

Fluctuations in Ion Pairs and Their Stabilities in Proteins

Sandeep Kumar¹ and Ruth Nussinov^{2,3*}

¹Laboratory of Experimental and Computational Biology, National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, Maryland

²Intramural Research Support Program, SAIC Frederick, National Cancer Institute, Frederick Cancer Research and Development Center, Laboratory of Experimental and Computational Biology, Frederick, Maryland

³Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

ABSTRACT This report investigates the effect of systemic protein conformational flexibility on the contribution of ion pairs to protein stability. Toward this goal, we use all NMR conformer ensembles in the Protein Data Bank (1) that contain at least 40 conformers, (2) whose functional form is monomeric, (3) that are nonredundant, and (4) that are large enough. We find 11 proteins adhering to these criteria. Within these proteins, we identify 22 ion pairs that are close enough to be classified as salt bridges. These are identified in the high-resolution crystal structures of the respective proteins or in the minimized average structures (if the crystal structures are unavailable) or, if both are unavailable, in the “most representative” conformer of each of the ensembles. We next calculate the electrostatic contribution of each such ion pair in each of the conformers in the ensembles. This results in a comprehensive study of 1,201 ion pairs, which allows us to look for consistent trends in their electrostatic contributions to protein stability in large sets of conformers. We find that the contributions of ion pairs vary considerably among the conformers of each protein. The vast majority of the ion pairs interconvert between being stabilizing and destabilizing to the structure at least once in the ensembles. These fluctuations reflect the variabilities in the location of the ion pairing residues and in the geometric orientation of these residues, both with respect to each other, and with respect to other charged groups in the remainder of the protein. The higher crystallographic B-factors for the respective side-chains are consistent with these fluctuations. The major conclusion from this study is that salt bridges observed in crystal structure may break, and new salt bridges may be formed. Hence, the overall stabilizing (or, destabilizing) contribution of an ion pair is conformer population dependent. *Proteins* 2001;43:433–454.

© 2001 Wiley-Liss, Inc.

Key words: salt bridge; ion pair; electrostatics; conformer; NMR ensemble; protein stability

© 2001 WILEY-LISS, INC.

INTRODUCTION

Understanding protein flexibility is important for protein function and for rational drug design.¹ Protein flexibility can imply two distinct phenomena. The first is large-scale molecular motions, in which two (or more) parts of a protein move as rigid bodies with respect to one another. This type of motion is often observed in molecular events such as substrate/ligand binding or allostery. This large-scale flexibility can usually be traced to a small segment of the protein functioning as a hinge, as already shown in the pioneering work of McCammon and Go and colleagues.^{2,3} This can be referred to as segmental flexibility.⁴ The second type involves small-scale protein motions. Proteins exist in a range of conformations around their native states, with the population times changing with the environment.^{5–9} The small conformational variations among these represent the systemic flexibility of a protein molecule. In contrast with segmental flexibility, systemic flexibility arises from fluctuations in atomic coordinates through the entire protein molecule. These fluctuations involve changes in both hydrophobic and electrostatic interactions. Changes in electrostatic interactions, such as formation/breakage of salt bridges, affect protein stability.^{10–14} These may result in pK_a shifts in the charged residues and be important for enzyme catalysis.¹⁵ The goal of the present analysis is to investigate the effect of protein flexibility around the native state on protein electrostatic interactions. To carry out such an investigation, one needs a range of conformers, which provide an adequate sampling of the conformational space around the native state.

Grant sponsor: Binational Science Foundation (BSF), Israel; Grant number: 95-00208; Grant sponsor: Ministry of Science; Grant sponsor: Center of Excellence (Israel Academy of Sciences); Grant sponsor: Magnet; Grant sponsor: Tel Aviv University Basic Research and Adams Brain Center; Grant sponsor: National Cancer Institute, National Institutes of Health; Grant number: NO1-CO-56000.

*Correspondence to: Ruth Nussinov, Intramural Research Support Program, SAIC Frederick, National Cancer Institute, Frederick Cancer Research and Development Center, Bldg 469, Room 151, Frederick, MD 21702. E-mail: ruthn@ncifcrf.gov

Received 23 September 2000; Accepted 20 February 2001

CONFORMER ENSEMBLES: MOLECULAR DYNAMICS AND NMR

There are two ways to obtain ensembles of conformers, each with its advantages and disadvantages. The first is to generate the ensembles through molecular dynamics simulations. The number of conformers that can be obtained using molecular dynamics is, in principle, very large, covering a broad conformational space. The problem is that molecular dynamic simulations are currently limited to nanosecond time scales, and multiple long simulations are needed to obtain such a conformational sampling around the protein native state.¹⁶

The second way is to take already available nuclear magnetic resonance (NMR) conformer ensembles deposited in the Protein Data Bank (PDB).¹⁷ A typical protein NMR experiment yields an ensemble of conformers satisfying a list of nuclear Overhauser effect (nOe) distance restraints among hydrogen atoms close in space (1.8–5.0 Å¹⁸), as well as other constraints. Dynamic information on proteins is implicit both in these restraints and in regions in which they are absent. Hence, it is reasonable to expect that ensembles of NMR conformers will provide information regarding the inherent flexibility of a protein molecule. However, there are potential shortcomings in the quality of conformers, owing both to insufficient data and to structure calculation protocols. Furthermore, it is often difficult to separate the true dynamic behavior of the protein from artifacts, as a result of these shortcomings.

Nevertheless, a number of reports have appeared in the literature indicating that the conformational sampling is valid and that the quality of the NMR conformers is reasonable. MacArthur and Thornton¹⁹ compared protein structures derived from NMR data and X-ray crystallography. These investigators found that protein cores observed in NMR structures are well defined and compare well with those of X-ray structures with a resolution of 2.0–2.3 Å. However, they also observed that there is a greater disorder on the surface of NMR-derived protein structures. This may be due either to the inherent flexibility of the protein molecule in solution (as compared with the crystalline state), particularly of surface residues, or to fewer nOe restraints for surface residues. More recently, Philippopoulos and Lim²⁰ compared an ensemble of the *Escherichia coli* ribonuclease H1 (RNase H1) conformers derived from NMR experiments both with an ensemble derived from molecular dynamics simulations, and with two X-ray structures. These workers have shown that the NMR average structure is in better agreement with the high-resolution (1.48-Å) X-ray structure of the RNase H1, as compared with either the lower-resolution (2.05-Å) crystal structure or with molecular dynamics simulation. Furthermore, the 15 conformers of the NMR ensemble have sampled more conformational space of the RNase H1 than the 1.7-ns molecular dynamics simulations. Abseher et al.²¹ also showed that the space covered by NMR structure ensembles overlaps significantly with that of ensembles generated in long molecular dynamics simulations. A PubMed search on the Internet indicates an increasing number of reports showing agreement between results on

protein flexibility obtained by NMR heteronuclear relaxation experiments and molecular dynamic studies. A recent review²² on studies of protein dynamics by NMR has described the value of various NMR techniques to study protein backbone, as well as side-chain dynamics. Hence, we have used the NMR ensembles of conformers. Analysis of the available high-resolution crystal structures provides further corroborative evidence.

A relevant question in such an investigation is the extent to which an NMR ensemble is an ensemble in the statistical mechanical sense. Certainly, NMR structures cannot be treated as thermodynamic ensembles. The various NMR conformers are geometric entities that abide by stereochemical and experimental limitations, i.e., geometries obtained by optimizing a cost function consisting of some force field energy and experimental distance restraints. This function is not a Hamiltonian, and probabilities (population times) of the individual conformers do not follow a Boltzmann distribution. Further, the total number of conformers obtained in a typical NMR experiment is usually quite small (of the order of 10²–10³). Only a limited set of these conformers with a function value below some arbitrary threshold is deposited in the PDB (typically 10–50).

With respect to our analysis, despite the fact that NMR ensembles are not “true” ensembles in the thermodynamic sense, the sampling is broad, and covers considerable conformational space, enabling it to provide hints with regard to the relationship between conformer fluctuations and the stabilizing/destabilizing salt bridge (or, ion pair) contribution to the protein structure.

SALT BRIDGES AND ION PAIRS IN CONFORMER ENSEMBLES

The electrostatic interactions studied here are salt bridges and ion pairs. We define an ion pair as two oppositely charged (Asp and Glu versus Arg, Lys, and His) residues. An ion pair is described as a salt bridge if (1) the side-chain charged group centroids of the two oppositely charged residues are within a 4-Å distance, and (2) at least one pair of side-chain nitrogen and oxygen atoms in the two residues are within a distance of 4 Å.¹² Throughout this study, we have followed this convention to describe electrostatic interaction between two oppositely charged residues as an ion pair or a salt bridge.

Reports in the literature with regard to specific ion pairs, or limited sets of ion pairs, have clearly indicated that some are stabilizing, whereas others are destabilizing.^{10,12,23–27} To address the problem in a statistically meaningful way, and to examine the factors affecting ion pair stability, we have recently carried out an extensive analysis of salt bridges in a large dataset of structurally nonredundant high-resolution crystal structures of proteins whose functional form is monomeric.¹² This large-scale analysis has demonstrated that salt bridges can be stabilizing or destabilizing to the protein, depending on three factors: (1) the buried/exposed location of the ion pairing residues in the protein structure; (2) the distance and geometric orientation of the side-chain charged groups

with respect to each other; and (3) the interaction of the charged groups of the salt bridge with the charged groups in the remainder of the protein.¹² In particular, the analysis has indicated that most ion pairs obeying the geometric definition of a salt bridge¹² are stabilizing to the proteins. By contrast, those with distances exceeding the two 4.0-Å limits largely contribute to destabilize the protein structure. This observation of the sensitivity of the electrostatic contributions of ion pairs to variations in the geometry and in the location, i.e., the environment around the ion pairing residues, leads directly to questions relating to their stabilizing/destabilizing contributions in ensembles of conformers. Continuum electrostatic calculations are sensitive to conformational details of the protein molecule. Recently, we have carried out a detailed examination of the electrostatic interactions in one particular case, the c-Myc–Max leucine zipper.²⁸ In the 40 NMR conformers of this heterodimer, each of the ion pairs, and the intra- and intersubunit network between the monomeric α -helices, fluctuate between being stabilizing and being destabilizing. These results have prompted the current large-scale study.

The present study presents an analysis of electrostatic interactions in an ensemble of NMR conformers of different proteins. NMR experiments on these proteins have yielded ensembles consisting of at least 40 conformers. The biochemically functional state of these proteins is monomeric, and all contain ≥ 50 residues. Inspection of the database reveals that the number of proteins whose structures have been solved by NMR decreases rapidly with the number of conformers in the ensemble. Only a handful of cases display NMR ensembles containing at least 40 conformers. We have examined all of these, if they conform to our criteria (i.e., size, functionally monomeric, and nonredundant). We do not expect these NMR conformer ensembles to either sample the whole conformational space around the protein native state or represent all the conformers in solution. However, they may reflect, at least to some extent, the protein flexibility around the native state.

Using a continuum electrostatics-based methodology,¹⁰ we have computed the electrostatic contributions of 22 ion pairs in ensembles of 11 proteins. Our procedure is first to identify the salt bridges in the crystal structure (if available) and in the energy minimized average structure. If neither of these is available, we select the “most representative” conformer as described below. The charged residue pairs in these 22 salt bridges constitute the 22 ion pairs in our study. For each of these, we have calculated the stabilizing/destabilizing contribution in each of the 40 or more conformers. A total of 1,201 sets of such calculations have been carried out. We have also monitored the environment and the relative orientations of the charged residues in these pairs across the conformer ensembles. Our results are consistent for all the cases we have examined and indicate distinct trends, with the strengths of the electrostatic interactions between the ion pairing residues varying considerably in the different conformers. In the vast majority of the cases, the fluctuations make these interac-

tions flip-flop between being stabilizing and being destabilizing. In the individual conformers, the residues forming these 22 ion pairs frequently move out of, and come back to be within salt bridging distances. We have also examined the complementary situation, i.e., in each of the conformers we searched for the charged residue pairs that form salt bridges in the conformers, but not in the representative (crystal, minimized, or most representative conformer) of the set. As expected, owing to protein dynamics, in the individual conformers different pairs of charged residues, which are farther apart in the representative conformers, frequently get close enough to form salt bridges. In many other conformers, no salt bridges are formed.

For several of these 11 proteins, high-resolution X-ray crystal structures and/or minimized average structures are also available in the PDB. Two high-resolution crystal structures are available for CheY and the B1 domain of streptococcal protein G. The electrostatic stabilities as well as the geometries of the ion pairs vary in these crystal structures as well. The available crystallographic B-factor values for the side-chain atoms in the charged residues constituting the ion pairs in CheY, B1 domain of protein G, and Cyanovirin-N indicate their larger mobilities. For the B1 domain of protein G, two NMR conformer ensembles are available in the PDB. The fluctuations in ion pair geometry and electrostatic stability in the two ensembles show similar behavior.

MATERIALS AND METHODS

Definition of Ion Pair and Salt Bridge

A pair of oppositely charged amino acids Asp or Glu with Arg, Lys, and His is considered an ion pair. The geometry of an ion pair refers to spatial position of the side-chain functional groups, in the ion pairing residues, with respect to each other. It is characterized in terms of two quantities: (1) the distance, r , between the centroids of the side-chain functional groups in the two oppositely charged residues; and (2) the angular orientation, θ , of the side-chain charged groups in the two ion pairing residues computed as the angle between two unit vectors, with each unit vector joining a C $^{\alpha}$ atom and the side-chain charged group centroid in a charged residue.¹²

Two ion pairing residues are considered to form a salt bridge when (1) the side-chain functional group centroids lie within 4.0 Å of each other, and (2) at least one pair of Asp/Glu side-chain carboxyl oxygen and Arg/Lys/His side-chain nitrogen atoms are within 4.0-Å distance.¹² The locations of residues forming ion pairs are characterized in terms of their solvent accessible surface areas (ASA).^{29,30} The solvent probe radius is 1.4 Å.

Database of Proteins and Ion Pairs

Conformational ensembles of only a few proteins contain large numbers of conformers. The number of NMR solved protein structures decreases rapidly with the number of conformers in the ensemble. In October 1999, only 26 NMR structure ensembles contained at least 40 conformers in the PDB; 24 of these ensembles are either for monomeric proteins or for single protein domains and contain at least

TABLE I. Database Composition in Present Study

Protein name (no. of residues)	PDB entry containing the ensemble	No. of conformers	Representative conformer ^a	Ion pairs(s) ^b	References
β -Spectrin PH domain (106)	1MPH	50	Conformer 12	D40–K42 K91–E95	45
CheY (128)	1CEY	46	1CHN, 3CHY (X-tal)	D12–K109 D57–K109 D41–K45	33, 34, 59
c-Myb DNA-binding domain Repeat 1 (R1)(52)	1MBF	50	1MBE (AV)	E47–R73 D48–R81 E49–K52	37
Repeat 2 (R2) (52)	1MBH	50	1MBG (AV)	K92–E99 D100–R133	
Repeat 3 (R3) (52)	1MBK	50	1MBJ (AV)	E150–R153 E151–R176	
CRIP (76)	1IML	48	Conformer 16	K2–D7	46
CSE-I (64)	1B3C	40	2B3C (AV)	K13–E27	38
Cyanovirin-N (101)	2EZN	40	2EZM (AV) 3EZM (X-tal)	E68–K84	35, 39
Horse heart cytochrome c (reduced form) (104)	2GIW	40	1GIW (AV)	E61–K99	40
HMG1 Box 2 (79)	1HSN	49	1HSM (AV)	K62–E66	41
ISL-1 (66)	1BW5	50	Conformer 6	K54–D58	47
B1 domain of protein G (56)	1GB1 3GB1	60 32	2GB1 (AV) 1PGA, 1PGB (X-tal)	K4–E15 E27–K28 D47–K50	36, 42, 43
U1 SNRPA (116)	1FHT	43	Conformer 36	E108–R110	48

X-tal, PDB entry(ies) contain the crystal structure(s) of the protein; AV, the PDB entry contains the minimized average structure calculated from the NMR conformer ensemble of the protein.

