and

Wuethrich 1978; Brutscher et al. 1997). Their values determined from simulated trajectories may be directly compared with the experimental data obtained from NMR measurements. The secondary chemical shifts calculated for the simulations initiated from the native state of ubiquitin and performed at 360 K (Table 1), as well as for the continuation run LM₁360, are displayed in Figure 8. Although small deviations are found, the good qualitative agreement between the experimental secondary chemical shifts and the ones calculated for the X-ray and the set of 10 NMR structures provides support for the reliability of this analysis. The positions of β -strands (indicated by negative values) and of helices (indicated by positive values) are found to correspond very well. Consistent with the results of the previous analyses, the secondary chemical shifts calculated from the simulation NW360 are very similar to the experimental ones, apart from a small discrepancy at residues 62–65 (located at the N terminus of the β -sheet strand β_5). The latter difference may be related to the fact that in the simulation, β_5 is pulled away from the rest of the protein. In all other simulations performed at 360 K, the calculated secondary chemical shifts suggest a disappearance of the α -helix (from its C-terminal end), most pronounced in simulation LW360 (see also Fig. 1). In the continuation run LM₁360, the calculated chemical shifts indicate a complete disappearance of the α -helix, while only the 3_{10} -helix appears to remain present.

Figure 7. Comparison between experimental (horizontal axis) and calculated (vertical axis) amide to C_{α} protons' 3J -coupling constants for the X-ray structure (Vijay-Kumar et al. 1987), the set of 10 NMR model structures (Cornilescu et al. 1998), and the simulations NW360, NM360, LW360, and LM360 (see Table 1). The calculations were performed using the Karplus relationship of Equation 1 (Wang and Bax 1996), based on the last 2 ns of the simulations. The experimental data are taken from Wang and Bax (1996).

In summary, the above analysis leads to the following observations:

1. Ubiquitin remains remarkably stable in water at neutral pH (in agreement with experiment), even at an elevated temperature of 360 K.
2. At 300 K, the native structure is essentially stable under all simulation conditions, with an approximated ranking $LW300 > NW300 \sim NM300 > LM300$ in terms of decreasing stability.
3. At this temperature, methanol induces a stabilization of the secondary structure and a reduction of the tertiary interactions.
4. A reduction of the pH further destabilizes the protein in the presence of methanol, but stabilizes it in pure water.
5. At 360 K, the deviations from the native structure become large (except in water at neutral pH), with an approximate ranking $NW360 > NM360 > LM360 \gg LW360$ in terms of decreasing stability.
6. At this temperature, a reduction of the pH induces the largest stability changes (disappearance of the α -helix, probably related to the absence of the salt-bridge interactions anchoring it to the β -sheet).
7. Methanol slightly destabilizes the protein at neutral pH, but slightly stabilizes it at low pH.

Simulations initiated from the A-state model

Most of the analyses reported so far have also been performed for the simulations A300 and A360 initiated

from the A-state model (Table 1), and the resulting average values are reported in Tables 2 and 3. Expectedly, most of the data characterizing deviations from the native structure (RMSD, radius of gyration, interproton distance violations, and correlation with experimental 3J -coupling constants) present extreme values compared to the simulations started from the native structure of the protein. In contrast, the average numbers of backbone–backbone hydrogen bonds are comparable to the ones observed in the latter simulations (especially at 300 K). This is because the overall secondary-structure content of the A-state is similar to that of the native state and is largely preserved during the simulations (especially at 300 K).

C_{α} secondary chemical shifts, calculated for the model A-state and the simulations A300 and A360, are displayed in Figure 9 together with the corresponding experimental values. The evolution of the secondary structure as a function of time is also shown for comparison. In addition, snapshots taken at successive 0.3-ns intervals along the trajectory of simulation A300 are displayed in Figure 10. The simulation A300 is in good qualitative agreement with experiment in several respects. The N-terminal half of the protein remains native-like in its secondary structure along the entire simulations including the β -sheet strands β_1 and β_2 (residues 1–17) and the α -helix (residues 21–35) (Brutscher et al. 1997). However, after ~ 1 ns, the long (nonnative) α -helix at the C-terminal half of the protein is broken into two smaller helix pieces (residues 40–50 and 54–71) separated by a region with no regular structure. The results

Figure 8. C_{α} secondary chemical shifts ($\Delta\delta$) from experiment (303 K, at pH 5.7, in water) (Wand et al. 1996), and calculated for the X-ray structure (Vijay-Kumar et al. 1987), the set of 10 NMR model structures (Cornilescu et al. 1998), and the simulations NW360, NM360, LM360, LW360, and LM1360 (see Table 1; Fig. 2). The calculations were performed using the SHIFTS program (Xu and Case 2001), based on the last 2 ns of the simulations. The residue-type-specific random-coil chemical shifts in water were taken from Richarz and Wüthrich (1978) (without correction for the methanol–water environment).

