# J|A|C|S
## ARTICLES

# Protein Side-Chain Dynamics and Residual Conformational Entropy

Nikola Trbovic,[‡] Jae-Hyun Cho,[‡] Robert Abel,[†] Richard A. Friesner,[†] Mark Rance,[§] and Arthur G. Palmer III*,[‡]

*Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, Department of Chemistry, Columbia University, New York, New York 10027, and Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati, Cincinnati, Ohio 45267*

Received August 14, 2008; E-mail: agp6@columbia.edu

***Abstract:*** Changes in residual conformational entropy of proteins can be significant components of the thermodynamics of folding and binding. Nuclear magnetic resonance (NMR) spin relaxation is the only experimental technique capable of probing local protein entropy, by inference from local internal conformational dynamics. To assess the validity of this approach, the picosecond-to-nanosecond dynamics of the arginine side-chain $N^\epsilon-H^\epsilon$ bond vectors of *Escherichia coli* ribonuclease H (RNase H) were determined by NMR spin relaxation and compared to the mechanistic detail provided by molecular dynamics (MD) simulations. The results indicate that arginine $N^\epsilon$ spin relaxation primarily reflects persistence of guanidinium salt bridges and correlates well with simulated side-chain conformational entropy. In particular cases, the simulations show that the aliphatic part of the arginine side chain can retain substantial disorder while the guanidinium group maintains its salt bridges; thus, the $N^\epsilon-H^\epsilon$ bond-vector orientation is conserved and side-chain flexibility is concealed from $N^\epsilon$ spin relaxation. The MD simulations and an analysis of a rotamer library suggest that dynamic decoupling of the terminal moiety from the remainder of the side chain occurs for all five amino acids with more than two side-chain dihedral angles (R, K, E, Q, and M). Dynamic decoupling thus may represent a general biophysical strategy for minimizing the entropic penalties of folding and binding.

## Introduction

Protein side-chain dynamics represent a major component of protein conformational entropy.[1] Perturbations of dynamics upon folding or molecular interaction are therefore believed to contribute to the thermodynamics of these processes.[2] Nuclear magnetic resonance (NMR) relaxation of backbone and side-chain spin probes has been used to infer conformational entropy from internal protein dynamics.[2-4] Side-chain methyl C—H bond vectors, the N—H bond vectors in the carboxyamide group of asparagine and glutamine, and the secondary amines ($N^\epsilon-H^\epsilon$) of arginine and tryptophan are routinely used as spin-relaxation probes of local side-chain dynamics.[2,5] Methods for NMR spin-relaxation studies of the dynamics of the terminal groups of other amino acids are in development.[6,7]

The full spectrum of possible enthalpy—entropy partitioning has been experimentally observed when using NMR spin relaxation to probe changes in entropy upon biologically relevant transitions: (1) enthalpically driven processes with global rigidification of the protein; (2) allosteric entropic compensation, in which the binding site becomes more rigid while more distal regions become more flexible; and (3) entirely entropically driven processes with apparently increased global conformational entropy.[4,8]

Arginine side chains often participate in molecular recognition by proteins, because the terminal positively charged guanidinium group affords multiple salt-bridge geometries.[9,10] The long arginine side chain also can bear substantial residual conformational entropy. Arginine $^{15}N^\epsilon$ spin relaxation has been used to investigate arginine side-chain dynamics with applications to protein—protein[11] and protein—nucleic acid[3,11-14] interactions. The underlying structural determinants and mechanisms

‡ Department of Biochemistry and Molecular Biophysics, Columbia University.
† Department of Chemistry, Columbia University.
§ Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati.

(1) Pickett, S. D.; Sternberg, M. J. *J. Mol. Biol.* **1993**, *231*, 825–839.
(2) Igumenova, T. I.; Frederick, K. K.; Wand, A. J. *Chem. Rev.* **2006**, *106*, 1672–1699.
(3) Berglund, H.; Baumann, H.; Knapp, S.; Ladenstein, R.; Härd, T. *J. Am. Chem. Soc.* **1995**, *117*, 12883–12884.
(4) Jarymowycz, V. A.; Stone, M. J. *Chem. Rev.* **2006**, *106*, 1624–1671.
(5) Kempf, J. G.; Loria, J. P. *Cell Biochem. Biophys.* **2003**, *37*, 187–211.
(6) Iwahara, J.; Jung, Y. S.; Clore, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 2971–2980.