^aThe “most representative” conformer is indicated by the conformer number in the NMR conformer ensemble.

^bAn ion pair is indicated by the constituent residue names and numbers. Residue names are given in the single letter code.

50 residues. Fourteen (out of 24) NMR ensembles for 11 different proteins contain at least one ion pair with the charged residue side-chain functional groups close enough to be classified as a salt bridge, either in their minimized average structures or in the “most representative” conformers, or in the corresponding high-resolution crystal structures (if available). These 11 proteins contain a total of 22 ion pairs (salt bridges). Table I lists the 11 proteins and 22 ion pairs we have studied.

The 11 proteins differ from each other. They do not share sequence or structural homologies. For the *E. coli* chemotaxis protein CheY, two high-resolution crystal structures (1CHN and 3CHY) are available. For the B1 domain of protein G, there are two sets of NMR ensembles (\langle 1GB1 \rangle and \langle 3GB1 \rangle). \langle 1GB1 \rangle contains 60 conformers, while \langle 3GB1 \rangle contains 32. Two high-resolution crystal structures (1PGA and 1PGB) are also available

for this protein. All have been used here. The choice of 40 as the minimum number of conformers in an ensemble is arbitrary. Our goal is to have large conformer ensembles in as many proteins as feasible; 40 conformers is a compromise, considering the available NMR data. Table II summarizes the NMR experiments performed to obtain 14 NMR conformer ensembles of these 11 proteins.

Computation of the Electrostatic Energies of Ion Pairs

The electrostatic stability of each of the 22 ion pairs in the NMR conformer ensembles, in the minimized average and in the crystal structures of the 11 proteins, is calculated using the method of computer mutations of the charged residue side-chains to their hydrophobic isosteres.¹⁰ Hydrophobic isosteres are the charged residue

TABLE II. Summary of NMR Experimental Studies on Proteins in Our Database*

Protein name	Conformer ensemble name and software	Spectrometer field strength (MHz)	Probes	Total restraints (nOe)	RMSD ^a (Å)
β -Spectrin PH domain ^b	(1MPH)	600	¹ H	563	1.00 \pm 0.19 (BB)
	X-PLOR				1.61 \pm 0.19 (H)
CheY ^c	(1CEY)	600	¹ H ¹³ C ¹⁵ N	1,032	1.34 (BB)
	FELIX				
c-Myb DNA-binding domain repeats R1, R2, and R3; 52 residues each ^d	(1MBF)	500	¹ H ¹⁵ N	640	0.44 \pm 0.11 (BB) R1
	(1MBH)			924	0.38 \pm 0.08 (BB) R2
	(1MBK) EMBOSS			536	0.40 \pm 0.10 (BB) R3 1.00 \pm 0.17 (H) R1 0.92 \pm 0.16 (H) R2 0.99 \pm 0.17 (H) R3 0.54 \pm 0.14 (BB)
CRIP ^e	(1IML)	600	¹ H ¹⁵ N	1,186	0.54 \pm 0.14 (BB)
	FELIX				
CSE-I	(1B3C)	600	¹ H	877	1.10 (BB)
	X-PLOR				
Cyanovirin-N	(2EZN)	500	¹ H	2,509	0.15 \pm 0.02 (BB)
	CNS	600	¹⁵ N		0.45 \pm 0.03 (H)
			750		¹³ C
Horse heart cytochrome c (reduced form)	(2GIW)	800	¹ H	2,620	0.67 \pm 0.10 (BB)
	AMBER	600		1.16 \pm 0.10 (H)	
	DYANA				
HMG1 Box 2	(1HSN)	600	¹ H	1,228	0.625 (BB)
	FELIXII	500	¹⁵ N		1.07 (H)
ISL-1	(1BW5)	600	¹ H ¹⁵ N ¹³ C	1,305	0.83 (BB)
	X-PLOR				1.48 (H)
B1 domain of protein G ^f	(1GB1)	600	¹ H	914	0.27 \pm 0.03 (BB)
	(3GB1)				0.65 \pm 0.05 (A)
	DISGEO				
U1 SNRP A	(1FHT)	500	¹ H	1,995	0.65 \pm 0.13 (BB)
	X-PLOR	600	¹⁵ N ¹³ C		1.16 \pm 0.16 (H)

*Magnetic field strengths of spectrometers on which the spectra were recorded, probes used in the experiments, computer programs used for structure calculation and refinement, total number of nuclear Overhauser effect (nOe) restraints used, and the precision of the resulting conformer coordinates with respect to the average conformer are shown. In all cases, several additional restraints have been used along with the nOe restraints for structure calculation and refinement.

^aRMSD, root-mean-square deviation (in the atomic coordinates); BB, backbone atoms; H, heavy (nonhydrogen) atoms; A, all atoms have been involved in the deviation calculations.

^bAutomated NOESY interpretation method with ambiguous distance constraints.

^cThe authors have produced ¹³C/¹⁵N-labeled CheY. ¹⁵N relaxation has been used to monitor protein dynamics.

^dThe authors have also solved the structures of whole DNA-binding domain R1R2R3 and minimum DNA binding domain R2R3.

^eThe authors have produced ¹⁵N-labeled CRIP.

^fTwo sets of conformer ensembles are available from the same group. Refer to text for details.

side-chains, with their partial atomic charges set to zero. This method has frequently been used in the literature.^{4,11–13,25,27,28} Predictions based on this method have been consistent with experiments.³¹ The total electrostatic free energy contribution $\Delta\Delta G_{\text{tot}}$ of an ion pair is the sum of three free energy terms: desolvation energy penalty $\Delta\Delta G_{\text{dsolv}}$, bridge energy term $\Delta\Delta G_{\text{brd}}$, and protein energy term $\Delta\Delta G_{\text{prt}}$ ($\Delta\Delta G$ is Gibbs free energy). $\Delta\Delta G_{\text{dsolv}}$ is the energy penalty incurred by ion pairing residues due to the

desolvation of the charged side-chains in the folded state of the protein with respect to the unfolded state. $\Delta\Delta G_{\text{brd}}$ represents the electrostatic interaction between the charged groups in ion pairing residue side-chains. $\Delta\Delta G_{\text{prt}}$ reflects the electrostatic interactions between the charged groups of the ion pair and those in the remainder of the protein. We have also computed another thermodynamic quantity called association energy, $\Delta\Delta G_{\text{assoc}}$, which represents the stability of the ion pair without taking into account its

electrostatic interaction with the other polar groups in its environment.

A detailed protocol for continuum electrostatic calculations has been described previously.^{12,27} We have used this protocol except with a finer grid spacing (0.5 Å per grid step). For the crystal structures, we have fixed the hydrogen atoms using the BIOPOLYMER module of INSIGHT II. We have energy minimized the crystal structure keeping the nonhydrogen atom positions fixed. The minimization process consisted of 100 steps of steepest descent followed by 500 steps of conjugate gradient. Energy minimization is carried out using the CFF91 force field in the DISCOVER module of INSIGHT II. This procedure improves the accuracy of the continuum electrostatic calculations.³² The calculations are carried out for all ion pairs listed in Table I, in the NMR conformer ensembles, minimized average structures and crystal structures of the 11 proteins. All calculations have been carried out at pH 7.0 and at zero ionic strengths. The electrostatic contribution of a triad of charged residues, Asp 12, Asp 57, and Lys 109 in CheY was computed in a similar manner.

RESULTS

The proteins investigated in this report are enumerated in Table I. The PDB file names, the number of conformers available for each protein, and which representative conformer was used for the initial identification of the salt bridges are listed. Our first choice was the crystal structure. If unavailable, we used the NMR energy minimized average structure. In cases in which neither were available in the PDB, we used the “most representative” one. Table II summarizes the NMR experiments and the quality of the data obtained from these experiments.

High-resolution crystal structures are available for three proteins in our database. These proteins are bacterial chemotaxis protein CheY (1CHN, 1.76 Å and 3CHY, 1.66 Å resolution),^{33,34} Cyanovirin-N (3EZM, 1.5 Å resolution),³⁵ and the B1 domain of streptococcal protein G (1PGA, 2.07 Å and 1PGB, 1.92-Å resolution).³⁶ The minimized average structures are available for six proteins in our database. These proteins are c-Myb DNA binding domain repeats R1, R2, and R3³⁷, CSE-I,³⁸ Cyanovirin-N,³⁹ Horse heart cytochrome c (reduced form),⁴⁰ and HMG1 Box 2⁴¹ and the B1 domain of streptococcal protein G.^{42,43} Both crystal and energy minimized average structures are available for Cyanovirin-N and for the B1 domain of protein G. For the protein G B1 domain, the two available sets of NMR conformer ensembles are used (Table I). The “most representative” conformer for an NMR ensemble is identified by the NMRCLUST procedure⁴⁴ in OLDERADO (On Line Database of Ensemble Representatives and Domains) available at the European Bioinformatics Institute (EBI) internet site. The “most representative” conformers are used for four proteins, β-spectrin pleckstrin homology domain,⁴⁵ CRIP,⁴⁶ ISL-1,⁴⁷ and U1 small nuclear ribonucleoprotein A.⁴⁸ Neither crystal structures nor minimized average structures are available for these four proteins.

The DNA binding domain of c-Myb is composed of three tandem repeats designated as R1, R2 and R3. Each repeat has its own NMR conformer ensemble. We have used all three. Homologies between these repeats vary between 31% and 46%. The three repeats have similar overall architectures. In total, these three repeats contain seven ion pairs. Structural alignment of the minimized average structures of the three c-Myb repeats R1, R2, and R3 shows that ion pairs E47–R73 in c-Myb R1 and E151–R176 in c-Myb R3 occupy equivalent positions in space. Ion pairs D48–R81 in c-Myb R1 and D100–R123 in c-Myb R2 are also spatially equivalent. Hence, out of 22 ion pairs in our databases, 20 are nonequivalent.

All the ion pairs used in this analysis are listed in Table I. Throughout this report, we use the bracketed PDB entry name ($\langle \cdot \cdot \cdot \rangle$) to denote an ensemble. For example, $\langle 1MPH \rangle$ denotes the ensemble of 50 NMR conformers of β-spectrin pleckstrin homology domain present in the PDB entry 1MPH. Our database contains a total of 14 NMR conformer ensembles and five high-resolution crystal structures.

Movement of Charged Residues

The average ion pair geometries in the NMR conformer ensembles are given in Table III. For comparison, the geometries of these ion pairs in the crystal structures, the minimized average structures and the “most representative” conformers are also given in Table III. Inspection of the data shows that our database has sampled a broad spectrum of ion pair movements. For example, the charged residues in the equivalent ion pairs, D48–R81 in the ensemble of c-Myb R1 ($\langle 1MBF \rangle$) and D100–R133 in the ensemble of c-Myb R2 ($\langle 1MBH \rangle$), tend to remain at almost constant distances. On the other hand, the charged residues in ion pairs K49–K52 in c-Myb R1 ($\langle 1MBF \rangle$), K92–E99 in c-Myb R2 ($\langle 1MBH \rangle$), E150–R153 in c-Myb R3 ($\langle 1MBK \rangle$) and E108–R110 in U1 small nuclear ribonucleoprotein A ($\langle 1FHT \rangle$) show extensive movements. The side-chain centroid distances between these ions show standard deviations as large as 2.6–3.6 Å about their mean values (Table III; see also Table IV). Charged residues in the remaining ion pairs show movements to different extents between these two extremes.

Figure 1 illustrates two examples of the fluctuations in ion pair geometries, E47–R73 in c-Myb R1 ($\langle 1MBF \rangle$) and D100–R143 in c-Myb R2 ($\langle 1MBH \rangle$). The side-chain charged group centroids of the ion pairing residues E47–R73 in c-Myb R1 often get close to within the salt bridging distance, and then move apart. Table V lists the number of conformers in which the ion pairing residues in our database are close enough to be classified as salt bridges. It can be seen that the 22 ion pairs in our database that form salt bridges in the representative conformers more often move apart than stay together in the conformers. Only 6 ion pairs survive as salt bridges in the majority of the conformers of their respective ensembles. These are K91–E95 in $\langle 1MPH \rangle$ (β-spectrin), E47–R73 and D48–R81 in $\langle 1MBF \rangle$ (c-Myb R1), D100–R133 in $\langle 1MBH \rangle$ (c-Myb R2), E151–R176 in $\langle 1MBK \rangle$ (c-Myb R3), and E61–K99 in $\langle 2G1W \rangle$

TABLE III. Geometries of 22 Ion Pairs in Our Database*

Ion pair ^a	Geometry [r (Å), θ (°)]			
	Ensemble average ^b	Crystal structure	Minimized average structure	MR conformer ^c
D40–K42	$\langle 1MPH \rangle$ 5.40 ± 1.49, 129.4 ± 20.0			3.34, 113.3
K91–E95	$\langle 1MPH \rangle$ 3.49 ± 0.92, 107.9 ± 15.6			2.70, 114.8
D12–K109	$\langle 1CEY \rangle$ 7.80 ± 1.84, 49.0 ± 21.0	1CHN 3.22, 54.3 3CHY 4.66, 56.8		
D57–K109	$\langle 1CEY \rangle$ 6.39 ± 1.55, 128.5 ± 24.4	1CHN 3.78, 117.7 3CHY 3.20, 90.1		
D41–K45	$\langle 1CEY \rangle$ 5.86 ± 1.27, 153.1 ± 14.3	1CHN 3.72, 138.9 3CHY 4.56, 150.6		
E47–R73	$\langle 1MBF \rangle$ 3.97 ± 0.74, 66.1 ± 14.6		1MBE 3.55, 76.3	
D48–R81	$\langle 1MBF \rangle$ 3.57 ± 0.03, 104.4 ± 4.9		1MBE 3.56, 105.8	
E49–K52	$\langle 1MBF \rangle$ 6.58 ± 2.56, 127.7 ± 28.2		1MBE 2.68, 156.2	
K92–E99	$\langle 1MBH \rangle$ 8.62 ± 3.61, 68.86 ± 24.52		1MBG 3.12, 5.3	
D100–R133	$\langle 1MBH \rangle$ 3.44 ± 0.05, 96.3 ± 3.9		1MBG 3.45, 93.3	
E150–R153	$\langle 1MBK \rangle$ 6.09 ± 2.63, 116.7 ± 31.9		1MBJ 3.54, 113.8	
E151–R176	$\langle 1MBK \rangle$ 3.91 ± 0.69, 82.15 ± 16.6		1MBJ 3.56, 87.8	
K2–D7	$\langle 1IML \rangle$ 7.81 ± 2.49, 130.2 ± 30.0			3.99, 135.4
K13–E27	$\langle 1B3C \rangle$ 4.65 ± 1.36, 73.2 ± 27.1		2B3C 2.82, 64.5	
E68–K84	$\langle 2EZN \rangle$ 5.90 ± 1.25, 147.0 ± 20.3	3EZM 3.94, 118.1	2EZM 5.17, 156.0	
E61–K99	$\langle 2GIW \rangle$ 4.12 ± 1.44, 112.4 ± 19.3		1GIW 2.73, 114.4	
K62–E66	$\langle 1HSN \rangle$ 6.11 ± 1.44, 150.0 ± 14.0		1HSM 3.12, 133.5	
K54–D58	$\langle 1BW5 \rangle$ 5.22 ± 1.32, 113.6 ± 26.1			3.37, 115.4
K4–E15	$\langle 1GB1 \rangle$ 5.51 ± 1.58, 149.7 ± 14.3	1PGA 3.28, 134.5	2GB1 4.95, 156.0	
K4–E15	$\langle 3GB1 \rangle$ 5.43 ± 1.83, 134.5 ± 19.8	1PGB 2.59, 119.2		
E27–K28	$\langle 1GB1 \rangle$ 9.46 ± 1.58, 97.3 ± 24.5	1PGA 3.65, 134.9	2GB1 9.78, 98.2	
E27–K28	$\langle 3GB1 \rangle$ 8.92 ± 1.45, 108.9 ± 20.4	1PGB 8.87, 92.1		
D47–K50	$\langle 1GB1 \rangle$ 7.43 ± 1.48, 139.2 ± 18.0	1PGA 3.47, 168.9	2GB1 7.82, 140.8	
D47–K50	$\langle 3GB1 \rangle$ 6.49 ± 1.58, 144.1 ± 20.8	1PGB 4.00, 138.2		
E108–R110	$\langle 1FHT \rangle$ 10.36 ± 2.86, 85.8 ± 47.6			3.86, 145.4

*Ion pair geometries are represented by two parameters, r (Å) and θ (°), defined in Materials and Methods. See Table IV and text for further details.