of simulation A300 agree well with the model proposed by Brutscher et al. (1997). According to these investigators, the A-state is a highly flexible assembly of three relatively rigid motional units with a defined secondary structure. The secondary chemical shifts calculated from simulation A300 also agree very well with the experimental values. The calculated secondary chemical shifts for residues 51–53 are, as expected from the secondary-structure analysis, close to zero or negative. The agreement with the experimental values for this region provides support for the kink within the C-terminal α -helix observed during the simulation. The situation is different at 360 K, where only the β -hairpin and the α -helix remain until the end of the simulation. Thus, in agreement with experimental findings for the A-state and in contrast with corresponding observations for the native state, the structure of the A-state appears to be significantly destabilized at elevated temperature (Harding et al. 1991).

Although the compatibility of the schematic A-state model proposed by Brutscher et al. (1997) and of the simulation results at 300 K with experimental data is

reasonable, the structures reported in Figures 1 and 10 should be taken with some caution. Owing to the limited amount of experimental data (underdetermination), the enhanced segmental mobility in this state (compared to the native one), and the short timescale of the simulation (3 ns), these structures probably represent illustrative conformations in the corresponding conformational ensemble rather than accurate structural models. A full structural relaxation of the initial structure based on such an approximate A-state model would probably require orders of magnitude (i.e., prohibitively) longer simulations.

Protein–solvent interactions

The radial distribution functions for the water (OW) and methanol oxygen (OM) atoms around the center of geometry of the protein are shown in Figure 11 for the simulations performed in pure water and in the 60% methanol–water mixture. In pure water, the density profiles for the water oxygen atoms are similar for the four systems (different pH values and temperatures), and show an essentially monotonic progression from 0 to 1 between the protein surface and bulk solvent. When methanol is added, the local water density (relative to

Figure 9. C_{α} secondary chemical shifts ($\Delta\delta$) (four top panels) from experiment (300 K, at pH 2 in a 60% methanol–water mixture) (Brutscher et al. 1997) and calculated for the model A-state and the simulations A300 and A360 (see Table 1), and secondary-structure content as a function of time (two bottom panels) for the two latter simulations. See the legends for Figures 4 and 8.

Figure 10. Ribbon diagrams of ubiquitin highlighting the secondary-structure content as a function of time for the simulation A300 (see Table 1). Snapshots are taken at successive 0.3-ns intervals.

bulk) is reduced close to the protein surface because of partial replacement by methanol molecules (with the emergence of a density peak for methanol around 1.5–2.0 nm for simulations NM300, LM300, and LM360). The curves clearly suggest a preferential solvation of the protein by methanol in the mixture, that is, the solvent composition in the immediate vicinity of the protein is richer in methanol compared to the bulk composition. For the simulations performed at neutral pH, the preferential-solvation effect is less pronounced at 360 K compared to 300 K, likely a consequence of the increased thermal motion (leading to a homogenization of the concentration gradient). At 300 K, the pH value seems to have little influence on the solvation. However, the effect of pH is very pronounced at 360 K. In simulation LM360, a large peak in the methanol density appears at ~ 0.3 – 0.5 nm, indicating that the protein core has become partly solvent-accessible. The solvation of the exposed (hydrophobic) core residues almost exclusively involves methanol molecules (despite its larger size compared to the water molecule), which is probably related to the higher hydrophobicity of this solvent.

The average numbers of protein–solvent hydrogen bonds for the different simulations (Table 1; Fig. 2) are reported in Table 4, sorted by solvent type and donor–acceptor relationship. For the simulations at 300 K initiated from the native state of the protein, the maximum total number of protein–solvent hydrogen bonds (166) is found for simulation NW300, with a predominance of hydrogen bonds involving water as a donor. The second largest total number of hydrogen bonds (146) is observed in simulation NM300, with the same predom-

inance of hydrogen bonds involving the solvent molecules as donors. The protein–solvent hydrogen bonds are almost equally distributed between water and methanol, despite the smaller total number of methanol molecules in solution (Table 1) and the fact that methanol can only donate a single hydrogen bond (as opposed to two for water). In the corresponding simulations performed at low pH, the total number of protein–solvent hydrogen bonds is significantly lower (124 and 102 for simulations LW300 and LM300, respectively). However, while the number of hydrogen bonds involving the solvent as a donor largely decreases, the number of those involving the solvent as an acceptor slightly increases. In contrast with the simulations performed at neutral pH, there is now a predominance of hydrogen bonds involving the solvent acting as an acceptor. This observation is certainly related to the different charge states of the protein at the different pH values (total charge +1e and charged acidic residues at neutral pH; total charge +13e and neutral acidic residues at low pH). For the four different environmental conditions, the number of protein–solvent hydrogen bonds decreases upon increasing the temperature from 300 K to 360 K.