(7) Paquin, R.; Ferrage, F.; Mulder, F. A.; Akke, M.; Bodenhausen, G. *J. Am. Chem. Soc.* **2008**, *130*, 15805–15807.
(8) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448*, 325–329.
(9) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1–9.
(10) Hoffman, M. M.; Khrapov, M. A.; Cox, J. C.; Yao, J.; Tong, L.; Ellington, A. D. *Nucleic Acids Res.* **2004**, *32*, D174–181.
(11) Pascal, S. M.; Yamazaki, T.; Singer, A. U.; Kay, L. E.; Forman-Kay, J. D. *Biochemistry* **1995**, *34*, 11353–11362.
(12) Slijper, M.; Boelens, R.; Davis, A. L.; Konings, R. N.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *Biochemistry* **1997**, *36*, 249–254.

of arginine side-chain motion, however, remain obscure. Both side-chain solvent accessibility[15] and electrostatic interactions[11–13,16] have been proposed as the primary structural determinants of arginine side-chain dynamics. The lack of mechanistic insight results from the limited atomistic detail afforded by NMR spin-relaxation parameters. Several studies of protein−substrate complexes reported substantial $N^\epsilon$ flexibility for arginine side chains believed to be involved in specific substrate interactions.[3,12,13] The observations were interpreted as indicative of either alternative electrostatic interactions or stable interactions with concerted motions of protein and ligand. An analogous observation was made for lysine side chains in a protein−DNA complex and attributed to alternative hydrogen-bond acceptors based on the NMR-derived structure.[17] In contrast, Forman-Kay and co-workers observed broadening of aliphatic resonances for arginine side chains with essentially rigid $N^\epsilon$−$H^\epsilon$ bond vectors in a protein−peptide complex.[11] They offer three explanations: exchange between hydrogen-bonded and non-hydrogen-bonded conformations, alternative hydrogen-bond acceptors, or stable $N^\epsilon$ interactions with aliphatic disorder accommodating internal relocation of the complex.

Molecular dynamics (MD) simulations provide a full atomistic description of protein motions. The method is therefore perfectly poised for mechanistic interpretation of NMR spin-relaxation data, provided that quantitative agreement and thus experimental validation of the simulations is achieved.[4,18–21] In addition to biophysical insight, MD simulations allow investigation of the implicit assumptions of spin-relaxation-based entropy determination, such as homogeneity of local dynamics and uncoupling of local motions.[22,23] An early study comparing spin-relaxation parameters and MD simulations divided asparagine, glutamine, and arginine side-chain dynamics into three classes, depending upon the value of the generalized order parameter $S^2$, which reports on the amplitude of motion of the bond vector under investigation on a time scale typically less than or similar to the time scale of overall rotational tumbling of the protein.[2] The three classes represented no dihedral disorder ($S^2 > 0.7$), one disordered side-chain dihedral angle ($0.7 \geq S^2 \geq 0.3$), and two or more disordered side-chain dihedral angles ($S^2 < 0.3$).[24] This partitioning suggests quantitative correlation between NMR spin-relaxation probes of side-chain dynamics and side-chain conformational entropy.

The 18 kDa nucleotidyl-transferase *Escherichia coli* ribonuclease H (RNase H) has been the subject of several NMR[25–30] and MD[31–33] studies of protein dynamics. Here we report a comparison between arginine $N^\epsilon$ spin relaxation at four static magnetic fields and extensive MD simulations for RNase H. Good agreement between simulated and experimental order parameters encourages use of the simulations for atomistic interpretation of the experimental data. The analysis confirms that $N^\epsilon$ dynamics are primarily sensitive to the presence of salt bridges formed by the guanidinium group. Importantly, the arginine side chain can occupy several rotameric states while forming stable salt bridges, thereby retaining substantial residual entropy. The simulations together with analysis of a rotamer library[34] suggest that dynamic decoupling of the side-chain terminus from the aliphatic part can be expected for all amino acids with more than two side-chain dihedral angles. The results imply that spin-relaxation-based determination of side-chain conformational entropy is limited by inherent sensitivity to local interactions, and that dynamic decoupling may be a general biophysical strategy for maximizing residual side-chain conformational entropy.