^aAn ion pair is indicated by the constituent residue names and numbers.

^bThe ensemble averages presented do not take into account the relative conformer populations. Hence, these values may not reflect the ion pair geometry and location in solution. Ensemble average denotes the average ion pair geometries in the NMR conformer ensemble. Standard deviations for the average values are also given.

^cMR conformer stands for the “most representative” conformer in the NMR conformer ensemble (given in Table I).

(Horse heart cytochrome c). The remaining 16 ion pairs survive as salt bridges only in a minority of the conformers. The ion pair E27–K28, which forms a salt bridge in one crystal structure (1PGA), does not survive in any conformer in either of the two NMR ensembles of the B1 domain of protein G. Still in the same protein, ion pair D47–K50 which forms a salt bridge in both crystal structures (1PGA and 1PGB) survives in just one conformer in each of the NMR ensembles $\langle 1GB1 \rangle$ and $\langle 3GB1 \rangle$. A similar behavior is shown by ion pairs D12–K109 and D57–K109 in $\langle 1CEY \rangle$ (CheY), K2–D7 in $\langle 1IML \rangle$ (CRIP), and E108–R110 in $\langle 1FHT \rangle$ (U1 small nuclear ribonucleoprotein A) (Table V). All the ensembles, except $\langle 1MPH \rangle$, $\langle 1MBF \rangle$, and $\langle 1MBH \rangle$, contain conformers in which no salt bridges are formed. In the ensembles $\langle 1IML \rangle$, $\langle 1B3C \rangle$, $\langle 2EZN \rangle$, $\langle 1HSN \rangle$, $\langle 1BW5 \rangle$, $\langle 1GB1 \rangle$, $\langle 3GB1 \rangle$, and $\langle 1FHT \rangle$, most of the conformers do not contain any salt bridges (Table V).

As expected, we also see the opposite trends: Pairs of oppositely charged residues (ion pairs) which are not close enough to form salt bridges in the representatives (crystal structure, energy minimized average structure or most

representative conformer) come close together to form salt bridges in other conformers in the ensembles. The conformer ensembles of β -spectrin, c-Myb R1, R2, and R3 and cytochrome c are particularly rich in such salt bridges. The majority of the conformers in the ensembles of $\langle 1MPH \rangle$, $\langle 1MBF \rangle$, $\langle 1MBH \rangle$, $\langle 1MBK \rangle$, and $\langle 2GIW \rangle$ contain several salt bridges formed by charged residues which are farther apart in their minimized average structures and in the “most representative” conformers (Table V). We denote such salt bridges as ‘alternative’ salt bridges. In some conformers they are formed in lieu of the salt bridges in the representative conformer; In most others, they are formed in addition. Figure 2a–c provides an example, in c-Myb R3. Figure 2a shows salt bridges E150–R153 and E151–R176 in the minimized average structure (1MBJ) of c-Myb R3, denoted “original” salt bridges. In four conformers, numbers 6, 13, 23, and 40, only these two original salt bridges are formed. In others, there is a spectrum of situations: in conformer 22, both original salt bridges are lost and no alternative salt bridges are formed. In five conformers, 11, 27, 28, 29, and 39, only one of the original salt bridges is

TABLE IV. Location of Charged Residues Forming 22 Ion Pairs*

Residue (conformation) ^a	Accessible surface area (ASA) (%)			
	Ensemble average ^b	Crystal structure	Minimized average structure	MR conformer
D40 (H)	(1MPH) 20.9 ± 5.0			18.8
K42 (H)	(1MPH) 77.6 ± 3.2			75.3
K91 (C)	(1MPH) 62.1 ± 10.0			50.4
E95 (H)	(1MPH) 17.8 ± 4.4			14.5
D12 (B)	(1CEY) 3.6 ± 4.0	1CHN 1.7, 3CHY 1.5		
D41 (H)	(1CEY) 21.9 ± 5.0	1CHN 21.1, 3CHY 23.2		
K45 (H)	(1CEY) 38.9 ± 6.4	1CHN 27.2, 3CHY 34.7		
D57 (B)	(1CEY) 8.3 ± 5.5	1CHN 2.1, 3CHY 1.4		
K109 (B)	(1CEY) 20.9 ± 8.3	1CHN 5.0, 3CHY 17.0		
E47 (H)	(1MBF) 13.3 ± 4.3		1MBE 21.2	
D48 (H)	(1MBF) 9.8 ± 4.0		1MBE 8.3	
E49 (H)	(1MBF) 36.8 ± 8.8		1MBE 32.1	
K52 (H)	(1MBF) 33.5 ± 7.2		1MBE 29.7	
R73 (C)	(1MBF) 11.3 ± 6.3		1MBE 5.6	
R81 (H)	(1MBF) 19.7 ± 5.2		1MBE 24.7	
K92 (H)	(1MBH) 58.2 ± 19.3		1MBG 45.2	
E99 (H)	(1MBH) 11.9 ± 6.3		1MBG 5.1	
D100 (H)	(1MBH) 4.8 ± 2.3		1MBG 2.6	
R133 (H)	(1MBH) 16.7 ± 3.4		1MBG 18.7	
E150 (H)	(1MBK) 62.7 ± 8.7		1MBJ 56.2	
E151 (H)	(1MBK) 17.1 ± 4.6		1MBJ 15.7	
R153 (H)	(1MBK) 52.3 ± 8.9		1MBJ 52.1	
R176 (C)	(1MBK) 23.2 ± 9.8		1MBJ 26.7	
K2 (B)	(1IML) 53.1 ± 8.2			40.6
D7 (B)	(1IML) 81.6 ± 7.2			77.5
K13 (C)	(1B3C) 26.3 ± 8.3		2B3C 21.9	
E27 (H)	(1B3C) 27.3 ± 9.9		2B3C 21.8	
E68 (B)	(2EZM) 37.9 ± 6.6	3EZM 16.7	2EZM 40.0	
K84 (B)	(2EZM) 66.1 ± 5.0	3EZM 67.3	2EZM 66.3	
E61 (H)	(2GIW) 45.9 ± 7.1		1GIW 46.0	
K99 (H)	(2GIW) 46.8 ± 5.3		1GIW 46.4	
K62 (H)	(1HSM) 47.1 ± 7.0		1HSM 38.5	
E66 (H)	(1HSM) 44.8 ± 5.4		1HSM 35.3	
K54 (H)	(1BW5) 13.4 ± 6.9			8.3
D58 (H)	(1BW5) 27.5 ± 12.1			22.3
K4 (B)	(1GB1) 40.8 ± 4.1	1PGA 29.4	2GB1 42.5	
K4 (B)	(3GB1) 36.9 ± 3.6	1PGB 28.5		
E15 (B)	(1GB1) 59.6 ± 4.1	1PGA 56.0	2GB1 58.8	
E15 (B)	(3GB1) 61.5 ± 3.3	1PGB 57.0		
E27 (H)	(1GB1) 45.3 ± 2.6	1PGA 28.4	2GB1 45.8	
E27 (H)	(3GB1) 47.1 ± 5.4	1PGB 34.5		
K28 (H)	(1GB1) 64.4 ± 3.9	1PGA 54.1	2GB1 65.1	
K28 (H)	(3GB1) 67.5 ± 5.6	1PGB 58.3		
D47 (B)	(1GB1) 76.8 ± 3.7	1PGA 53.8	2GB1 78.3	
D47 (B)	(3GB1) 70.9 ± 4.0	1PGB 58.2		
K50 (B)	(1GB1) 50.2 ± 7.2	1PGA 27.4	2GB1 54.8	
K50 (B)	(3GB1) 30.0 ± 3.9	1PGB 28.4		
E108 (C)	(1FHT) 70.2 ± 16.3			48.7
R110 (C)	(1FHT) 81.9 ± 13.9			52.6

^aThe backbone conformations of the residues are given in parentheses: H, B, and C denote α -helix, β -strand and coil conformations, respectively.

^bThe location of the charged residues constituting the 22 ion pairs in the protein is given in terms of percentage accessible surface area (ASA), as well as the average ASA of the residue in the NMR conformer ensemble. The ensemble averages presented do not take into account the relative conformer populations. Hence, these values may not reflect the ion pair geometry and location in solution.

preserved, and no alternative salt bridges are formed. In 12 conformers, 9, 15, 16, 24, 30, 31, 32, 35, 44, 48, 49, and 50, both original salt bridges are preserved, and at least one additional alternative salt bridge is formed. Figure 2b shows conformer 30 in the ensemble (1MBK), as an

example. Three alternative salt bridges (E149–R153, E168–K171, and D178–K182) are formed in addition to the original ones. In the remaining 28 conformers, one or both of the original salt bridges are lost and at least one alternative salt bridge is formed. Figure 2c shows an

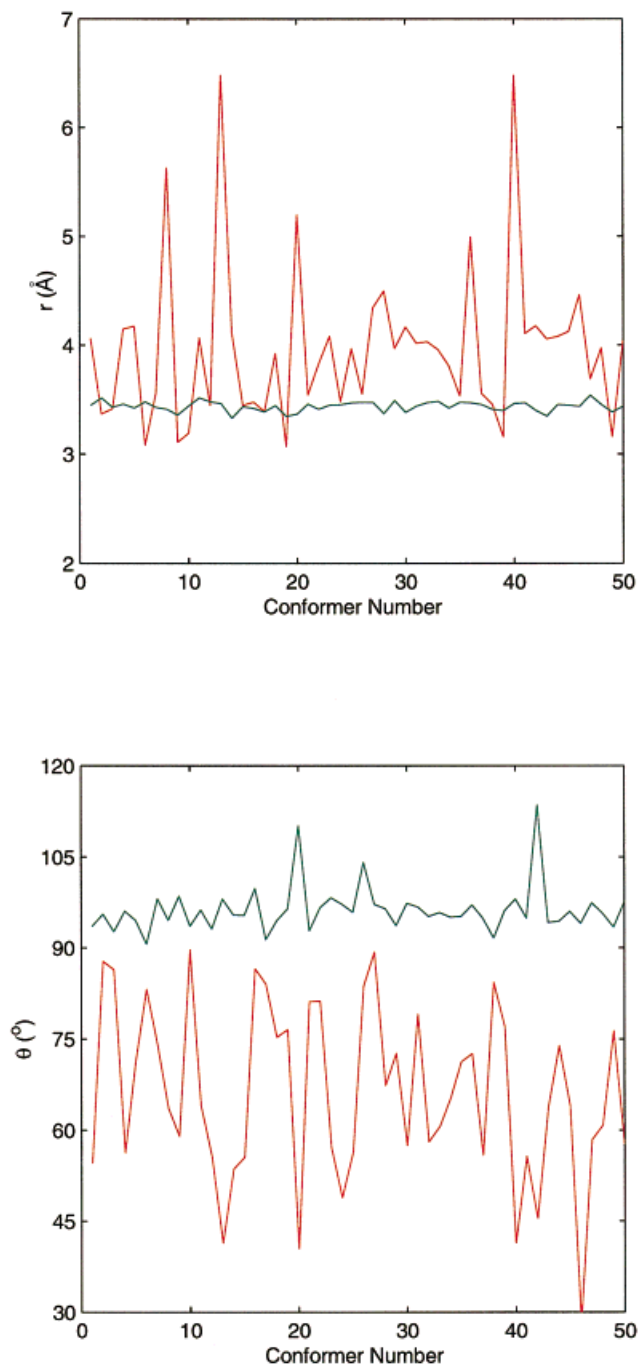


Fig. 1. Example of conformer-dependent fluctuations of ion pair geometries in NMR ensembles. Two ion pairs are chosen to illustrate our observations. The ion pair E47–R73 of c-Myb R1 is shown in red, while the ion pair D100–R133 of c-Myb R2 is shown in green. The plots show the variations in the geometric parameters r (Å) and θ ($^{\circ}$) for the ion pairs in different NMR conformers. The ion pair geometry parameters are defined in Materials and Methods. The ion pair E47–R73 (red) shows extensive fluctuations in its geometry. Most of the ion pairs in our database show similar behavior. The ion pair D100–R133 does not show much fluctuation in its geometry. However, such ion pairs are rare in our database.

example of conformer 1 in $\langle 1\text{MBK} \rangle$. In this conformer both original salt bridges are lost and two alternative ones (K143–E149 and D178–K182) are formed.