In the simulations A300 and A360 initiated from the A-state model, although residues are more significantly exposed to the solvent, the total number of solute–solvent hydrogen bonds is only slightly increased (by ~ 20) compared to the values observed in the corresponding simulations LM300 and LM360 initiated from the native structure. This can be explained by the fact that the hydrophilic side chains are already largely exposed to solvent in the native structure so that the A-state differs

Figure 11. Radial distribution functions $g(r)$ for the water (OW) and methanol oxygen (OM) atoms shown as a function of the distance r from the center of geometry of the protein for the simulations NW300, NW360, NM300, NM360, LW300, LW360, LM300, and LM360 (see Table 1). Averaging is performed over the last 2 ns of the simulations.

Table 4. Average number of protein–solvent hydrogen bonds calculated for the different simulations

	$\langle \text{hb} \rangle_{\text{P} \rightarrow \text{Met}}$	$\langle \text{hb} \rangle_{\text{Met} \rightarrow \text{P}}$	$\langle \text{hb} \rangle_{\text{P} \rightarrow \text{Wat}}$	$\langle \text{hb} \rangle_{\text{Wat} \rightarrow \text{P}}$
NW300	—	—	61 (14)	105 (29)
NM300	33 (7)	41 (3)	30 (4)	42 (12)
LW300	—	—	78 (17)	46 (24)
LM300	41 (13)	6 (3)	36 (6)	19 (10)
NW360	—	—	55 (11)	95 (26)
NM360	27 (5)	35 (4)	28 (3)	38 (10)
LW360	—	—	72 (18)	41 (21)
LM360	41 (12)	9 (5)	34 (8)	19 (10)
LM ₁ 300	46 (14)	10 (6)	39 (9)	21 (12)
LM ₁ 360	45 (16)	11 (6)	39 (10)	22 (13)
LM ₁ 400	37 (12)	10 (6)	35 (8)	18 (11)
LM ₂ 300(360)	49 (18)	11 (7)	42 (11)	22 (13)
LM ₂ 300(400)	50 (19)	10 (7)	40 (9)	23 (16)
A300	46 (13)	10 (5)	40 (9)	22 (13)
A360	48 (16)	12 (6)	38 (8)	22 (12)

See Table 1 and Figure 2. The number of hydrogen bonds involving methanol as an acceptor ($\text{P} \rightarrow \text{Met}$) or as a donor ($\text{Met} \rightarrow \text{P}$), and water as an acceptor ($\text{P} \rightarrow \text{Wat}$) or as a donor ($\text{Wat} \rightarrow \text{P}$) are reported. The average number of hydrogen bonds involving protein backbone atoms is indicated in parentheses. Averaging is performed over the last 2 ns of the simulations.

predominantly by the exposure of hydrophobic side chains less prone to hydrogen bonding.

The distributions of the methanol-to-water molecule ratio (normalized to the bulk values) around side chains

of different types (charged, polar, or nonpolar) are displayed in Figure 12 for simulation A300. In all cases, the proportion of methanol molecules is higher in the neighborhood of the side chains compared to the bulk (ratio up to ~ 2.5 times the bulk ratio in the first peak). Around the positively charged side chains, the ratio abruptly falls down to the bulk value after this first peak. Around other types of side chains, the ratio remains above the bulk value even up to a distance of ~ 1.5 nm.

Conclusion

In this study, a series of explicit-solvent MD simulations of ubiquitin under different environmental conditions (presence of methanol cosolvent, pH, and temperature) and initiated from different structures (native, model of A-state) is reported. During simulations performed under native conditions, the protein remains remarkably stable (even at an elevated temperature of 360 K), while the generated configurational ensemble is compatible with experimental (NMR) data. This provides additional validation for the GROMOS force-field parameter set 45A4. The presence of the methanol cosolvent enhances the stability of secondary-structure elements (except the C-terminal end of the α -helix at 360 K) but weakens tertiary interactions. This is consistent with the expected ability of methanol to mitigate a hydrophobic environment in a nonspecific way. However, large deviations from the