## Materials and Methods

**1. NMR Spectroscopy.** $R_1$, $R_2$, and {$^1$H}−$^{15}$N nuclear Overhauser enhancement (NOE) parameters were measured for the 10 arginine $N^\epsilon$−$H^\epsilon$ moieties in a 0.75 mM sample of [U-$^2$H,$^{15}$N] RNase H at pH 5.5 and 300 K using established sensitivity-enhanced gradient-selected pulse sequences[35] (Tables S1−S3, Supporting Information). $R_1$ and NOE were measured at 9.4, 11.7, 14.1, and 16.4 T. $R_1$ experiments used the following time points: 0.01, 0.2, 0.4, 0.8, and 1.6 s at 9.4 T; 0.01, 0.25, 0.5, 1, and 2 s at 11.7 T; 0.01, 0.25, 0.5, 1.25, and 2.5 s at 14.1 T; and 0.01, 0.3, 0.75, 1.5, and 3 s at 16.4 T. NOE experiments used saturation periods of 4 s at 9.4 and 11.7 T, 5.8 s at 14.1 T, and 7.8 s at 16.4 T; control experiments used a 12 s recycle delay. Three NOE experiments were conducted at each field, and average NOE rates were used for the analysis. $R_2$ rates were measured with Carr−Purcell−Meiboom−Gill experiments at 9.4, 11.7, and 14.1 T, using a spin−echo delay of 1.0 ms and time points at 0.006, 0.03, 0.06, 0.1, and 0.15 s. Two spectra were acquired for each time point in all $R_1$ and $R_2$ experiments. The $R_2$ experiment at 14.1 T was repeated, and average rates from the two experiments were used for the analysis. Rates were extracted from spectra processed with an exponential line broadening of 6 Hz in $F_2$ and a cosine bell in $F_1$ for all arginine $N^\epsilon$−$H^\epsilon$ resonances. For the unassigned resonance near that of R138, Lorentzian-to-Gaussian resolution enhancement was used in $F_2$, with an exponential line broadening of −6 Hz and a Gaussian broadening of 8 Hz. Errors were derived from jackknife

(13) Wilkinson, T. A.; Botuyan, M. V.; Kaplan, B. E.; Rossi, J. J.; Chen, Y. *J. Mol. Biol.* **2000**, *303*, 515–529.
(14) Wilkinson, T. A.; Zhu, L.; Hu, W.; Chen, Y. *Biochemistry* **2004**, *43*, 16153–16160.
(15) Buck, M.; Boyd, J.; Redfield, C.; MacKenzie, D. A.; Jeenes, D. J.; Archer, D. B.; Dobson, C. M. *Biochemistry* **1995**, *34*, 4041–4055.
(16) Cai, M.; Huang, Y.; Prakash, O.; Wen, L.; Dunkelbarger, S. P.; Huang, J. K.; Liu, J.; Krishnamoorthi, R. *Biochemistry* **1996**, *35*, 4784–4794.
(17) Foster, M. P.; Wuttke, D. S.; Radhakrishnan, I.; Case, D. A.; Gottesfeld, J. M.; Wright, P. E. *Nat. Struct. Biol.* **1997**, *4*, 605–608.
(18) Maragakis, P.; Lindorff-Larsen, K.; Eastwood, M. P.; Dror, R. O.; Klepeis, J. L.; Arkin, I. T.; Jensen, M. O.; Xu, H.; Trbovic, N.; Friesner, R. A.; Palmer, A. G.; Shaw, D. E. *J. Phys. Chem. B* **2008**, *112*, 6155–6158.
(19) Showalter, S. A.; Brüschweiler, R. *J. Chem. Theory Comput.* **2007**, *3*, 961–975.
(20) Showalter, S. A.; Brüschweiler-Li, L.; Johnson, E.; Zhang, F.; Brüschweiler, R. *J. Am. Chem. Soc.* **2008**, *130*, 6472–6478.
(21) Trbovic, N.; Kim, B.; Friesner, R. A.; Palmer, A. G. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 684–694.
(22) Akke, M.; Brüschweiler, R.; Palmer, A. G. *J. Am. Chem. Soc.* **1993**, *115*, 9832–9833.
(23) Prompers, J. J.; Brüschweiler, R. *J. Phys. Chem. B* **2000**, *104*, 11416–11424.
(24) Smith, L. J.; Mark, A. E.; Dobson, C. M.; van Gunsteren, W. F. *Biochemistry* **1995**, *34*, 10918–10931.

(25) Butterwick, J. A.; Palmer, A. G. *Protein Sci.* **2006**, *15*, 2697–2707.
(26) Butterwick, J. A.; Patrick Loria, J.; Astrof, N. S.; Kroenke, C. D.; Cole, R.; Rance, M.; Palmer, A. G. *J. Mol. Biol.* **2004**, *339*, 855–871.
(27) Kroenke, C. D.; Loria, J. P.; Lee, L. K.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 7905–7915.
(28) Kroenke, C. D.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 10119–10125.
(29) Mandel, A. M.; Akke, M.; Palmer, A. G., III *J. Mol. Biol.* **1995**, *246*, 144–163.
(30) Mandel, A. M.; Akke, M.; Palmer, A. G., III *Biochemistry* **1996**, *35*, 16009–16023.
(31) Philippopoulos, M.; Lim, C. *J. Mol. Biol.* **1995**, *254*, 771–792.
(32) Philippopoulos, M.; Lim, C. *Proteins: Struct., Funct., Genet.* **1999**, *36*, 87–110.
(33) Philippopoulos, M.; Mandel, A. M.; Palmer, A. G.; Lim, C. *Proteins: Struct., Funct., Genet.* **1997**, *28*, 481–493.
(34) Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 389–408.
(35) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy*; 2nd ed.; Academic Press: San Diego, CA, 2007.