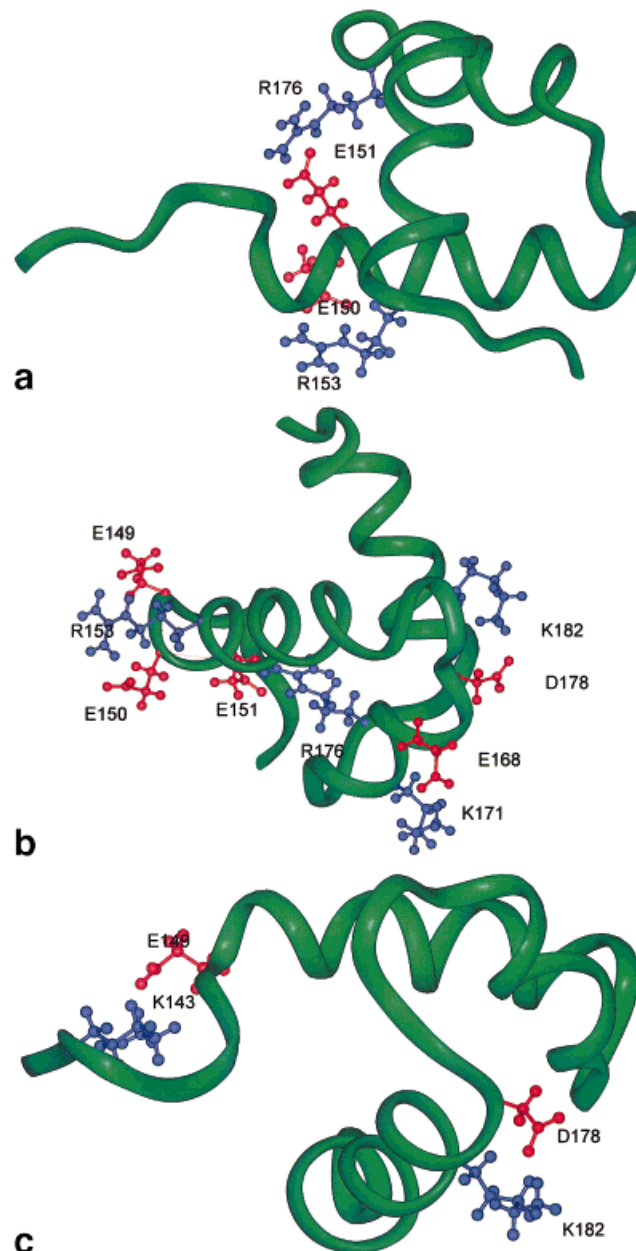


Fig. 2. Example illustrating the movement of charged residues with respect to each other in the NMR conformer ensemble. The protein used here is c-Myb DNA binding domain repeat 3 (R3). **a**: The average structure of the ensemble (PDB file 1MBJ) shows two salt bridges formed by ion pairs E150–R153 and E151–R176. Conformers 6, 13, 23, and 40 in the ensemble $\langle 1\text{MBK} \rangle$ contain only these two salt bridges. **b**: Three additional salt bridges—E149–R153, E168–K171, and D178–K182—are formed in conformer 30 along with E150–R153 and E151–R176. **c**: Conformer 1 in $\langle 1\text{MBK} \rangle$ contains two salt bridges K143–E149 and D178–K182, while the salt bridges E150–R153 and E151–R176 are broken because of the movement of the charged residues.

Ion pair geometries vary between the crystal structures as well. The crystal structures of CheY in the absence of Mg^{2+} (3CHY) and in its presence (1CHN) show variable geometries for ion pairs D12–K109 and D57–K109. Similarly, K4–E15, E27–K28, and D47–K50 differ in the orthorhombic and trigonal crystals of the B1 domain of

TABLE V. Movement of Ion Pairing Residues With Respect to Each Other*

Protein ensemble	Ion pair	Total no. of conformers	No. of conformers in which ion pairing residues form salt bridge	No. of conformers in which ion pairing residues do not form salt bridge	No. of conformers that do not contain any salt bridge	No. of conformers that contain other salt bridge(s)
⟨1MPH⟩	D40–K42	50	15	35	0	50
⟨1MPH⟩	K91–E95	50	40	10	—	—
⟨1CEY⟩	D12–K109	46	1	45	16	22
⟨1CEY⟩	D57–K109	46	3	43	—	—
⟨1CEY⟩	D41–K45	46	7	39	—	—
⟨1MBF⟩	E47–R73	50	27	23	0	50
⟨1MBF⟩	D48–R81	50	50	0	—	—
⟨1MBF⟩	E49–K52	50	16	34	—	—
⟨1MBH⟩	K92–E99	50	11	39	0	48
⟨1MBH⟩	D100–R133	50	50	0	—	—
⟨1MBK⟩	E150–R153	50	16	34	1	40
⟨1MBK⟩	E151–R176	50	36	14	—	—
⟨1IML⟩	K2–D7	48	2	46	31	17
⟨1B3C⟩	K13–E27	40	12	28	26	6
⟨2EZN⟩	E68–K84	40	5	35	23	14
⟨2G1W⟩	E61–K99	40	26	14	2	36
⟨1HSN⟩	K62–E66	49	6	43	27	19
⟨1BW5⟩	K54–D58	50	9	41	39	3
⟨1GB1⟩	K4–E15	60	12	48	44	5
⟨3GB1⟩	K4–E15	32	10	22	20	5
⟨1GB1⟩	E27–K28	60	0	60	—	—
⟨3GB1⟩	E27–K28	32	0	32	—	—
⟨1GB1⟩	D47–K50	60	1	59	—	—
⟨3GB1⟩	D47–K50	32	1	31	—	—
⟨1FHT⟩	E108–R110	43	2	41	24	17

*The charged residues move with respect to one another in the NMR conformer ensembles. In each conformer ensemble, the ion pairing residues that form salt bridges in the representative structure (crystal structure(s)/minimized average structure/most representative conformer) move away from one another in most of the conformers. Many conformers do not contain any salt bridge. In other conformers, new salt bridges are formed between the charged residues, which are farther apart in the representative conformer. These salt bridges are formed both in addition and instead of the salt bridges in the representative conformer. All these ion pairs listed form salt bridges in the representative conformer. An ion pair is indicated by the constituent residue names and numbers. The last two columns indicate the whole ensemble.

streptococcal protein G.³⁶ E27–K28 shows the largest change (Table III). E68–K84 residues in the cyanovirin-N (⟨2EZN⟩) ensemble and in the minimized average structure (2EZM) are farther apart than in the crystal structure (3EZM). This protein swaps domains under crystallization conditions.³⁵ The different geometries observed in the presence/absence of Mg^{2+} , the different crystal forms, and the domain swapping, further illustrate that the most highly populated conformers vary with the conditions, leading to population shifts. Because different salt bridges are formed in different conformers, what we observe in experiments in solution or in the crystalline state, is a reflection of the salt bridges present in the most highly populated states.

Additional evidence for the movement of the charged residues comes from their accessible surface areas (ASA).^{29,30} ASA points to the location of a residue in the protein. A small ASA value for a residue indicates that it is mostly buried in the protein core. A large ASA implies that it is at (or near) the surface. Table IV lists the ASA for the charged residues involved in the 22 ion pairs in the ensembles, crystal structures, the minimized average structures and in the “most representative” conformers. The locations of these residues in the proteins fluctuate. For

most of the charged residues, the standard deviations in the ASAs are ~15–60% of the mean ASA values in the conformer ensembles. We have computed only the arithmetic averages, since data on populations are not available in the conformer coordinates files. Hence, the average ion pair geometries and locations presented in the present report do not reflect the situation in solution.

Taken together, these observations indicate that the residues forming salt bridges in the representative conformers may move apart and get closer in different conformers of the ensembles. Other charged residues which did not form salt bridge(s) in the representative conformers can come together to form salt bridges in the other conformers.

Continuum Electrostatics Calculations

We have computed the electrostatic stabilities of each ion pair in all conformers. Each calculation yields the following energy terms: the total electrostatic contribution to the free energy of ion pair formation $\Delta\Delta G_{\text{tot}}$, the desolvation energy penalty ($\Delta\Delta G_{\text{dslv}}$) for the charged residues in the folded state of the protein, the bridge energy term ($\Delta\Delta G_{\text{brd}}$) for the electrostatic interaction among the charges within the ion pairing residues and the protein energy term ($\Delta\Delta G_{\text{prt}}$) for the electrostatic interaction of

charges in the ion pair with the charges in the remainder of the protein. These energy terms are related to one another by the following equation:

$$\Delta\Delta G_{tot} = \Delta\Delta G_{dslv} + \Delta\Delta G_{brd} + \Delta\Delta G_{prt}$$

We have also computed the association energy ($\Delta\Delta G_{assoc}$) for each ion pair. Table VI presents the average, minimum, and maximum values of these terms for each ion pair in the ensembles. The strengths of the ion pairs fluctuate extensively between conformers. The electrostatic contributions of 21 out of the 22 ion pairs in the crystal structures, minimized average structures and the “most representative” conformers fall within two standard deviations (SD) of the corresponding average values in the NMR ensembles. However, in most (13) of the ion pairs, they are within 1 SD. The average energy values for an ion pair are the mean of its values in the given ensemble.

In the following discussion, we describe the variability in the extent of the fluctuations in the stabilities of the ion pairs, with the DNA binding domain of c-Myb³⁷ as an example. Among the ion pairs in this domain, E47–R73 shows the largest fluctuations in its stability across the ensemble. On the other hand, ion pair D100–R133 fluctuates the least. Figure 3 shows the fluctuations in their energy terms. E47–R73 provides a typical example of the behavior of ion pairs in our database. All energy terms, the average ASA, the side-chain centroid distances between the charged residues and the relative orientations of the charged groups with respect to each other fluctuate extensively (Fig. 1).

Ion pair D100–R173 in c-Myb R2 does not fluctuate much (Fig. 3). Its equivalent D48–R81 in c-Myb R1 does not fluctuate either. Both ion pairs remain close enough to be classified as salt bridges in all 50 conformers of their respective ensembles, $\langle 1MBF \rangle$ and $\langle 1MBH \rangle$. These salt bridges are formed between two helices and are buried in the core. Residues D48 and D100 are less than 10% exposed (Table IV), while R81 and R133 have ASA of less than 20%. These residues show small fluctuations in their ASAs. In the minimized average structure of c-Myb R2 (1MBG) the salt bridge D100–R133 is surrounded by Trp 95, Thr 96, Lys 97, Gln 101, Val 103, Ile 104, Arg 125, Lys 128, Gln 129, Cys 130, Glu 132, Trp 134, His 135, Asn 136, His 137, Leu 138, and Pro 140. This indicates a fairly polar environment around this salt bridge, with 11 of the surrounding residues being either polar or charged. The residues that make significant contact with D100–R133 are Lys 97 (with Asp 100), Gln 129, Arg 132, Asn 136, and Glu 138 (with Arg 133). Trp 95 and Trp 134 also surround this salt bridge. Each repeat (R1, R2, and R3) in c-Myb DNA binding domain contains three Trp residues separated by 18 or 19 residues. Mutating these conserved Trp residues in R2 and R3 to alanine leads to a severe loss of DNA-binding activity.³⁷ The presence of these structural tryptophans along with the polar environment around the charged residues D100 and R133 may have suppressed the movement of these residues. The absence of fluctuations in geometries as well as the electrostatic stabilities of these ion pairs represent exceptions in our database.

An ion pair network consisting of charged residues Asp 12, Asp 57 and Lys 109 also shows extensive fluctuations in the NMR conformer ensemble of CheY (ICEY) and in the two high-resolution crystal structures for this protein (Table VI).

Table VII lists the number of conformers in which a given ion pair is stabilizing or destabilizing. E47–R73 in c-Myb R1 (Fig. 3) is stabilizing in 39 (out of 50) conformers and destabilizing in the remaining 11 conformers. 19 out of the 22 ion pairs in our database interconvert between being stabilizing and destabilizing at least once. One (E61–K99 in $\langle 2G1W \rangle$), the reduced form of horse heart cytochrome c) of the remaining three ion pairs that do not interconvert, shows extensive fluctuations nonetheless. The two ion pairs which do not show fluctuations are D48–R81 in $\langle 1MBF \rangle$ (c-Myb R1) and D100–R133 in $\langle 1MBH \rangle$ (c-Myb R2). These two ion pairs are equivalent. The reasons for the observed behavior of these ion pairs were discussed above.

In all 22 ion pairs, one of the two species, stabilizing or destabilizing, appears to be dominant. Table VII indicates that 12 ion pairs have stabilizing electrostatic contributions in most of the conformers. For the remaining 10, the electrostatic contribution is destabilizing in most of the conformers. This observation may explain why such interactions are reported to be either stabilizing or destabilizing by experiments and by calculations. Since data on the relative populations of each conformer in an NMR ensemble are unavailable, it is unclear whether the dominant species are also the largest contributor to the experimental/theoretical observation. In particular, the population times of the conformers vary with conditions.^{6,9} Furthermore, our analysis indicates that the overall electrostatic contribution of an ion pair is more likely to be stabilizing when the ion pairing residues are close enough to form salt bridges. This is due to stronger electrostatic interaction between the side-chain functional groups ($\Delta\Delta G_{brd}$) in the salt bridges, as compared with the situation when these groups are farther apart (data not shown). The electrostatic interaction between the charged residue side-chain functional groups is the main factor in neutralizing the unfavorable desolvation penalty ($\Delta\Delta G_{dslv}$) incurred by these residues in the folded state of the protein.¹²

DISCUSSION

The main goal of the present study is to analyze the sensitivity of electrostatic strengths of charge–charge interactions to the variations in the orientation of charges with respect to one another as well as their location in the protein. This study has evolved from our previous studies which indicated that the location of the charged residues and the geometries of their interaction are important factors in determining the electrostatic stabilities of salt bridges and ion pairs.^{12,27,28} In particular, analysis of intra- and interhelical ion pairs and their network in the c-Myc–Max leucine zipper NMR conformer ensemble indicated that their contributions toward the stability of the protein fluctuate between being stabilizing and destabiliz-