Figure 12. Distributions of the methanol-to-water molecule ratio (normalized to the bulk values) around side chains of different types for the simulation A300 (see Table 1). Side-chain types are sorted as follows: acidic (Asp C_β , Glu C_γ ; neutral, polar), basic (His C_γ , Lys C_ϵ , Arg C_δ ; positively charged), amide (Gln C_γ , Asn C_β ; polar), alcohol (Ser C_β , Thr C_β ; polar), aliphatic (Leu C_γ , Ile C_γ , Met C_γ ; nonpolar), and aromatic and heterocyclic (Phe C_γ , Pro C_γ ; nonpolar). Averaging is performed over the last 2 ns of the simulations.

protein native structure only occur when the pH is lowered. They are characterized mostly by the disappearance of the α -helix and the loss of tertiary interactions between the corresponding peptide segment and the rest of the protein, probably related to the disruption of the salt bridges anchoring the α -helix to the β -sheet. Surprisingly, the simultaneous addition of methanol and decrease of the pH appears to have less impact on the protein stability than a pH decrease alone (in water), indicating that the effect of methanol on the structure strongly depends on the charge state of the protein. In particular, the charge state of the protein largely influences the density profiles of water at the protein surface and the number and distribution of solute-solvent hydrogen bonds. These results indicate that the influences of pH and methanol cosolvent are not simply additive, but that a more complex interplay is present between these two environmental factors. Finally, the simulations initiated from the A-state model present good agreement with the available (NMR) data and are consistent with the model proposed by Brutscher et al. (1997). The A-state consists of a β -hairpin, a native-like α -helix, and two nonnative α -helices connected by flexible linkers and defining four relatively rigid independent motional units.

Materials and Methods

All simulations (see Table 1; Fig. 2) were carried out using the GROMOS96 simulation program (van Gunsteren et al. 1996; Scott et al. 1999) together with the united-atom parameter set 45A4 (Schuler et al. 2001; Soares et al. 2005; Lins and Hünenberger 2005), the SPC water model (Berendsen et al. 1981), and a GROMOS-compatible methanol model (Walser et al. 2000). Initial coordinates for the native structure of ubiquitin were taken from the 0.18-nm resolution crystallographic structure 1UBQ (Dastidar and Mukhopadhyay 2005) of the Protein Data Bank (Berman et al. 2000). Initial coordinates for a partially folded structure assumed representative of the A-state were built based on the schematic model proposed by Brutscher et al. (1997) (see Fig. 1 therein). To assess the compatibility of this model structure with experimental data, the α -carbon (C_α) and carbonyl carbon (C_O) chemical shifts were calculated for all residues using the SHIFTS program (Xu and Case 2001). These were found to agree well with the corresponding experimental values (Brutscher et al. 1997; Fig. 9).

Two simulations of ubiquitin were performed starting from the folded structure under conditions experimentally known to either preserve the native state (at 300 K and neutral pH in pure water) or to induce a transition to the A-state (at 300 K and low pH in a 60% methanol-water mixture, i.e., the solution resulting from the mixing of six volumes of methanol with four volumes of water). Two additional simulations were performed using intermediate conditions (at 300 K and low pH in pure water or at 300 K and neutral pH in the 60% methanol-water mixture). Finally, a fifth simulation was started from the modeled A-state structure under conditions experimentally known to stabilize the A-state (at 300 K and low pH in the 60% methanol-water mixture). The indications "neutral" and "low" pH refer to pH values of 5.6 (crystallographic structure determination of the

native state) (Vijay-Kumar et al. 1987) and 2 (structural investigation of the partially folded A-state) (Brutscher et al. 1997), respectively. In practice, the neutral pH is represented through positively charged basic residues (including the N terminus) and negatively charged acidic residues (including the C terminus), while the low pH is represented through positively charged basic residues and neutral acidic residues. These choices lead to total protein charges of +1e and +13e, respectively, for the neutral- and low-pH conditions. These charges were neutralized by inclusion of one or 13 chloride anions, respectively, in the simulated systems. Finally, another set of five simulations was performed using the same starting structures and simulation conditions, but at an elevated temperature of 360 K. The systems and conditions corresponding to these 10 simulations are summarized in Table 1.