calculations. Spin-relaxation data were analyzed using reduced spectral density mapping[36] with an N–H bond length of 1.04 Å.[37] Data measured at the three lower fields were used to estimate the spectral density $J(0)$, while the parameters measured at the three higher fields were used for the higher frequency components of the spectral density function, $J(\omega_N)$ and $J(0.87\omega_H)$.[26,27] Spectral densities of all arginines except R46 and R106 were fitted using two order parameters and an internal correlation time, while R46 and R106 required only one order parameter and an internal correlation time. Fitting used the nonlinear least-squares regression algorithm implemented in MATLAB. The correlation times for overall tumbling were obtained using a previously reported axial diffusion tensor.[38]

**2. MD Simulations.** The crystal structure of RNase H (PDB code 2RN2),[39] 9387 TIP3P[40] water molecules, and seven Cl⁻ ions were placed in an orthorhombic box to accommodate a minimum water shell thickness of 1 nm, with the protein described by the AMBER99SB force field.[41] The system was energy minimized in two steps by first minimizing only the solvent and then minimizing the whole system. The system was equilibrated to 300 K over 110 ps at constant temperature and constant pressure (NPT), using periodic boundary conditions and a cutoff of 1 nm for both particle-mesh-Ewald and Lennard-Jones interactions; protein coordinates were constrained for the first 10 ps. The SHAKE[42] algorithm was applied for constraining vibrations of bonds involving hydrogen atoms, and a time step of 1 fs was used. The system was then equilibrated for a further 1 ns at NPT. Eight structures with temperatures close to 300 K were extracted from the equilibration and used as starting structures for eight 20 ns constant energy (NVE) production runs, using identical simulation parameters. One of these simulations was extended to 117 ns. The procedure was repeated with the TIP4P water model,[40] yielding an additional set of eight 20 ns trajectories. All unconstrained simulations were conducted with the pmemd module, and all constrained simulations were conducted with the sander module in the AMBER 9.0 molecular dynamics package.

**3. Simulated Order Parameters.** Simulated generalized order parameters were calculated from

$$S^2 = \frac{1}{2}\left(3\sum_{i=1}^{3}\sum_{j=1}^{3}\langle\mu_i\mu_j\rangle^2 - 1\right) \quad (1)$$

in which $\mu_1$, $\mu_2$, and $\mu_3$ are the $x$, $y$, and $z$ components of the relevant bond vector scaled to unit magnitude, $\mu$, respectively.[43] Angular brackets indicate averaging over the snapshots in a 10 ns block, after superposition of backbone heavy atoms to remove the effects of overall tumbling. The block length is chosen to mimic the time scale limitation of NMR spin relaxation (RNase H has an overall tumbling time of 9.7 ns).[18] To assess the separability of internal and overall protein motion, the rotationally invariant iRED approach[44] also was used to calculate order parameters from 10 ns blocks. The results obtained from the two methods agree within errors (not shown).

**4. Conformational Entropy.** Arginine side-chain conformational entropy was calculated from dihedral-angle distributions sampled during the simulations. In principle, this entropy can be computed exactly from computer simulations by numerically evaluating

$$S = -R \int P(\bar{\chi}) \ln P(\bar{\chi})\, d\bar{\chi} \quad (2)$$

in which $P(\bar{\chi})$ is the four-dimensional probability density function of the four dihedral angles of the arginine side chain.[45] Recent increases in computational resources allow direct computation of adequately converged estimates of $P(\bar{\chi})$ without the need for approximations of the underlying form and structure of the probability density function.