TABLE VI. Stabilities of 22 Ion Pairs in the 11 Proteins in Our Database*

Ion pair and protein ^a	Conformer ensemble ^b	$\Delta\Delta G_{dslv}$ (kcal/mol)	$\Delta\Delta G_{brd}$ (kcal/mol)	$\Delta\Delta G_{prt}$ (kcal/mol)	$\Delta\Delta G_{tot}$ (kcal/mol)	$\Delta\Delta G_{assoc}$ (kcal/mol)
D40–K42 in β -spectrin	$\langle 1MPH \rangle_{av}$	2.61 ± 0.84	-1.82 ± 2.10	-12.03 ± 3.01	-11.24 ± 3.12	-1.49 ± 1.87
	Min., Max.	1.59, 5.86	-6.42, -0.23	-18.80, -2.47	-16.21, +0.28	-5.53, +0.44
	Conformer 12	2.30	-4.53	-13.11	-15.35	-4.05
K91–E95 in β spectrin	$\langle 1MPH \rangle_{av}$	6.17 ± 2.70	-3.09 ± 1.61	-6.27 ± 2.96	-3.19 ± 2.03	-0.75 ± 0.66
	Min., Max.	2.04, 11.28	-6.67, -0.47	-13.59, +2.63	-7.45, +4.06	-2.09, +0.33
	Conformer 12	11.28	-6.40	-8.44	-3.55	-0.44
D12–K109 in CheY	$\langle 1CEY \rangle_{av}$	11.25 ± 3.29	-1.36 ± 1.08	-2.54 ± 4.62	$+7.36 \pm 4.06$	-0.38 ± 0.26
	Min., Max.	4.60, 16.96	-5.41, -0.37	-14.86, +5.22	-1.46, +15.08	-1.62, -0.13
	X-tal 1CHN	14.27	-10.15	-4.75	-0.62	-3.37
	X-tal 3CHY	10.85	-2.79	-8.25	-0.19	-0.80
D57–K109 in CheY	$\langle 1CEY \rangle_{av}$	8.36 ± 2.19	-1.59 ± 1.10	$+0.31 \pm 4.60$	$+7.08 \pm 4.63$	-0.58 ± 0.52
	Min., Max.	4.96, 14.26	-6.76, -0.46	-10.41, +10.97	-1.95, +19.41	-3.56, -0.19
	X-tal 1CHN	11.86	-4.63	-4.52	+2.72	-1.23
	X-tal 3CHY	8.77	-7.57	-4.25	-3.05	-3.59
D12–K109–D57 in CheY ^c	$\langle 1CEY \rangle_{av}$	15.46 ± 4.24	$+1.08 \pm 2.89$	-7.88 ± 4.94	$+8.66 \pm 5.92$	$+0.18 \pm 1.18$
	Min., Max.	7.36, 23.76	-5.04, +9.04	-16.88, +1.24	-2.68, +21.07	-3.49, +3.04
	X-tal 1CHN	17.51	-8.49	-8.83	+0.19	-4.36
	X-tal 3CHY	13.85	-6.02	-12.77	-4.94	-3.85
D41–K45 in CheY	$\langle 1CEY \rangle_{av}$	3.98 ± 1.35	-1.14 ± 0.92	-2.58 ± 2.10	$+0.26 \pm 2.01$	-0.62 ± 0.59
	Min., Max.	2.06, 10.68	-3.97, -0.35	-5.69, +3.80	-3.59, +6.51	-2.48, -0.18
	X-tal 1CHN	4.25	-3.10	-4.75	-3.60	-1.73
	X-tal 3CHY	4.12	-1.78	-4.89	-2.55	-0.89
E47–R73 in c-Myb R1	$\langle 1MBF \rangle_{av}$	8.69 ± 2.60	-6.59 ± 3.01	-5.54 ± 3.38	-3.44 ± 4.19	-3.76 ± 2.03
	Min., Max.	3.92, 14.94	-11.80, -1.60	-14.44, -0.87	-15.15, +4.73	-7.03, -0.34
	1MBE	7.11	-8.26	-6.90	-8.05	-5.63
D48–R81 in c-Myb R1	$\langle 1MBF \rangle_{av}$	9.56 ± 1.15	-11.40 ± 0.85	-2.27 ± 1.15	-4.11 ± 1.01	-5.79 ± 0.34
	Min., Max.	6.62, 12.86	-12.60, -6.98	-6.22, -0.94	-8.12, -0.96	-6.51, -4.55
	1MBE	8.95	-11.56	-1.73	-4.34	-6.16
E49–K52 in c-Myb R1	$\langle 1MBF \rangle_{av}$	3.55 ± 1.78	-1.60 ± 1.80	-5.70 ± 3.06	-3.74 ± 2.37	-1.25 ± 1.48
	Min., Max.	0.59, 7.91	-5.59, -0.11	-13.96, -0.50	-8.73, +1.84	-4.34, -0.08
	1MBE	1.28	-5.04	-2.57	-6.34	-4.54
K92–E99 in c-Myb R2	$\langle 1MBH \rangle_{av}$	8.03 ± 2.52	-1.76 ± 2.63	-9.65 ± 3.11	-3.37 ± 3.37	-0.90 ± 1.35
	Min., Max.	3.82, 14.92	-10.04, -0.06	-16.57, -5.06	-9.44, +4.46	-4.25, +0.06
	1MBG	12.77	-8.72	-5.24	-1.19	-3.52
D100–R133 in c-Myb R2	$\langle 1MBH \rangle_{av}$	10.13 ± 1.19	-12.37 ± 0.69	-2.12 ± 1.09	-4.36 ± 0.91	-6.43 ± 0.30
	Min., Max.	7.72, 12.37	-13.91, -10.97	-5.71, -0.66	-7.19, -2.66	-7.08, -5.46
	1MBG	11.34	-13.10	-1.79	-3.55	-6.41
E150–R153 in c-Myb R3	$\langle 1MBK \rangle_{av}$	1.50 ± 1.02	-2.03 ± 1.76	-1.85 ± 2.02	-2.38 ± 2.58	-1.83 ± 1.63
	Min., Max.	0.10, 4.66	-6.05, -0.10	-6.97, +0.46	-7.29, +3.45	-5.71, -0.09
	1MBJ	0.78	-6.89	-0.65	-6.75	-6.68
E151–R176 in c-Myb R3	$\langle 1MBK \rangle_{av}$	6.58 ± 2.05	-6.85 ± 2.45	-5.16 ± 3.75	-5.43 ± 2.87	-4.29 ± 1.39
	Min., Max.	3.78, 14.65	-14.42, -0.75	-19.56, +0.48	-11.37, +1.25	-6.09, -0.38
	1MBJ	5.77	-8.47	-2.80	-5.50	-5.81
K2–D7 in CRIP	$\langle 1IML \rangle_{av}$	1.20 ± 0.97	-0.57 ± 0.36	-0.13 ± 0.68	$+0.50 \pm 1.07$	-0.44 ± 0.30
	Min., Max.	0.23, 3.93	-1.81, -0.10	-2.09, +1.34	-1.51, +2.76	-1.51, -0.09
	Conformer 16	0.36	-1.47	-0.26	-1.37	-1.27
K13–E27 in CSE-I	$\langle 1B3C \rangle_{av}$	4.68 ± 2.44	-2.39 ± 1.59	-0.52 ± 2.05	$+1.76 \pm 2.98$	-1.36 ± 0.96
	Min., Max.	1.26, 10.05	-6.11, -0.49	-8.29, +3.27	-3.68, +9.78	-3.72, -0.28
	2B3C	4.14	-6.08	+0.30	-1.64	-4.31
E68–K84 in cyanovirin-N	$\langle 2EZN \rangle_{av}$	1.72 ± 1.51	-0.95 ± 0.55	-0.26 ± 0.85	$+0.52 \pm 1.57$	-0.68 ± 0.41
	Min., Max.	0.47, 5.08	-2.94, -0.49	-4.21, +0.72	-1.21, +4.09	-2.01, -0.34
	2EZM	0.43	-0.90	-0.04	-0.51	-0.70
	X-tal 3EZM	3.28	-1.65	-1.06	+0.57	-1.27
E61–K99 in horse heart cytochrome c (Reduced)	$\langle 2GIW \rangle_{av}$	1.51 ± 1.13	-3.49 ± 2.05	-2.35 ± 2.54	-4.33 ± 2.02	-3.02 ± 1.94
	Min., Max.	0.40, 5.88	-7.14, -0.45	-10.56, -0.02	-8.38, -0.44	-6.59, -0.29
	1GIW	0.52	-5.99	-0.36	-5.84	-5.60
K62–E66 in HMG1 Box 2	$\langle 1HSM \rangle_{av}$	1.70 ± 1.31	-0.98 ± 0.65	-0.11 ± 0.98	$+0.61 \pm 2.03$	-0.73 ± 0.54
	Min., Max.	0.40, 7.34	-3.45, -0.41	-2.29, +5.31	-1.60, +11.81	-2.59, -0.29
	1HSM	2.47	-3.38	+0.22	-0.69	-2.42

TABLE VI. (Continued)

Ion pair and protein ^a	Conformer ensemble ^b	$\Delta\Delta G_{dslv}$ (kcal/mol)	$\Delta\Delta G_{brd}$ (kcal/mol)	$\Delta\Delta G_{prt}$ (kcal/mol)	$\Delta\Delta G_{tot}$ (kcal/mol)	$\Delta\Delta G_{assoc}$ (kcal/mol)
K54–D58 in ISL-I	$\langle 1BW5 \rangle_{av}$	4.76 ± 2.18	-2.01 ± 1.40	$+0.48 \pm 1.45$	$+3.23 \pm 2.49$	-1.09 ± 0.78
	Min., Max.	1.92, 14.76	-8.17, -0.39	-2.14, +7.16	-1.21, +13.13	-3.65, -0.25
	Conformer 6	4.35	-4.22	+0.63	+0.77	-2.71
K4–E15 in protein G B1 domain	$\langle 1GB1 \rangle_{av}$	1.48 ± 0.96	-1.17 ± 0.96	-0.08 ± 0.61	$+0.23 \pm 1.44$	-0.92 ± 0.81
	Min., Max.	0.30, 4.68	-3.99, -0.24	-1.38, +2.38	-3.17, +3.94	-3.45, -0.17
	$\langle 3GB1 \rangle_{av}$	2.85 ± 1.57	-1.51 ± 1.25	-0.43 ± 0.75	$+0.90 \pm 2.03$	-1.09 ± 1.01
	Min., Max.	0.37, 6.42	-4.54, -0.22	-1.82, +2.57	-2.92, +6.27	-3.83, -0.16
	2GB1	0.60	-1.06	+0.09	-0.37	-0.81
	X-tal 1PGA	1.56	-2.61	-0.20	-1.24	-1.88
	X-tal 1PGB	1.84	-5.86	-0.26	-4.28	-5.15
E27–K28 in protein G B1 domain	$\langle 1GB1 \rangle_{av}$	1.61 ± 1.07	-0.27 ± 0.13	0.00 ± 0.62	$+1.35 \pm 1.14$	-0.21 ± 0.09
	Min., Max.	0.46, 4.36	-0.67, -0.11	-2.61, +2.55	+0.08, +6.24	-0.46, -0.10
	$\langle 3GB1 \rangle_{av}$	1.41 ± 0.85	-0.32 ± 0.29	-0.44 ± 0.54	$+0.64 \pm 0.75$	-2.41 ± 0.17
	Min., Max.	0.64, 3.53	-1.76, -0.13	-1.85, +0.31	-0.85, +2.44	-1.08, -0.11
	2GB1	1.11	-0.21	+0.21	+1.11	-0.18
	X-tal 1PGA	4.07	-2.87	-4.55	-3.36	-1.95
	X-tal 1PGB	3.22	-0.27	-3.83	-0.87	-0.20
D47–K50 in protein G B1 domain	$\langle 1GB1 \rangle_{av}$	1.17 ± 1.02	-0.42 ± 0.37	-0.03 ± 0.45	$+0.71 \pm 1.07$	-0.30 ± 0.29
	Min., Max.	0.21, 5.49	-2.19, -0.12	-1.69, +1.07	-2.09, +4.77	-1.94, -0.09
	$\langle 3GB1 \rangle_{av}$	3.15 ± 1.67	-0.81 ± 0.62	-1.67 ± 1.37	$+0.67 \pm 1.26$	-0.55 ± 0.48
	Min., Max.	1.07, 7.23	-3.05, -0.22	-4.53, +0.22	-2.52, +2.92	-2.34, -0.16
	2GB1	0.51	-0.32	+0.07	+0.27	-0.25
	X-tal 1PGA	2.64	-3.07	-4.66	-5.09	-2.28
	X-tal 1PGB	2.02	-2.00	-4.55	-4.54	-1.37
E108–R110 in U1 SNRP A	$\langle 1FHT \rangle_{av}$	1.01 ± 0.76	-0.34 ± 0.41	-0.90 ± 1.10	-0.23 ± 1.22	-0.29 ± 0.37
	Min., Max.	0.01, 2.80	-2.43, -0.04	-3.60, +1.27	-3.22, +2.28	-2.17, 0.00
	Conformer 36	1.70	-2.43	-0.72	-1.45	-2.17

* $\Delta\Delta G_{dslv}$, $\Delta\Delta G_{brd}$, $\Delta\Delta G_{prt}$, $\Delta\Delta G_{tot}$, and $\Delta\Delta G_{assoc}$ denote desolvation energy, bridge energy, protein energy, total energy, and association energy terms respectively. (See description of these energy terms in the text.) Standard deviations for the average values are also given. Min and Max denote the minimum and maximum values of these terms in the ensemble. Note that the ensemble averages presented are simply the arithmetic averages. They do not take into account the relative conformer populations. As such, they do not represent the average ion pair stability in solution. The values of these energy terms are also given in the crystal/minimized average structure and in the “most representative” conformer. Prefix X-tal to a PDB entry indicates that it is the crystal structure; otherwise, it is the minimized average structure. The “most representative” conformer is indicated by its conformer number in the ensemble.

^aAn ion pair is indicated by the constituent residue names and numbers.

^bThe NMR conformer ensemble is indicated by its Protein Data Bank (PDB) entry name. $\langle \dots \rangle_{av}$ denotes the average values of the energy terms in the NMR conformer ensemble.

^cTriad of charged residues Asp 12, Asp 57, and Lys 109 in CheY.

ing in a conformer-dependent manner. The origin of the fluctuations in electrostatic strengths of these ion pairs was traced to the fluctuations in the location of the charged residues and their orientations with respect to one another.²⁸ However, the solution structure of the c-Myc–Max leucine zipper was determined using only ¹H NMR. It was therefore unclear whether the observed fluctuations in electrostatic interactions would be seen in other proteins as well. We use NMR conformer ensembles from several proteins to examine the consistency in this relationship between the location and the orientation of charge–charge interactions and their electrostatic strengths.

Validity of the Observed Fluctuations

It can be debated whether a conformer ensemble obtained by NMR experiments truly reflects the dynamics of protein behavior in solution. The NMR conformer ensemble contains different structures that are compatible with the existing, and the missing nOe and other restraints.¹⁸ However, frequently, only partial datasets are

available. Hence, NMR derived structures are usually underdetermined. This, along with the procedures followed in the structural assignment, may lead to difficulties in distinguishing and quantifying differences between the true dynamic behavior of the molecule and inaccuracies in the structure determination. Since negatively charged side-chain carboxylate (COO⁻) groups in Asp and Glu lack protons, ¹H NMR studies can not directly localize these groups. Also, the ¹H chemical shift assignments are often incomplete for the long side-chains of positively charged residues, Arg and Lys. Furthermore, many commonly used NMR structure refinement programs do not contain electrostatic terms. A salt bridge present in the crystal structure or average energy minimized structure may not be retained in most of the NMR conformers in the absence of sufficient nOe restraints and explicit electrostatic term to model the salt bridge. Moreover, information on relative populations of different NMR conformers is absent in the PDB files. Hence, it cannot be concluded that the NMR conformers that retain the salt bridges seen in crystal/