All simulations were performed under periodic boundary conditions based on truncated-octahedron computational boxes. The protein was solvated using solvent (and cosolvent) molecules from equilibrated configurations of pure water (density 0.972 g/cm³) (Berendsen et al. 1981) or of a 60% methanol-water mixture (density 0.897 g/cm³) (Mikhail and Kimel 1961), respecting a minimum distance of 0.23 nm between solute and solvent atoms and a minimum distance of 1.6 nm (simulations starting from the native state) or 1.0 nm (simulations starting from the A-state model) between solute atoms and the closest box wall. This procedure led to truncated-octahedron boxes derived from cubes of approximate edge lengths 7.8 nm or 13.4 nm, respectively. After energy minimization and equilibration of the solvent molecules with positionally constrained solute atoms, the chloride counterions were added by replacing in turn the water molecules with the highest electrostatic potential at their oxygen atom. The relaxation of the solvent and counterions around the protein was then performed in a stepwise fashion at constant volume. Starting from initial velocities based on a Maxwell distribution at 60 K and harmonic positional restraints on all protein atoms with a force constant of 5.0×10^4 kJ/(mol · nm²), the temperature was progressively increased to 120, 180, 240, and 300 K while the harmonic force constant was simultaneously reduced to 2.5, 1.25, 0.6, 0.3×10^4 kJ/(mol · nm²) at intervals of 20 ps (totaling 100 ps of equilibration). At this point, production MD simulations were performed at a constant pressure of 1 bar and at the temperature of interest (300 or 360 K) for 3 or 10 ns (Table 1).

The final configuration of the 10-ns simulation at 360 K starting from the native structure and performed under conditions promoting the A-state (LM360) was used as the starting point for additional simulations with the same environmental conditions but at different temperatures (Fig. 2). In a first series of 5-ns simulations, the temperature was set to 300, 360, or 400 K (the lowest one as an attempt to quench the system into the A-state, and the highest one meant to speed up the conformational transformation). In subsequent 5-ns simulations, the two simulations at the highest temperatures (360 K and 400 K) were continued at 300 K (again as an attempt to quench the system after an enhanced conformational search).

For all simulations, the equations of motion were integrated based on the leapfrog scheme using a 2-fs timestep. Bonds between hydrogen and heavy atoms as well as the geometry of the solvent molecules were constrained using the SHAKE algorithm (Ryckaert et al. 1977) with a relative geometric tolerance of 10^{-4} . The temperatures of solute (including ions and cosolvent) and solvent were controlled by coupling them separately to a Berendsen thermostat (Berendsen et al. 1984) with a relaxation time of 0.1 ps. The pressure was maintained

by coupling to a Berendsen barostat via isotropic coordinate scaling (Berendsen et al. 1984) with a coupling constant of 0.5 ps and a compressibility of $4.575 \times 10^{-4} \text{ kJ}/(\text{mol} \cdot \text{nm}^3)^{-1}$. Nonbonded interactions were handled using a twin-range approach (van Gunsteren et al. 1996; Scott et al. 1999) with short- and long-range cutoff distances of 0.8 and 1.4 nm, respectively. The pair list for short-range nonbonded interactions was updated every 10 fs, together with the intermediate-range nonbonded forces, virial and energy. The mean effect of electrostatic interactions beyond the long-range cutoff distance was included using a reaction-field approximation (Barker and Watts 1973) based on a relative dielectric permittivity of 66 for the solvent (Glattli et al. 2002). Configurations were recorded every 0.5 ps for analysis.

The analysis of the simulations was performed in terms of C_α root-mean-square atomic positional deviation from the crystallographic native structure (RMSD), C_α root-mean-square atomic positional fluctuations (RMSF), secondary-structure content, radius of gyration (rgyr), protein–protein and protein–solvent hydrogen bonds (hb), NOE distance bound violations, NMR 3J -coupling constants, secondary chemical shifts, and radial distribution functions.

The definition used for assessing the presence of a hydrogen bond was based on a geometric criterion. If the distance between the hydrogen of the donor and the oxygen of the acceptor is below 0.25 nm, and the donor–hydrogen–acceptor angle is larger than 135° , the atoms are assumed to participate in a hydrogen bond. The secondary-structure assignment was performed according to the criteria proposed by Kabsch and Sander (1983).

A set of 2727 interproton distance bounds was available from experimentally determined NOE intensities for ubiquitin under native conditions (temperature 304 K, pH 6.6, water) (Cornilescu et al. 1998). These distance bounds include pseudo-atom distances and multiplicity corrections as given by Wuethrich et al. (1983). Interproton distances were calculated along the different MD trajectories based on r^{-3} -averaging (Tropp 1980). Distance violations were then computed as the deviation between these average interproton distances and the corresponding NOE-derived upper bounds. For the experimental NOE cross-peak intensities where no stereospecific assignment could be made or where individual hydrogen atoms were experimentally indistinguishable (leading to a distance bound that can be associated to any member in a list of possible proton pairs), the distance violations were calculated for each possible pair based on the simulation of the protein under native conditions, and the bound was assigned to the pair resulting in the smallest distance violation. The resulting list of NOE-derived distances contains proton-pair degeneracies. For each degeneracy only the smallest upper-bound distance was selected, reducing the 2727 experimental NOE-derived bounds to a set of 1661. This final set was used to calculate the violations associated with specific proton pairs during all simulations.