In this study $P(\bar{\chi})$ was approximated with a four-dimensional histogram of the dihedral-angle distribution sampled *in silico*. This discrete probability density function then was used to numerically evaluate eq 2. The choice of units for measuring dihedral angles when evaluating eq 2 affects only absolute entropy values, not differences. In the present work, dihedral angles were measured in degrees, yielding a maximum entropy of 46.78 cal/(mol·K) for a uniform dihedral-angle distribution. The dependence of the computed entropies on the bin size of the integration mesh is shown in Figure S1 (Supporting Information). Standard errors of the computed entropies were estimated with the bootstrap method (10 000 random bootstrap samples per computed entropy).[46]

**5. Hydrogen Bonds and Salt Bridges.** A hydrogen bond was considered formed between a donor hydrogen atom (attached to a nitrogen or oxygen atom) and an acceptor atom (nitrogen or oxygen atom) when the distance between donor hydrogen and acceptor was less than 3.0 Å, and simultaneously the angle between donor, donor hydrogen, and acceptor as well as the angle between donor hydrogen, acceptor, and the acceptor-bound atom were larger than 100°.[47] A given hydrogen atom was considered to participate in a salt bridge if it formed a hydrogen bond with the salt-bridge acceptor. Because the hydrogen atoms of the terminal amino group of lysine are equivalent, a lysine salt bridge was considered formed if its terminal nitrogen atom was closer than 4 Å to its partner oxygen.[48] The results are not significantly affected if this cutoff is reduced to 3.2 Å.

**6. Rotamer Nomenclature.** Rotameric conformations are labeled according to their four $\chi$ angles: p (gauche positive), ~60°; t (trans), ~180°; m (gauche negative), ~300°. For terminal $\chi$ angles in residues with planar functional groups, where the ideal rotamers deviate from these values, the mean value rounded to the nearest 10° is given.[34]

## Results

**1. $N^\epsilon$–$H^\epsilon$ Flexibility.** Flexibilities of the 10 arginine side-chain $N^\epsilon$–$H^\epsilon$ bond vectors of *E. coli* RNase H (Figure 1a) were examined using NMR spin relaxation and MD simulations. Resonances of all $N^\epsilon$–$H^\epsilon$ moieties are well resolved in $^1H$–$^{15}N$ correlation spectra (Figure 1b). Generalized order parameters and effective correlation times were obtained for all $N^\epsilon$ sites using reduced spectral density mapping. An average chemical shift anisotropy (CSA) value of −114.0 ppm was determined from the field dependence of the relaxation rate constants.[28] $N^\epsilon$ order parameters were also calculated from eight 20 ns MD simulations of the protein conducted using the AMBER99SB force field in both TIP3P and TIP4P explicit water. One of the TIP3P trajectories was extended to 117 ns. $N^\epsilon$ order parameters

(36) Farrow, N. A.; Zhang, O. W.; Szabo, A.; Torchia, D. A.; Kay, L. E. *J. Biomol. NMR* **1995**, *6*, 153–162.

(37) Case, D. A. *J. Biomol. NMR* **1999**, *15*, 95–102.

(38) Kroenke, C. D. Ph.D. thesis, Columbia University, 2001.

(39) Katayanagi, K.; Miyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Nakamura, H.; Ikehara, M.; Matsuzaki, T.; Morikawa, K. *J. Mol. Biol.* **1992**, *223*, 1029–1052.

(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(41) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

(42) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

(43) Chandrasekhar, I.; Clore, G. M.; Szabo, A.; Gronenborn, A. M.; Brooks, B. R. *J. Mol. Biol.* **1992**, *226*, 239–250.

(44) Prompers, J. J.; Brüschweiler, R. *J. Am. Chem. Soc.* **2002**, *124*, 4522–4534.

(45) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325–332.

(46) Efron, B.; Tibshirani, R. J. *An Introduction to the Bootstrap*; Chapman & Hall/CRC: Boca Raton, FL, 1993.

(47) Fabiola, F.; Bertram, R.; Korostelev, A.; Chapman, M. S. *Protein Sci.* **2002**, *11*, 1415–1423.