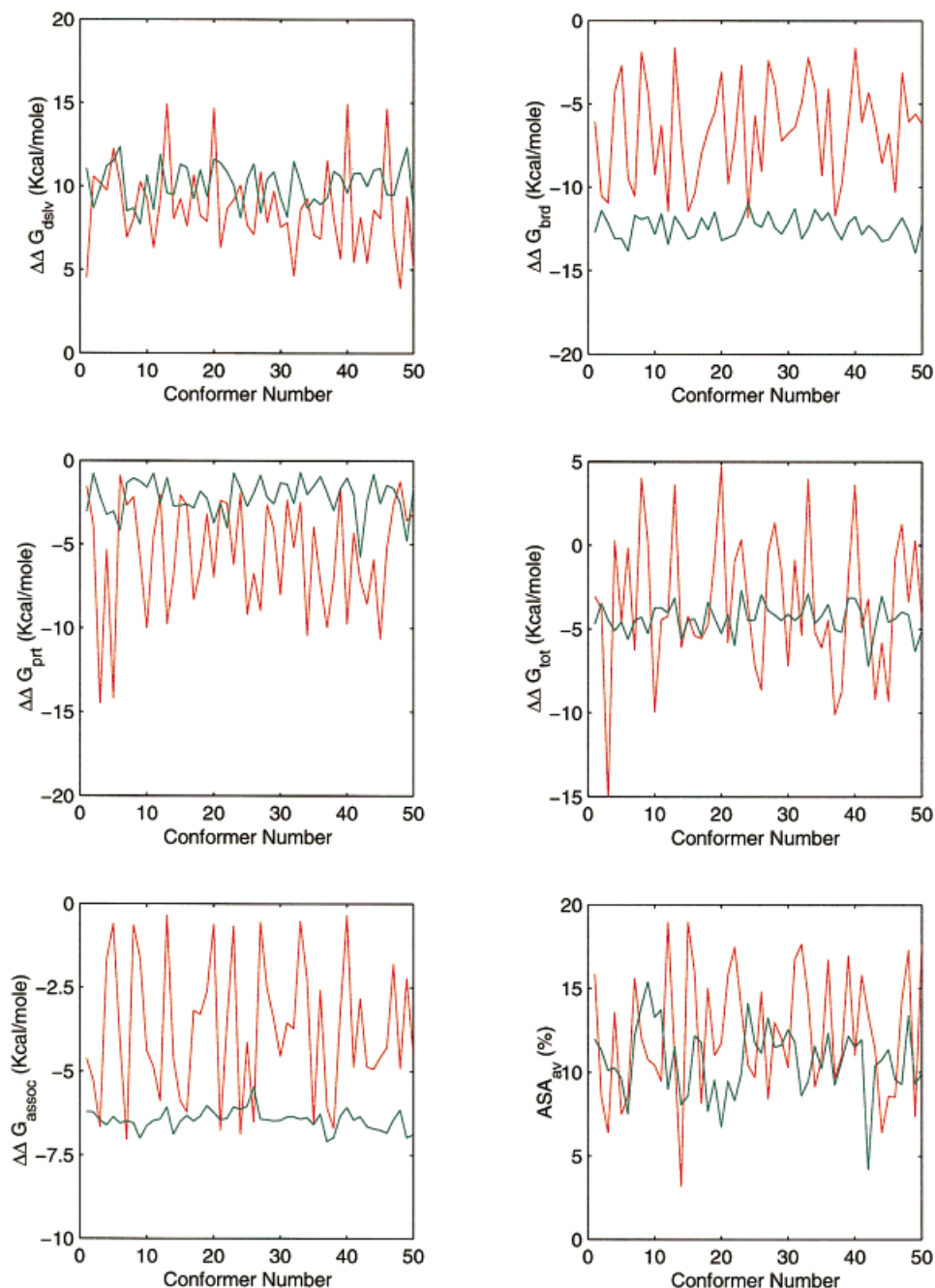


Fig. 3. Example of conformer-dependent fluctuations of ion pair stabilities in the NMR conformer ensembles. Two ion pairs, those of Fig. 1, are depicted. The ion pair E47–R73 of c-Myb DNA binding domain R1 is shown in red, while the ion pair D100–R133 of c-Myb DNA binding domain R2 is shown in green. The plots show variation in all energy terms, i.e., $\Delta\Delta G_{\text{dslv}}$, $\Delta\Delta G_{\text{bnd}}$, $\Delta\Delta G_{\text{prt}}$, $\Delta\Delta G_{\text{tot}}$ and $\Delta\Delta G_{\text{assoc}}$, and average ASA for each ion pair in the different NMR conformers. The energy terms and average ASA are described in the text. The ion pair E47–R73 (red) shows extensive fluctuations in its stability. See legend for Fig. 1 for further details.

average energy minimized structure or in the “most representative” conformer have high-population times. These difficulties should be borne in mind when interpreting Table V.

Despite these shortcomings in the NMR methodology, structure calculation and refinement protocols, it is becoming increasingly clear that at least part of the conforma-

tional variability of NMR structures may arise from protein internal motion. The local disorder in NMR structural ensembles has been compared with crystallographic temperature factors. The percentage of charged residue side-chains exhibiting greater than average disorder is similar between ensembles of NMR structures and crystallographic B-factors.⁴⁹ There are a number of reports in the

TABLE VII. Number of Conformers in Which 22 Ion Pairs Are Stabilizing or Destabilizing in the NMR Conformer Ensemble

Protein ensemble	Ion pair	Total no. of conformers	No. of conformers in which ion pair is stabilizing	No. of conformers in which ion pair is destabilizing
⟨1MPH⟩	D40–K42	50	49	1
⟨1MPH⟩	K91–E95	50	46	4
⟨1CEY⟩	D12–K109	46	3	43
⟨1CEY⟩	D57–K109	46	1	45
⟨1CEY⟩	D41–K45	46	21	25
⟨1MBF⟩	E47–R73	50	39	11
⟨1MBF⟩	D48–R81	50	50	0
⟨1MBF⟩	E49–K52	50	48	2
⟨1MBH⟩	K92–E99	50	42	8
⟨1MBH⟩	D100–R133	50	50	0
⟨1MBK⟩	E150–R153	50	38	12
⟨1MBK⟩	E151–R176	50	48	2
⟨1IML⟩	K2–D7	48	16	32
⟨1B3C⟩	K13–E27	40	14	26
⟨2EZN⟩	E68–K84	40	25	15
⟨2G1W⟩	E61–K99	40	40	0
⟨1HSN⟩	K62–E66	49	18	31
⟨1BW5⟩	K54–D58	50	5	45
⟨1GB1⟩	K4–E15	60	28	32
⟨3GB1⟩	K4–E15	32	12	20
⟨1GB1⟩	E27–K28	60	0	60
⟨3GB1⟩	E27–K28	32	3	29
⟨1GB1⟩	D47–K50	60	6	54
⟨3GB1⟩	D47–K50	32	7	25
⟨1FHT⟩	E108–R110	43	24	19

^aNMR conformer ensembles are indicated by their Protein Data Bank (PDB) file names (in brackets).

^bAn ion pair is indicated by the constituent residue names and numbers.

literature that compare the NMR ensembles with those derived from molecular dynamic simulations, showing that NMR conformer ensembles reflect protein dynamics in solution (e.g., ref. 20 and 21, and references therein). NMR is an important tool for studying protein dynamics in solution and for the characterization of states not accessible to X-ray crystallography.⁵⁰ Heteronuclear NMR relaxation measurements are useful in studying and overcoming ambiguities in NMR data due to protein flexibility.^{51,52} The determination of ¹⁵N and ¹³C relaxation parameters has become a routinely used method to characterize protein flexibility in terms of backbone and side-chain dynamics.²² Using NMR relaxation methods, Lee et al.⁵³ have studied the dynamics of calcium-saturated calmodulin in the formation of a complex with a peptide model of smooth muscle myosin light chain kinase. Their studies have resolved the motion of individual residues between the bound and unbound states and discriminated between backbone and side-chain perturbations. Recently, Goodman et al.⁵⁴ carried out a statistical analysis of protein dynamics using a database of backbone order parameters derived from ¹⁵N NMR relaxation data. Their studies indicate that amino acids with small side-chains show

greater backbone mobility than those with larger side-chains. Furthermore, the motion of a given backbone or side-chain NH group appears to depend upon the sizes of neighboring amino acids in the sequence, in the secondary and in the tertiary structures. ²H auto-correlation and ¹³C cross-correlation NMR experiments are useful in studying side-chain dynamics.⁵⁵ Recently, the Lipari–Szabo model-free analysis of protein dynamics has been gaining wide acceptance. This analysis facilitates interpretation of heteronuclear NMR relaxation data in terms of a generalized order parameter, S^2 . S^2 is a measure of the spatial extent of the internal motion of the protein.⁵⁶ A recent PubMed inquiry (on the Internet) has shown that an increasing number of molecular dynamics studies have been able to reproduce the fast motion dynamics estimated by the S^2 parameter in the well defined regions of the proteins.⁵⁷ In most NMR experiments, main-chain coordinates are usually more reliable than the side-chain atomic coordinates. Comparison of the NMR derived side-chain conformation coordinates with those predicted by applying the side-chain packing algorithms to the NMR derived main-chain coordinates and to crystal structures shows a good agreement in the well defined regions.⁵⁸

Complementation of homonuclear ¹H NMR spectra with multidimensional heteronuclear ¹⁵N/¹³C – ¹H NMR experiments often provides valuable data on ion pair interactions. For example, Moy et al.⁵⁹ have monitored the distance between the side-chain nitrogen atom of Lys 109 and oxygen atoms of the side-chain carboxyl group of Asp 57 in 46 simulated annealing NMR structure calculations for CheY. These investigators have used ¹⁵N–¹H^N vectors and long range nOe restraints to define the structure in this region of the protein. They conclude that the salt bridge between Asp 57 and Lys 109 is broken in the presence of Mg²⁺. This salt bridge had been observed in a high-resolution crystal structure of Mg²⁺ free form of CheY. Similarly, Ogata et al.³⁷ have reported the presence of intrarepeat salt bridges in c-Myb DNA binding domain based on NMR measurements. The solution structure of c-Myc–Max leucine zipper has also revealed the presence of six intra- and interhelical ion pairs and ruled out the presence of another interhelical ion pair that was predicted from the sequences of c-Myc and Max helices and coiled-coil geometry.⁶⁰ Liu et al.⁶¹ have detected side-chain side-chain hydrogen bonding interactions by using ¹³C and ¹⁵N labeled proteins.

Taken together, these reports indicate that NMR conformer ensembles reflect a certain extent of real protein flexibility, particularly in the well-defined regions. However, clearly, we can not rule out the possibility of errors due to inherent ambiguities in NMR data and potential artifacts in the structure calculation protocols. Hence, we have evaluated the validity of the observed fluctuations in several different ways.

First, we have attempted to assess the quality of the NMR structures in our database. To do that, we have extracted information on the actual experiments, such as the number of nOe, hydrogen bond, dihedral angle and dipolar coupling restraints, from the PDB file headers and

from the original papers describing the NMR experiments for each protein in our database. While some PDB file headers contain information on the number of the restraints, others provide restraint energy information for the conformers in the ensemble. Experimental details for each structure are given in the relevant literature reports cited in Table I and summarized in Table II. The available information is heterogeneous. However, many common points can be noted. Homonuclear (^1H) and heteronuclear (^{13}C and/or ^{15}N) relaxation experiments have been performed at field strengths of 500–800 MHz. A large number of intra-residue, inter-residue sequential, medium-range, and long-range nOe restraints have been used, along with several hydrogen bond and dihedral angle restraints. Most NMR structure calculations and refinements have been performed using distance geometry restraints and simulated annealing protocols (Table II). The quality of the NMR conformer ensembles can be assessed by the average of the root-mean-square deviations (RMSD) of the atomic coordinates of individual conformers with respect to the average conformer (Table II). It can be seen that the quality is good, especially in the more recent NMR structures that contain results from both homonuclear and heteronuclear relaxation experiments.

Table IV shows conformations of residues involved in salt bridge (ion pair) interactions in our database. All, except six, of these residues are located in either α -helical (H) or β -strand (B) regions of the proteins. Secondary structural regions are usually better defined than the loop regions. Most groups have deposited in the PDB only those conformers that show no restraint violation. Except for three, all residues constituting the 22 ion pairs in our database fall within well-defined regions of the proteins. The three residues that lie in disordered regions are Glu 108 and Arg 110 (ion pair Glu 108–Arg 110) in (1FHT) and Asp 58, belonging to ion pair K54–D58 in (1BW5).

For eight proteins in our database, heteronuclear (^{13}C and/or ^{15}N) experiments have been performed along with the homonuclear ^1H NMR experiments. Only ^1H NMR data are available for the remaining three proteins, β -spectrin PH domain, CSE-I and the B1 domain of protein G (Table II). These three proteins contain six ion pairs (Table I). Fluctuations seen in their stabilities are similar to those of other ion pairs in the database. Since we observed fluctuations in the stabilities of almost all ion pairs, we can safely conclude that our calculations are not affected to a significant extent by the differences in the experimental methods and that the quality of data used in our analysis is high.

Second, we have studied the crystallographic B-factors in those cases in which high-resolution structures are available, along with the NMR conformer ensembles. B-factors represent the electron density smear due to thermal motion and positional disorder of atoms.⁶² Hence, at high resolution, B-factor values may reflect protein flexibility. High-resolution protein crystal structures are available for CheY, cyanovirin-N, and the B1 domain of protein G. Table VIII lists the average B-factors for all and side-chain atoms of the ion pairing residues in these

structures. As expected, atoms belonging to side-chains show larger B-factor values. However, the average B-factors of these ion pairing residues in the protein crystal structures do not correlate with either the fluctuations in location (ASAs) of these residues or with ion pair geometries in the corresponding NMR ensembles. Similarly, no correlation is observed between the average B-factors of these residues and the fluctuations in the total electrostatic contribution ($\Delta\Delta G_{\text{tot}}$) of the corresponding ion pairs to stability in the ensembles. This may be due to different conformer populations in crystalline and solution.

Third, in the case of CheY and the B1 domain of protein G, two crystal structures are available for each protein (Table I). Table III indicates that the geometric orientations of the six ion pairs differ substantially between the crystal forms, the NMR conformer ensemble geometric average and the geometries in the average energy minimized structure. Table IV indicates the same for the location of the charged residues constituting these six ion pairs. The electrostatic strengths of these ion pairs differ in the two crystal forms (Table VI). The salt bridge D57–K109 interconverts between being stabilizing and destabilizing in the two crystal structures of CheY. In 1CHN, D57–K109 is destabilizing ($\Delta\Delta G_{\text{tot}} = +2.72$ kcal/mol). In contrast, it is stabilizing ($\Delta\Delta G_{\text{tot}} = -3.05$ kcal/mol) in 3CHY. The ion pair network formed by residues Asp 12, Asp 57, and Lys 109 shows fluctuations in the NMR ensemble, as well as in the two crystal structures of CheY. In 1CHN, this ion pair network is a weakly destabilizing salt bridge network ($\Delta\Delta G_{\text{tot}} = +0.19$ kcal/mol; Table VI). In 3CHY, it is stabilizing ($\Delta\Delta G_{\text{tot}} = -4.94$ kcal/mol). These observations indicate that ion pair geometries and their electrostatic stabilities fluctuate in crystal structures as well.