Averages for the $^3J_{H^N H^\alpha}$ coupling constants between the peptide and α -carbon protons were calculated using the generalized Karplus equation derived for polypeptides by Wang and Bax (1996):

$$^3J_{H^N H^\alpha} = 6.98 \cos^2(\phi - 60) - 1.38 \cos(\phi - 60) + 1.72, \quad (1)$$

where ϕ denotes the dihedral backbone angle C–N–C $_\alpha$ –C (in degrees).

The C_α and C_O chemical shifts for the α - and carbonyl-carbons were computed using the program SHIFTS (Xu and Case 2001), which uses a database of chemical shifts (calculated using density-functional theory) to predict ^{13}C shifts in proteins. The ^{13}C shifts can be used to identify regular

secondary-structure elements along a polypeptide chain, by subtracting the ^{13}C shift associated with the specific residue in a random-coil model (obtained from NMR measurements on model peptides) (Richarz and Wuethrich 1978), leading to so-called secondary chemical shifts. Secondary chemical shifts were calculated along the different trajectories and are reported as time averages.

Unless otherwise specified, all average properties reported in the figures and tables are computed separately over the first and last 2 ns of each simulation.

Acknowledgments

We thank Dr. Chris Oostenbrink for his help in using GROMOS++ and Prof. Othmar Steinhauser for insightful discussions. Financial support by the National Center of Competence in Research (NCCR) Structural Biology of the Swiss National Science Foundation (SNSF) is also gratefully acknowledged.

References

- Alexandrescu, A.T., Ng, Y.L., and Dobson, C.M. 1994. Characterization of a trifluoroethanol-induced partially folded state of α -lactalbumin. *J. Mol. Biol.* **235**: 587–599.
- Alonso, D.O. and Daggett, V. 1995. Molecular dynamics simulations of protein unfolding and limited refolding: Characterization of partially unfolded states of ubiquitin in 60% methanol and in water. *J. Mol. Biol.* **247**: 501–520.
- Arai, M. and Kuwajima, K. 2000. Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**: 209–282.
- Barker, J.A. and Watts, R.O. 1973. Monte-Carlo studies of the dielectric properties of water-like models. *Mol. Phys.* **26**: 789–792.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., and Hermans, J. 1981. *Interaction models for water in relation to protein hydration*. Jerusalem Symposia on Quantum Chemistry and Biochemistry, Vol. 14, pp. 331–342. Reidel, Dordrecht, Holland.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**: 3684–3690.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The Protein Data Bank. *Nucleic Acids Res.* **28**: 235–242.
- Brown, J.E. and Klee, W.A. 1971. Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* **10**: 470–476.
- Brutscher, B., Brueschweiler, R., and Ernst, R.R. 1997. Backbone dynamics and structural characterization of the partially folded A state of ubiquitin by ^1H , ^{13}C , and ^{15}N nuclear magnetic resonance spectroscopy. *Biochemistry* **36**: 13043–13053.
- Buck, M., Radford, S.E., and Dobson, C.M. 1993. A partially folded state of hen egg white lysozyme in trifluoroethanol: Structural characterization and implications for protein folding. *Biochemistry* **32**: 669–678.
- Bychkova, V.E. and Ptitsyn, O.B. 1993. The molten globule state of protein molecules is becoming a rule rather than exception. *Biofizika* **38**: 58–66.
- Cary, P.D., King, D.S., Crane-Robinson, C., Bradbury, E.M., Rabbani, A., Goodwin, G.H., and Johns, E.W. 1980. Structural studies on two high-mobility-group proteins from calf thymus, HMG-14 and HMG-20 (ubiquitin), and their interaction with DNA. *Eur. J. Biochem.* **112**: 577–580.
- Chin, D.T., Kuehl, L., and Rechsteiner, M. 1982. Conjugation of ubiquitin to denatured hemoglobin is proportional to the rate of hemoglobin degradation in HeLa cells. *Proc. Natl. Acad. Sci.* **79**: 5857–5861.
- Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C., and Gronenborn, A.M. 1990. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.* **112**: 4989–4991.
- Cordier, F. and Grzesiek, S. 2002. Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. *J. Mol. Biol.* **317**: 739–752.
- Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. 1998. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**: 6836–6837.