(48) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1983**, *168*, 867–885.
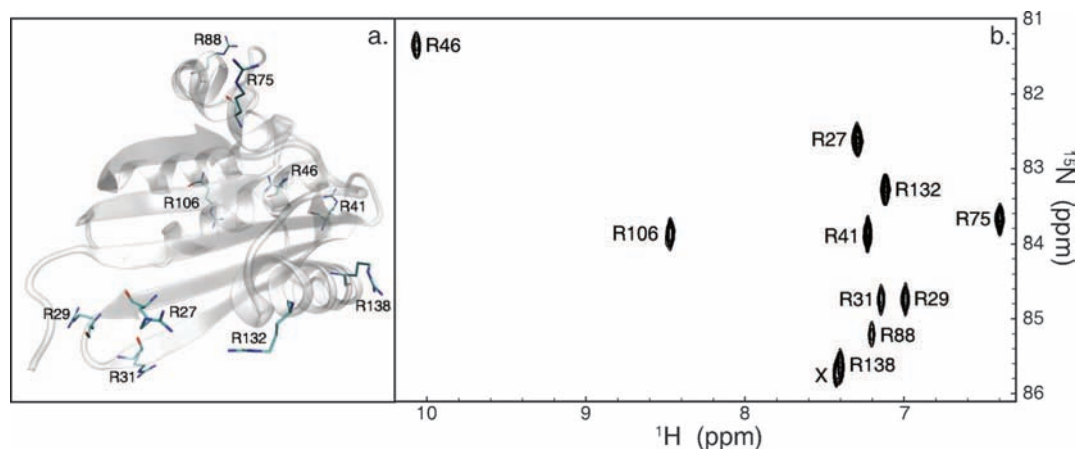
**Figure 1.** RNase H arginine residues. (a) Location of arginines in the crystal structure of RNase H (PDB code 2RN2).[39] (b) Section of the 10 ms time point spectrum from the 16.4 T $R_1$ experiment showing the arginine $N^\epsilon{-}H^\epsilon$ resonances of RNase H. X indicates a minor unassigned resonance (see text).
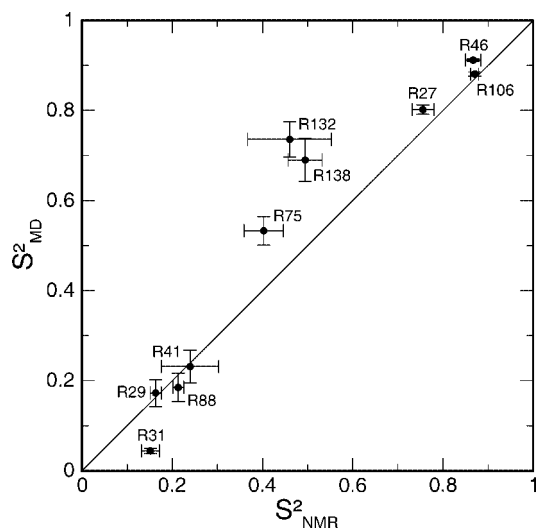


**Figure 2.** Experimental vs simulated arginine $N^\epsilon$ order parameters. Experimental error bars were derived as described in Materials and Methods. Simulated error bars indicate the standard error of the mean over sixteen 10 ns blocks.

agree within errors for simulations using the two water models. The choice of explicit solvent model is known to influence overall tumbling significantly; however, the effects of the solvent model on internal dynamics are a matter of ongoing research.[49] The thoroughly validated combination of AMBER99SB force field and TIP3P water model therefore was chosen for comparison with NMR spin-relaxation data on timescales faster than the experimentally determined tumbling time of RNase H.

Experimental and simulated $N^\epsilon$ order parameters agree well (correlation coefficient of 0.94) and cover a wide range of motional amplitudes (Figure 2). All comparisons between simulation and experiment refer to the eight 20 ns trajectories, unless otherwise stated. The unassigned resonance near that of R138 may represent a minor conformation of this side chain, with a population of approximately 20% based on intensity differences in the control spectra of the 16.4 T NOE experiment. The resonance has an experimental order parameter of $0.31 \pm 0.02$, even lower than the $N^\epsilon$ order parameter for R138, and was not investigated further.

**2. Structural Determinants.** As observed previously[15] and shown in Figure 3a, relative side-chain solvent accessibility (rsa),

calculated from the crystal structure of RNase H (PDB code 2RN2)[39] using NACCESS,[50] correlates with $N^\epsilon$ flexibility (correlation coefficient of $-0.81$). Some sites, however, most notably R106, deviate from this trend. The high rsa and $N^\epsilon$ rigidity of R106 are reconciled by the presence of a solvent-exposed salt bridge to the carboxylate of E57, formed 85% of the time in the simulations. $N^\epsilon$ salt bridges indeed appear to be an important factor of local dynamics (Figure 3b). The additional restriction observed for R132 and R138 can be explained by $N^\eta$ salt bridges, which can restrict the $N^\epsilon{-}H^\epsilon$ bond-vector orientation in the absence of prevalent $N^\epsilon$ interactions. The $N^\eta$ groups of R132 form a salt bridge to the carboxylate of E32 91% of the time, while the $N^\eta$ groups of R138 form a salt bridge to the carboxylate of D134 94% of the time *in silico*.

The intermediate flexibility of R75, however, cannot be explained by guanidinium salt bridges. The aliphatic part of the side chain of R75 maintains hydrophobic packing with W120 *in silico*, thereby somewhat restricting $N^\epsilon{-}H^\epsilon$ flexibility in the absence of any prevalent guanidinium group interactions.