Fourth, we compare two sets of NMR experimental data on the B1 domain of protein G.^{42,43} This comparison addresses concerns with regard to errors and artifacts in the NMR structure determination methodology. The first NMR structure for the B1 domain of protein G was solved by Gronenborn et al.⁴² The same group has reported further improvements in the accuracy of the structure by using a pseudopotential function for the radius of gyration.⁴³ The motivation for this refinement was their observation that NMR structures are usually poorly packed and expanded, as compared with X-ray structures. The NMR ensemble from the first experiment (1GB1) contains 60 conformers, while that from the second experiment (3GB1) contains 32. The B1 domain of protein G contributes three ion pairs, K4–E15, E27–K28, and D47–K50 to our database. Figure 4a–c plots the electrostatic energy terms and the average ASA for these ion pairs in the conformer ensembles (1GB1) and (3GB1). All terms for the three ion pairs show fluctuations. The nature of the fluctuations is similar in both ensembles for K4–E15 and E27–K28, although some variations can be noticed (Fig. 4a,b). Table III shows that the geometries of all three ion pairs are similar between the two ensembles and differ substantially from those in the crystal structures, 1PGA and 1PGB. The locations of the charged residues in the three

TABLE VIII. B-Factors of Ion Pairing Residues in Crystal Structures*

Protein name	Ion pairing residue	Atoms ^a	B-factor (Å ²)	
			Structure 1	Structure 2
CheY	D12	All	8.53 ± 1.73	10.00 ± 2.35
		Side-chain	9.57 ± 1.66	11.87 ± 1.77
	D41	All	12.85 ± 2.95	11.81 ± 1.33
		Side-chain	14.53 ± 3.36	12.93 ± 0.77
	K45	All	17.24 ± 3.97	20.40 ± 3.98
		Side-chain	19.49 ± 4.12	23.05 ± 3.37
	D57	All	8.69 ± 1.20	8.65 ± 1.65
		Side-chain	9.29 ± 1.41	9.96 ± 1.18
	K109	All	7.61 ± 2.07	10.54 ± 1.31
		Side-chain	6.08 ± 1.26	11.47 ± 0.96
B1 domain of protein G	K4	All	19.19 ± 9.97	15.21 ± 6.74
		Side-chain	26.45 ± 6.91	19.97 ± 5.22
	E15	All	17.54 ± 5.15	19.43 ± 7.26
		Side-chain	21.44 ± 3.13	24.65 ± 5.18
	E27	All	17.85 ± 5.50	21.71 ± 5.73
		Side-chain	21.88 ± 3.76	25.94 ± 3.79
	K28	All	14.76 ± 3.65	20.45 ± 4.28
		Side-chain	17.20 ± 1.75	23.59 ± 2.94
	D47	All	10.73 ± 2.58	22.84 ± 2.48
		Side-chain	12.47 ± 2.11	24.53 ± 1.35
	K50	All	10.52 ± 2.65	17.88 ± 4.59
		Side-chain	12.39 ± 1.96	21.25 ± 2.93
Cyanovirin-N	E68	All	17.54 ± 5.31	
		Side-chain	21.37 ± 3.82	
	K84	All	11.31 ± 3.24	
		Side-chain	13.57 ± 2.43	

*Average B-Factors of the Atoms in the Ion Pairing Residues in the Crystal Structures of CheY, Cyanovirin-N, and the B1 Domain of Protein G. In most cases, the atoms in the side-chains have greater B-factors indicating their larger mobilities.

^aAll, all atoms in the residue; side-chain, atoms beyond C^α.

ion pairs, except K50, also show similar fluctuations in the two ensembles (Table IV).

D47–K50 shows appreciable differences between the two ensembles (Fig. 4c). The ASA of K50 is smaller in the ensemble ⟨3GB1⟩ and in better agreement with that of the crystal structures, 1PGA and 1PGB (Table IV). The bridge and protein energy terms for this ion pair are more stabilizing in ⟨3GB1⟩ as compared with those in ⟨1GB1⟩ (Fig. 4c and Table VI). However, this gain in stabilization due to improved bridge and protein energy terms is almost nullified by the greater desolvation penalty paid by D47–K50 in the ensemble ⟨3GB1⟩ as compared with ⟨1GB1⟩. As a result, the total stabilities $\Delta\Delta G_{\text{tot}}$ have similar magnitudes in the two ensembles (Table VI and Fig. 4c). Interestingly, this ion pair has greater stability and improved geometry in both crystal structures, 1PGA and 1PGB, as compared with those in the NMR ensembles.

Methodology and Overall Ion Pair Stability

This study focuses on the fluctuations in electrostatic contribution to ion pair stability in NMR conformer

ensembles. At the same time, the results it obtains validate certain aspects of the continuum electrostatics methodology. Continuum electrostatics has been widely used in calculations of free energy changes due to formation of electrostatic interactions in proteins. Typically, these methods include atomic description of the protein (solute) but treat water (solvent) only in terms of its bulk properties.⁶³ In recent years, these methods have become increasingly popular as tools for studies of solvation free energies of proteins and organic molecules. The free energy change upon salt bridge formation has often been estimated using this methodology.^{4,10,11,13,25–28} Our present study indicates that this method is sensitive to the details of the conformation, and can be used to study dynamic charge–charge interactions in proteins.

The overall stability of an ion pair may include energy terms other than the electrostatic contribution. Almost all charged residue side-chains contain hydrophobic methyl or methylene groups. Hence, additional energy terms may be due to van der Waals and hydrophobic interactions. The

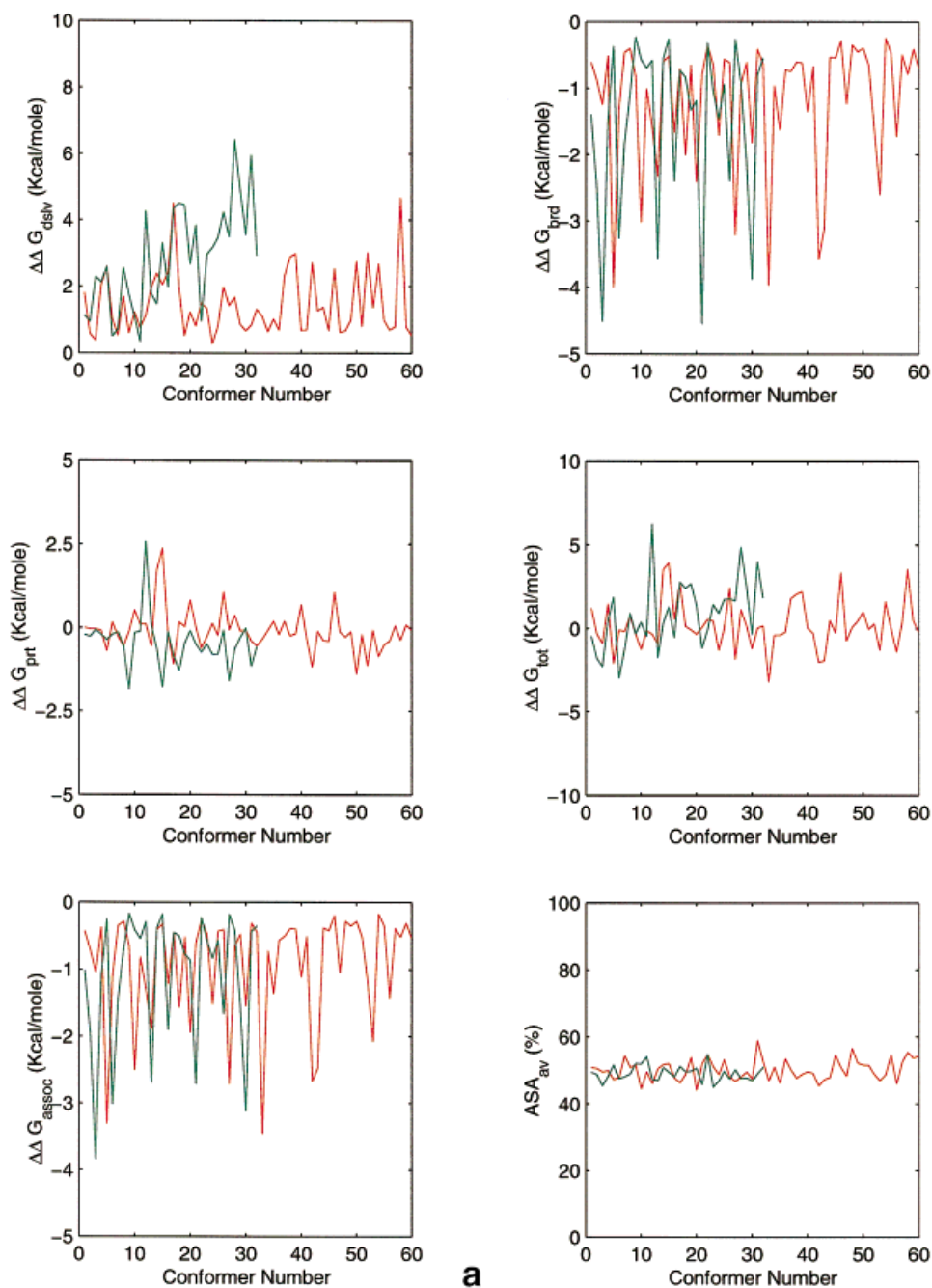


Fig. 4. Comparison of the fluctuations in stabilities of the ion pairs (a) K4–E15, (b) E27–K28, and (c) D47–K50 in two NMR conformer ensembles (1GB1) (red) and (3GB1) (green) of B1 domain of protein G. Additional pseudopotential restraints for the radius of gyration were used for (3GB1). The ensemble (1GB1) contains 60 conformers, while there are 32 conformers in (3GB1). The stability of the three ion pairs fluctuate in both ensembles. The fluctuations are roughly similar in ion pairs K4–E15 and E27–K28. Fluctuations in the different energy terms are different for the ion pair D47–K50 in the two ensembles. D47–K50 is more buried in protein globule, pays greater desolvation energy penalty, and has more stabilizing bridge and protein energy terms in (3GB1) than in (1GB1). However, the total free energy contribution of this ion pair fluctuates similarly in both ensembles. See text for details.

loss of side-chain conformational entropy in the folded protein may also be a factor in the overall ion pair stability. However, it remains unclear whether these additional energy terms may significantly reduce or enhance the observed fluctuations in the electrostatic contribution to

ion pair stability. Observations on ion pair movements and locations of charged residues in proteins reported here suggest that salt bridges in proteins may break and reform, pointing to a rather small energy difference between the two states.

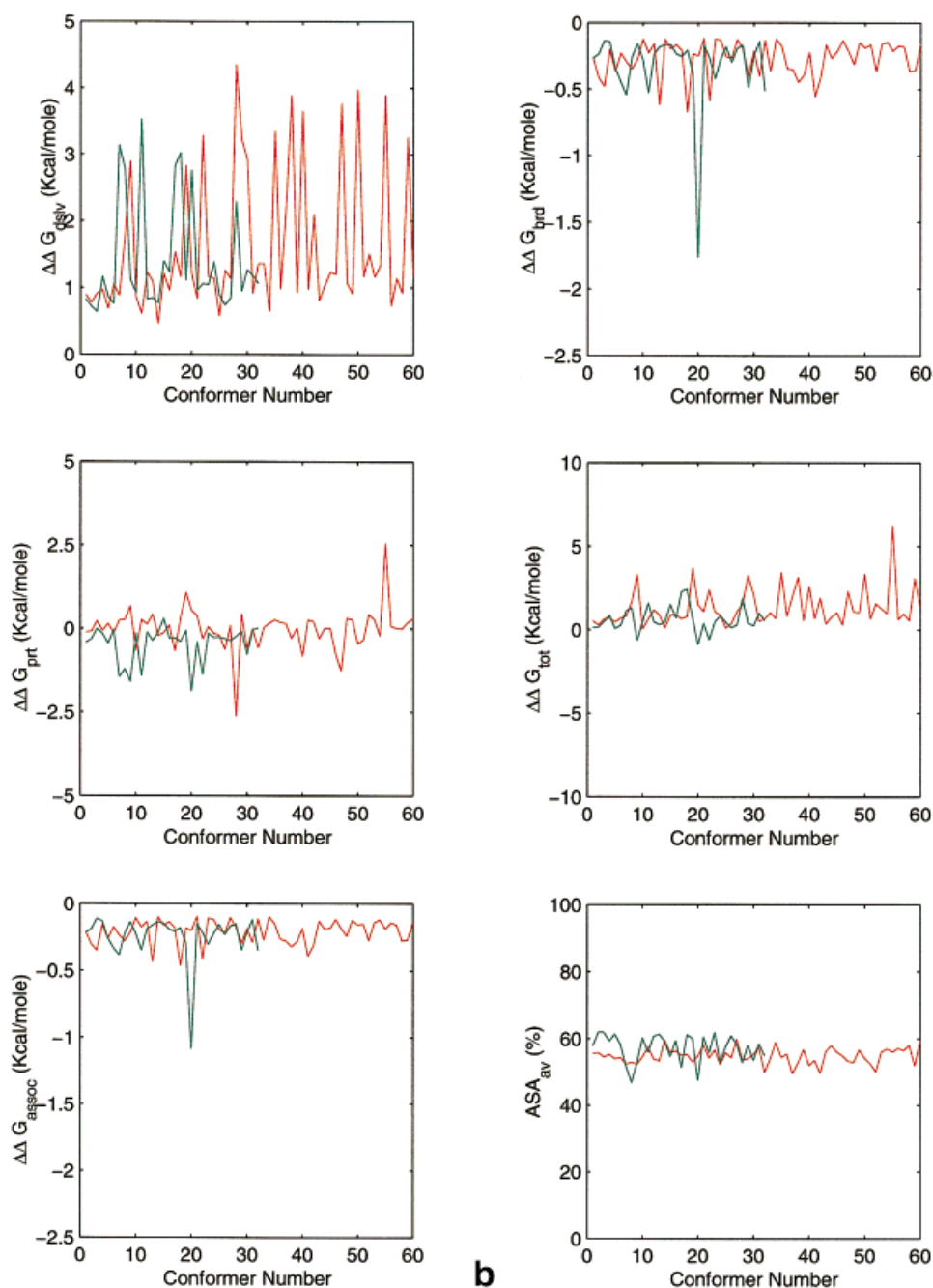


Figure 4. (Continued.)

CONCLUSIONS

The question of whether salt bridges are stabilizing or destabilizing to protein structures has been frequently addressed in the experimental and computational literature. Recently, a systematic study on a large number of salt bridges in high-resolution protein crystal structures has shown that the contributions of salt bridges/ion pairs to protein stability depend on the geometry of the charged pairing groups, and their environment in the structure.¹² Given this strong dependence of the stabilizing/destabiliz-

ing contribution on geometry, structural and the exposed/buried environment, here we investigate the electrostatic contributions of ion pairs toward protein stability in ensembles of conformers around their native state. These conformer ensembles can be obtained by molecular dynamics simulations or by NMR.