- Cox, J.P.L., Evans, P.A., Packman, L.C., Williams, D.H., and Woolfson, D.N. 1993. Dissecting the structure of a partially folded protein. Circular dichroism and nuclear magnetic resonance studies of peptides from ubiquitin. *J. Mol. Biol.* **234**: 483–492.
- Cremades, N., Sancho, J., and Freire, E. 2006. The native-state ensemble of proteins provides clues for folding, misfolding and function. *Trends Biochem. Sci.* **31**: 494–496.
- Daggett, V. and Fersht, A.R. 2003. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* **28**: 18–25.
- Dastidar, S.G. and Mukhopadhyay, C. 2005. Unfolding dynamics of the protein ubiquitin: Insight from simulation. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **72**: 051928/1–051928/10.
- Daura, X., Mark, A.E., and van Gunsteren, W.F. 1998. Parametrization of aliphatic CH_n united atoms of GROMOS96 force field. *J. Comput. Chem.* **19**: 535–547.
- Dyson, H.J. and Wright, P.E. 1998. Equilibrium NMR studies of unfolded and partially folded proteins. *Nat. Struct. Biol.* **5**: 499–503.
- Dyson, H.J. and Wright, P.E. 2005. Elucidation of the protein folding landscape by NMR. *Methods Enzymol.* **394**: 299–321.
- Glatli, A., Daura, X., and van Gunsteren, W.F. 2002. Derivation of an improved simple point charge model for liquid water: SPC/A and SPC/L. *J. Chem. Phys.* **116**: 9811–9828.
- Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D.H., Niall, H.D., and Boyse, E.A. 1975. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci.* **72**: 11–15.
- Harding, M.M., Williams, D.H., and Woolfson, D.N. 1991. Characterization of a partially denatured state of a protein by two-dimensional NMR: Reduction of the hydrophobic interactions in ubiquitin. *Biochemistry* **30**: 3120–3128.
- Hershko, A., Heller, H., Eytan, E., Kaklij, G., and Rose, I.A. 1984. Role of the α -amino group of protein in ubiquitin-mediated protein breakdown. *Proc. Natl. Acad. Sci.* **81**: 7021–7025.
- Hirota, N., Mizuno, K., and Goto, Y. 1997. Cooperative α -helix formation of β -lactoglobulin and melittin induced by hexafluoroisopropanol. *Protein Sci.* **6**: 416–421.
- Jahn, T.R. and Radford, E.S. 2005. The yin and yang of protein folding. *FEBS J.* **272**: 5962–5970.
- Jourdan, M. and Searle, M.S. 2001. Insights into the stability of native and partially folded states of ubiquitin: Effects of cosolvents and denaturants on the thermodynamics of protein folding. *Biochemistry* **40**: 10317–10325.
- Kabsch, W. and Sander, C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**: 2577–2637.
- Kim, P.S. and Baldwin, R.L. 1982. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**: 459–489.
- Kuwajima, K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* **6**: 87–103.
- Lenkinski, R.E., Chen, D.M., Glickson, J.D., and Goldstein, G. 1977. Nuclear magnetic resonance studies of the denaturation of ubiquitin. *Biochim. Biophys. Acta* **494**: 126–130.
- Lins, R.D. and Hünenberger, P.H. 2005. A new GROMOS force field for hexopyranose-based carbohydrates. *J. Comput. Chem.* **26**: 1400–1412.
- Mikhail, S.Z. and Kimel, W.R. 1961. Densities and viscosities of methanol-water mixtures. *J. Chem. Eng. Data* **6**: 533–537.
- Pan, Y. and Briggs, M.S. 1992. Hydrogen exchange in native and alcohol forms of ubiquitin. *Biochemistry* **31**: 11405–11412.
- Prompers, J.J. and Brueschweiler, R. 2000. Thermodynamic interpretation of NMR relaxation parameters in proteins in the presence of motional correlations. *J. Phys. Chem. B* **104**: 11416–11424.
- Prompers, J.J., Scheurer, C., and Bruschweiler, R. 2001. Characterization of NMR relaxation-active motions of a partially folded A-state analogue of ubiquitin. *J. Mol. Biol.* **305**: 1085–1097.
- Ptitsyn, O.B. 1987. Protein folding: Hypotheses and experiments. *J. Protein Chem.* **6**: 273–293.
- Ptitsyn, O.B. 1995. Molten globule and protein folding. *Adv. Protein Chem.* **47**: 83–229.
- Ptitsyn, O.B. and Semisotnov, G.V. 1991. The mechanism of protein folding. In *Conformations and forces in protein folding* (eds. B.T. Nall and K.A. Dill), pp. 155–168. American Association for the Advancement of Science, Washington, DC.
- Rechsteiner, M., ed. 1988. *Ubiquitin*. Plenum Press, New York.