**3. Order Parameters and Entropy.** Simulated side-chain conformational entropies calculated from eq 2 are shown in Figure S1 (Supporting Information). Although computed entropies decrease uniformly with decreasing bin size of the integration mesh, the correlation coefficient between entropies computed with a 20° mesh and those computed with a 5° mesh is 0.99, suggesting that this bias does not systematically alter relative side-chain entropies. In addition, the standard error of the computed entropies, as estimated by bootstrap statistics, which do not include the error due to the bias introduced by the integration bin size, is calculated to be less than 0.1 cal/(mol·K) for each of the computed entropies, regardless of bin size. These data indicate that torsional sampling *in silico* was sufficient to converge the four-dimensional probability density functions, and that qualitative trends in computed entropies are invariant to bin size.

Simulated side-chain conformational entropies correlate with $N^\epsilon$ order parameters (correlation coefficient of $-0.94$ for simulated and $-0.86$ for experimental order parameters and 20° entropy data, Figure 4). The correlation is poor, however, among the six side chains with $S^2_{MD} > 0.5$ (correlation coefficient of $-0.59$). R75, R132, and R138 exhibit lower side-chain entropy than R27, although they have lower $N^\epsilon$ order parameters than

(49) Wong, V.; Case, D. A. *J. Phys. Chem. B* **2008**, *112*, 6013–6024.

(50) Hubbard, S. J., Thornton, J. M., Department of Biochemistry and Molecular Biology, University College London, 1993.
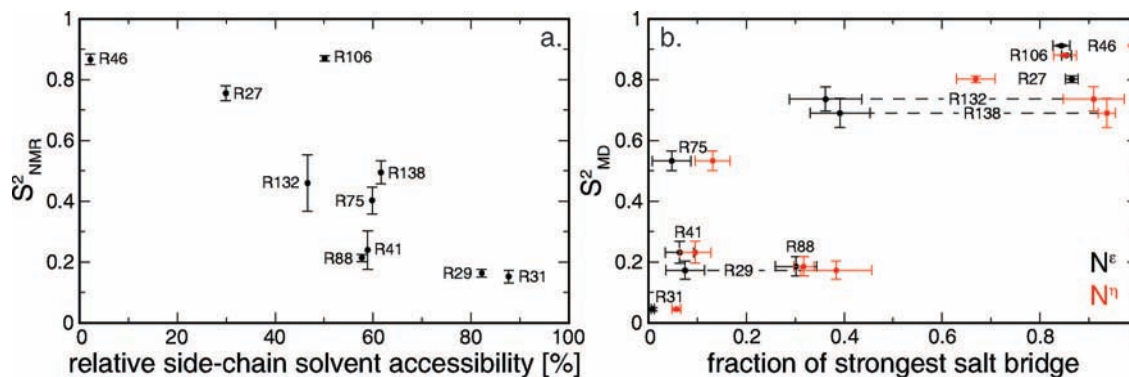
**Figure 3.** Structural determinants of $N^\epsilon$ flexibility. (a) Relative side-chain solvent accessibility, calculated from the crystal structure of RNase H (PDB code 2RN2)[39] using NACCESS,[50] vs experimental order parameters. (b) Average maximal fraction of time the $N^\epsilon-H^\epsilon$ moiety (black) or the two $N^\eta-H^\eta$ moieties (red) of an arginine participate in a salt bridge during a 20 ns trajectory. Standard errors of the mean of the salt-bridge fractions calculated over the eight 20 ns trajectories are shown. Dashed lines connect data points belonging to the same residue when needed for clarity.
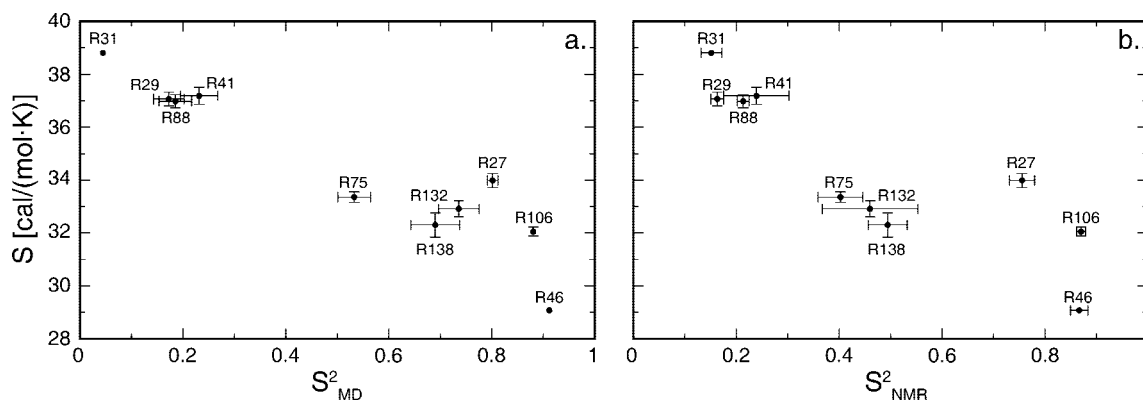