In the present study, we carry out an extensive, large-scale investigation of all proteins in the protein structural database which have at least 40 NMR conformers. We compute the electrostatic contribution of each ion pair in

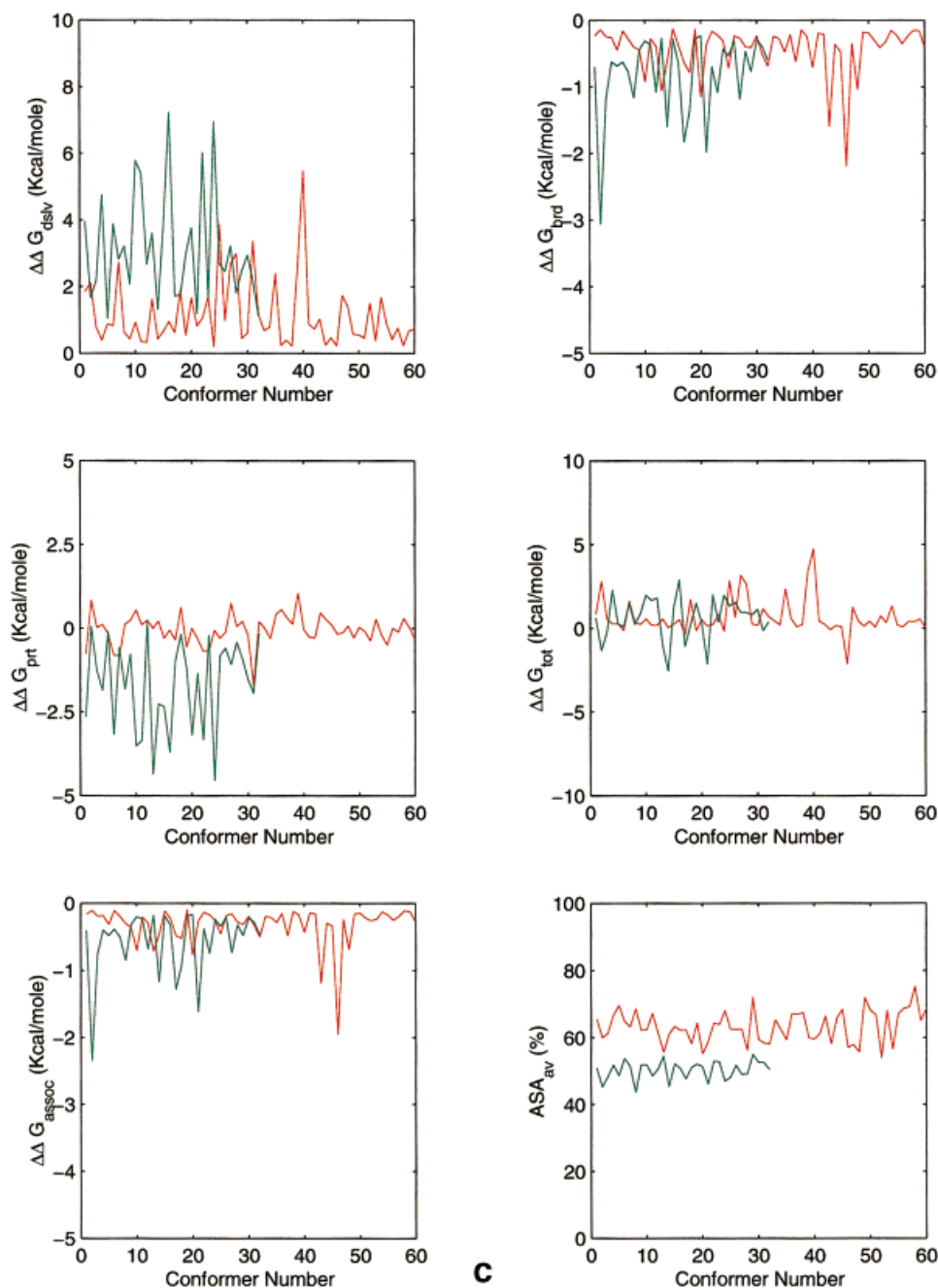


Figure 4. (Continued)

each conformer, resulting in a total of 1,201 ion pair stability calculations. Protein crystal structures and average energy minimized NMR structures have also been used, where available. The stabilities of almost all ion pairs fluctuate, interconverting between being stabilizing and destabilizing. These observations on NMR conformer ensembles and protein crystal structures provide hints to fluctuations in electrostatic interactions in solution, illustrating that the electrostatic contribution of an ion pair to protein stability is conformer dependent. If the interaction between the ion pairing residues is stabilizing in the conformers that have

high population times, that ion pair will be stabilizing to the protein structure. Salt bridges observed in protein crystal structures may often break and reform in solution as a result of the movement of the charged residues. Other pairs of charged residues may come together and form salt bridges either to compensate, or in addition to, the salt bridges observed in protein crystal structures.

ACKNOWLEDGMENTS

The authors thank Drs. Buyong Ma, Chung-Jung Tsai, Neeti Sinha, and Yuk Sham for helpful discussions. In

particular, we thank Dr. Jacob V. Maizel, Jr., for encouragement throughout this project. The personnel at FCRDC are thanked for their assistance. The research of R. Nussinov in Israel has been supported in part by grant number 95-00208 from Binational Science Foundation (BSF), Israel, by a grant from the Ministry of Science, by the Center of Excellence, administered by the Israel Academy of Sciences, by the Magnet grant, and by the Tel Aviv University Basic Research and Adams Brain Center grants. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-56000. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the U.S. government.

REFERENCES

- Carlson HA, McCammon JA. Accommodating protein flexibility in computational drug design. *Mol Pharmacol* 2000;57:213–218.
- McCammon JA, Gelin BR, Karplus M, Wolynes PG. The hinge-bending mode in lysozyme. *Nature* 1976;262:325–326.
- Go N, Naguti T, Nishikawa T. Dynamics of a small globular protein in terms of low frequency vibrational modes. *Proc Natl Acad Sci USA* 1983;80:3696–3700.
- Kumar S, Wolfson H, Nussinov R. Protein flexibility and electrostatic interactions. *IBM J Res Dev* 2001;in press.
- Tsai CJ, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels and protein function. *Protein Sci* 1999;8:1181–1190.
- Tsai CJ, Ma B, Nussinov R. Folding and binding cascades: shifts in energy landscapes. *Proc Natl Acad Sci USA* 1999;96:9970–9972.
- Kumar S, Ma B, Tsai CJ, Wolfson H, Nussinov R. Folding funnels and conformational transitions via hinge-bending motions. *Cell Biochem Biophys* 1999;31:141–164.
- Ma B, Kumar S, Tsai CJ, Nussinov R. Folding funnels and binding mechanisms. *Protein Eng* 1999;12:713–720.
- Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* 2000;9:10–19.
- Hendsch ZS, Tidor B. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci* 1994;3:211–226.
- Lounnas V, Wade RC. Exceptionally stable salt bridges in cytochrome P450cam have functional roles. *Biochemistry* 1997;36:5402–5417.
- Kumar S, Nussinov R. Salt bridge stability in monomeric proteins. *J Mol Biol* 1999;293:1241–1255.
- Xiao L, Honig B. Electrostatic contributions to the stability of hyperthermophilic proteins. *J Mol Biol* 1999;289:1435–1444.
- Sheinerman FB, Norel R, Honig B. Electrostatic aspects of protein–protein interactions. *Curr Opin Struct Biol* 2000;10:153–159.
- Warshel A, Naray-Szabo G, Sussman F, Hwang JK. How do serine proteases really work? *Biochemistry* 1989;28:3629–3637.
- Kazmirski SL, Li A, Daggett V. Analysis methods for comparison of multiple molecular dynamics trajectories: applications to protein unfolding pathways and denatured ensembles. *J Mol Biol* 1999;290:283–304.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr., Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: a computer based archival file for macromolecular structures. *J Mol Biol* 1977;112:535–542.
- Branden C, Tooze J. *Introduction to protein structure*. 2nd ed. New York: Garland; 1999.
- MacArthur MW, Thornton JM. Conformational analysis of protein structures derived from NMR data. *Proteins* 1993;17:232–251.
- Philippopoulos M, Lim C. Exploring the dynamic information content of a protein NMR structure: comparison of a molecular dynamics simulation with the NMR and X-ray structures of *Escherichia coli* ribonuclease H1. *Proteins* 1999;36:87–110.
- Abseher R, Horstink L, Hilbers CW, Nilges M. Essential spaces defined by NMR structure ensembles and molecular dynamics simulation show significant overlap. *Proteins* 1998;31:370–382.
- Ishima R, Torchia DA. Protein dynamics from NMR. *Nature Struct Biol* 2000;7:740–743.
- Horovitz A, Fersht AR. Cooperative interactions during protein folding. *J Mol Biol* 1992;224:733–740.
- Marqusee S, Sauer RT. Contribution of a hydrogen bond/salt bridge network to the stability of secondary and tertiary structures in lambda repressor. *Protein Sci* 1994;3:2217–2225.
- Xu D, Lin SL, Nussinov R. Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *J Mol Biol* 1997;265:68–84.
- Barril X, Aleman C, Orozco M, Luque FJ. Salt bridge interactions: stability of ionic and neutral complexes in the gas phase, in solution and in proteins. *Proteins* 1998;32:67–79.
- Kumar S, Ma B, Tsai CJ, Nussinov R. Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins* 2000;38:368–383.
- Kumar S, Nussinov R. Fluctuations between stabilizing and destabilizing electrostatic contributions of ion pairs in conformers of the c-Myc–Max leucine zipper. *Proteins* 2000;41:485–497.
- Lee BK, Richards FM. The interpretation of protein structures. Estimation of static accessibility. *J Mol Biol* 1971;55:379–400.
- Tsai CJ, Nussinov R. Hydrophobic folding units derived from dissimilar monomer structures and their interactions. *Protein Sci* 1997;6:24–42.
- Waldburger CD, Schildbach JF, Sauer RT. Are buried salt bridges important for protein stability and conformational specificity? *Nature Struct Biol* 1995;2:122–128.
- Nielsen JE, Anderson KV, Honig B, Hooft RWW, Klebe G, Vriend G, Wade RC. Improving macromolecular electrostatics calculations. *Protein Eng* 1999;12:657–662.
- Bellolell L, Prieto J, Serrano L, Coll M. Magnesium binding to the bacterial chemotaxis protein CheY results in large conformational changes involving its functional surface. *J Mol Biol* 1994;238:489–495.
- Volz K, Matsumura P. Crystal structure of *Escherichia coli* CheY refined at 1.7-Å resolution. *J Biol Chem* 1991;266:15511–15519.
- Yang F, Bewley CA, Louis JM, Gustafson KR, Boyd MR, Gronenborn AM, Clore GM, Wlodawer A. Crystal structure of Cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. *J Mol Biol* 1999;288:403–412.
- Gallagher T, Alexander P, Bryan P, Gilliland GL. Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. *Biochemistry* 1994;33:4721–4729.
- Ogata K, Morikawa S, Nakamura H, Hojo H, Yoshimura S, Zhang R, Aimoto S, Ametani Y, Hirata Z, Sarai A, Ishii S, Nishimura Y. Comparison of the free and DNA-complexed forms of the DNA-binding domain from C-Myb. *Nature Struct Biol* 1995;2:309–319.
- Jablonsky MJ, Jackson PL, Trent JO, Watt DD, Krishna NR. Solution structure of a β -neurotoxin from the new world scorpion *Centruroides sculpturatus* Ewing. *Biochem Biophys Res Commun* 1999;254:406–412.
- Bewley CA, Gustafson KR, Boyd MR, Covell DG, Bax A, Clore GM, Gronenborn AM. Solution structure of Cyanovirin-N, a potent HIV-inactivating protein. *Nature Struct Biol* 1998;5:571–578.
- Banci L, Bertini I, Huber JG, Spyroulias GA, Turano P. Solution structure of reduced horse heart cytochrome c. *J Biol Inorg Chem* 1999;4:21–31.
- Read CM, Cary PD, Crane-Robinson C, Driscoll PC, Norman DG. Solution structure of a DNA-binding domain from HMG1. *Nucleic Acids Res* 1993;21:3427–3436.
- Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 1991;253:657–661.
- Kuszewski J, Gronenborn AM, Clore GM. Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J Am Chem Soc* 1999;121:2337–2338.
- Kelley LA, Gardner SP, Sutcliffe MJ. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally-related subfamilies. *Protein Eng* 1996;9:1063–1065.
- Nilges M, Macias MJ, O'Donoghue SI, Oschkinat H. Automated NOESY interpretation with ambiguous distance restraints: the

- refined NMR solution structure of the pleckstrin homology domain from β -spectrin. *J Mol Biol* 1997;269:408–422.
46. Perez-Alvarado GC, Kosa JL, Louis HA, Beckerle MC, Winge DR, Summers MF. Structure of the cysteine-rich intestinal protein, CRIP. *J Mol Biol* 1996;257:153–174.
 47. Ippel JH, Larsson G, Behravan G, Zdunek J, Schleucher M, Lundqvist J, Lycksell PO, Wijmenga SS. The solution structure of the homeodomain of the rat insulin-gene enhancer protein Isl-1. Comparison with other homeodomains. *J Mol Biol* 1999;288:689–703.
 48. Avis JM, Allain FHT, Howe P, Varani G, Nagai K, Neuhaus D. Solution structure of the N-terminal Rnp domain of U1A protein: the role of C-terminal residues in structure stability and RNA binding. *J Mol Biol* 1996;257:398–411.
 49. Billeter M. Comparison of protein structures determined by NMR in solution and by X-ray diffraction in single crystals. *Q Rev Biophys* 1992;25:325–377.
 50. Fersht A. Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding. New York: WH Freeman; 1999.
 51. Wuthrich K. NMR of proteins and nucleic acids. New York: John Wiley & Sons; 1986.
 52. Constantine KL, Friedrichs MS, Wittekind M, Jamil H, Chu CH, Parker RA, Goldfarb V, Mueller L, Farmer BT II. Backbone and side chain dynamics of uncomplexed human adipocyte and muscle fatty acid binding proteins. *Biochemistry* 1998;37:7965–7980.
 53. Lee AL, Kinear SA, Wand AJ. Redistribution and loss of side chain entropy upon formation of a calmodulin–peptide complex. *Nature Struct Biol* 2000;7:72–77.
 54. Goodman JL, Pagel MD, Stone MJ. Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters. *J Mol Biol* 2000;295:963–978.
 55. Yang D, Mittermaier A, Mok Y-K, Kay LE. A study of protein side-chain dynamics from new ^2H auto-correlation and ^{13}C cross-correlation NMR experiments: application to the N-terminal SH3 domain of drk. *J Mol Biol* 1998;276:939–954.
 56. Jin D, Andrec M, Montelione GT, Levy RM. Propagation of experimental uncertainties using the Lipari–Szabo model-free analysis of protein dynamics. *J Biomol NMR* 1998;12:471–492.
 57. Horita DA, Zhang W, Smithgall TE, Gmeiner WH, Byrd RA. Dynamics of the Hck-SH3 domain: comparison of experiment with multiple molecular dynamics simulations. *Protein Sci* 2000;9:95–103.
 58. Chung SY, Subbiah S. Validation of NMR side chain conformations by packing calculations. *Proteins* 1999;35:184–194.
 59. Moy FJ, Lowry DF, Matsumura P, Dahlquist FW, Krywko JE, Dommaille PJ. Assignments, secondary structure, global fold, and dynamics of chemotaxis Y protein using three- and four-dimensional heteronuclear (^{13}C , ^{15}N) NMR spectroscopy. *Biochemistry* 1994;33:10731–10742.
 60. Lavigne P, Crump MP, Gagne SM, Hodges RS, Kay CM, Sykes BD. Insights into the mechanism of heterodimerization from ^1H -NMR solution structure of the c-Myc–Max heterodimeric leucine zipper. *J Mol Biol* 1998;281:165–181.
 61. Liu A, Hu W, Majumdar A, Rosen MK, Patel, DJ. NMR detection of side chain–side chain hydrogen bonding interactions in $^{13}\text{C}/^{15}\text{N}$ -labeled proteins. *J Biomol NMR* 2000;17:305–310.
 62. Parthasarathy S, Murthy MRN. Protein thermal stability: insights from atomic displacement parameters (B values). *Protein Eng* 2000;13:9–13.
 63. Honig B, Nicholls A. Classical electrostatics in biology and chemistry. *Science* 1995;268:1144–1149.