- Refield, C. 2004. NMR studies of partially folded molten-globule states. *Methods Mol. Biol.* **278**: 233–254.
- Richarz, R. and Wuethrich, K. 1978. Carbon-13 NMR chemical shifts of the common amino acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers* **17**: 2133–2141.
- Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comp. Phys.* **23**: 327–341.
- Sanchez, I.E. and Kiefhaber, T. 2003. Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**: 367–376.
- Schneider, D.M., Dellwo, M.J., and Wand, A.J. 1992. Fast internal main-chain dynamics of human ubiquitin. *Biochemistry* **31**: 3645–3652.
- Schuler, L.D., Daura, X., and van Gunsteren, W.F. 2001. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* **22**: 1205–1218.
- Scott, W.R.P., Hünenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennen, J., Torda, A.E., Huber, T., Krueger, P., and van Gunsteren, W.F. 1999. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103**: 3596–3607.
- Shiraki, K., Nishikawa, K., and Goto, Y. 1995. Trifluoroethanol-induced stabilization of the α -helical structure of β -lactoglobulin: Implication for non-hierarchical protein folding. *J. Mol. Biol.* **245**: 180–194.
- Soares, T.A., Daura, X., Oostenbrink, C., Smith, L.J., and van Gunsteren, W.F. 2004. Validation of the GROMOS force-field parameter set 45Alfa3 against nuclear magnetic resonance data of hen egg lysozyme. *J. Biomol. NMR* **30**: 407–422.
- Soares, T.A., Hünenberger, P.H., Kastenzholz, M.A., Kraeutler, V., Lenz, T., Lins, R.D., Oostenbrink, C., and van Gunsteren, W.F. 2005. An improved nucleic acid parameter set for the GROMOS force field. *J. Comput. Chem.* **26**: 725–737.
- Stockman, B.J., Euvrard, A., and Scahill, T.A. 1993. Heteronuclear three-dimensional NMR spectroscopy of a partially denatured protein: The A-state of human ubiquitin. *J. Biomol. NMR* **3**: 285–296.
- Thomas, S.T. and Makhataдзе, G.I. 2000. Contribution of the 30/36 hydrophobic contact at the C-terminus of the α -helix to the stability of the ubiquitin molecule. *Biochemistry* **39**: 10275–10283.
- Tjandra, N., Feller, S.E., Pastor, R.W., and Bax, A. 1995. Rotational diffusion anisotropy of human ubiquitin from ¹⁵N NMR relaxation. *J. Am. Chem. Soc.* **117**: 12562–12566.
- Tropp, J. 1980. Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: The effect of fluctuating internuclear distances. *J. Chem. Phys.* **72**: 6035–6043.
- van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P., and Tironi, I.G. 1996. *Biomolecular simulation: The GROMOS96 manual and user guide*. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland.
- Vijay-Kumar, S., Bugg, C.E., and Cook, W.J. 1987. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **194**: 531–544.
- Walser, R., Mark, A.E., van Gunsteren, W.F., Lauterbach, M., and Wipff, G. 2000. The effect of force-field parameters on properties of liquids: Parametrization of a simple three-site model for methanol. *J. Chem. Phys.* **112**: 10450–10459.
- Wand, A.J., Urbauer, J.L., McEvoy, R.P., and Bieber, R.J. 1996. Internal dynamics of human ubiquitin revealed by ¹³C-relaxation studies of randomly fractionally labeled protein. *Biochemistry* **35**: 6116–6125.
- Wang, A.C. and Bax, A. 1996. Determination of the backbone dihedral angles *j* in human ubiquitin from reparametrized empirical Karplus equations. *J. Am. Chem. Soc.* **118**: 2483–2494.
- Wilkinson, K.D. and Mayer, A.N. 1986. Alcohol-induced conformational changes of ubiquitin. *Arch. Biochem. Biophys.* **250**: 390–399.
- Wolynes, P.G. 2005. Energy landscapes and solved protein-folding problems. *Philos. Transact. Ser. A Math. Phys. Eng. Sci.* **363**: 453–464.
- Woolfson, D.N., Cooper, A., Harding, M.M., Williams, D.H., and Evans, P.A. 1993. Protein folding in the absence of the solvent ordering contribution to the hydrophobic interaction. *J. Mol. Biol.* **229**: 502–511.
- Wuethrich, K., Billeter, M., and Braun, W. 1983. Pseudostructures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**: 949–961.
- Xu, X.P. and Case, D.A. 2001. Automated prediction of ¹⁵N, ¹³C_α, ¹³C_β and ¹³C' chemical shifts in proteins using a density functional database. *J. Biomol. NMR* **21**: 321–333.