**Figure 4.** Simulated (a) and experimental (b) arginine $N^\epsilon$ order parameters vs side-chain conformational entropy computed using a 20° integration mesh. Bootstrap standard errors of the entropies are shown. Error bars not shown are similar to or smaller than the plotted points.

R27. Notably, R75, with a simulated $N^\epsilon$ order parameter of 0.53, displays slightly lower side-chain entropy than R27, with the third highest simulated $N^\epsilon$ order parameter of 0.80.

**4. Structural Analysis.** The side chains of both R27 and R75 occupy several rotameric states *in silico* (Figure 5a,b). The side chain of R27 undergoes concerted rotameric flips in $\chi_1$ and $\chi_3$ that keep its $N^\epsilon-H^\epsilon$ bond vector pointing in one direction (indicated by color coding of conformations in Figure 5). R27 therefore retains substantial side-chain conformational entropy while maintaining its salt bridge to the carboxylate of E6 (Figure 5a). The $\chi_1$ and $\chi_2$ dihedral angles of R75, on the other hand, are essentially constrained to the t rotamer by hydrophobic packing with the aromatic ring of W120 ($\chi_1$ is always t, $\chi_2$ is t 98% of the time). The guanidinium group of R75 is solvent exposed, and its $\chi_3$ and $\chi_4$ angles are disordered (Figure 5b). R75 therefore exhibits a relatively low $N^\epsilon$ order parameter, while its conformational side-chain entropy is substantially reduced compared to completely solvent-exposed arginine side chains.

Analysis of the other two arginine side chains with essentially rigid $N^\epsilon-H^\epsilon$ bond vectors, R46 and R106, reveals that R27 is not unique in its ability to maintain side-chain disorder in a salt bridge (Figure 5c,d). The solvent-exposed side chain of R106 interconverts freely in all simulations between two rotameric states which preserve $N^\epsilon-H^\epsilon$ orientation and thus its salt bridge. R106 undergoes concerted flips in $\chi_1$ and $\chi_3$, in analogy to R27 (Figure 5c). The side chain of R46, however, is buried. It remains in one rotameric state (mtp70) throughout the eight 20 ns simulations, but it occupies another state (ttt80) for the second half of the extended 117 ns simulation. This

second state also is accessed through concerted $\chi_1-\chi_3$ flips in order to maintain the salt-bridge interactions (Figure 5d). The side-chain conformational entropy of R46, calculated from the 117 ns trajectory, increases by 1.73 cal/(mol·K), further reducing the $S^2$-entropy correlation among the rigid sites (correlation coefficient of −0.57).

Arginine side chains engaged in $N^\eta$-dominated salt bridges show behavior similar to those in $N^\epsilon$ salt bridges. The main difference is the possibility of alternative salt-bridge configurations. The side chain of R132, for example, can form its salt bridge to E32 through its $N^\eta$ groups in two possible ways (conformations labeled in red and blue, respectively, in Figure 6). In each configuration, the side chain occupies several rotameric states, as observed for $N^\epsilon$ salt bridges. In contrast to $N^\epsilon$ salt bridges, however, interconversion between $N^\eta$ salt-bridge configurations reorients the $N^\epsilon-H^\epsilon$ vector substantially. Analogous behavior is observed for R138 (Figure S2, Supporting Information). The conformational space sampled by a solvent-exposed arginine side chain (R41) is shown in Figure S3 for comparison (Supporting Information).

Analysis of a rotamer library confirms the degeneracy of the conformational space of arginine with respect to $N^\epsilon-H^\epsilon$ orientation.[34] Twenty-seven of the 28 conformations in the library can be grouped into six clusters with $N^\epsilon-H^\epsilon$ orientations varying less than 30° between member conformations. Two of these clusters are shown in Figure 7. For any one of the five arginines forming salt bridges with $N^\epsilon$ participation, the majority of conformations sampled *in silico* fall into one of the six clusters. The remaining conformations deviate in at least one $\chi$
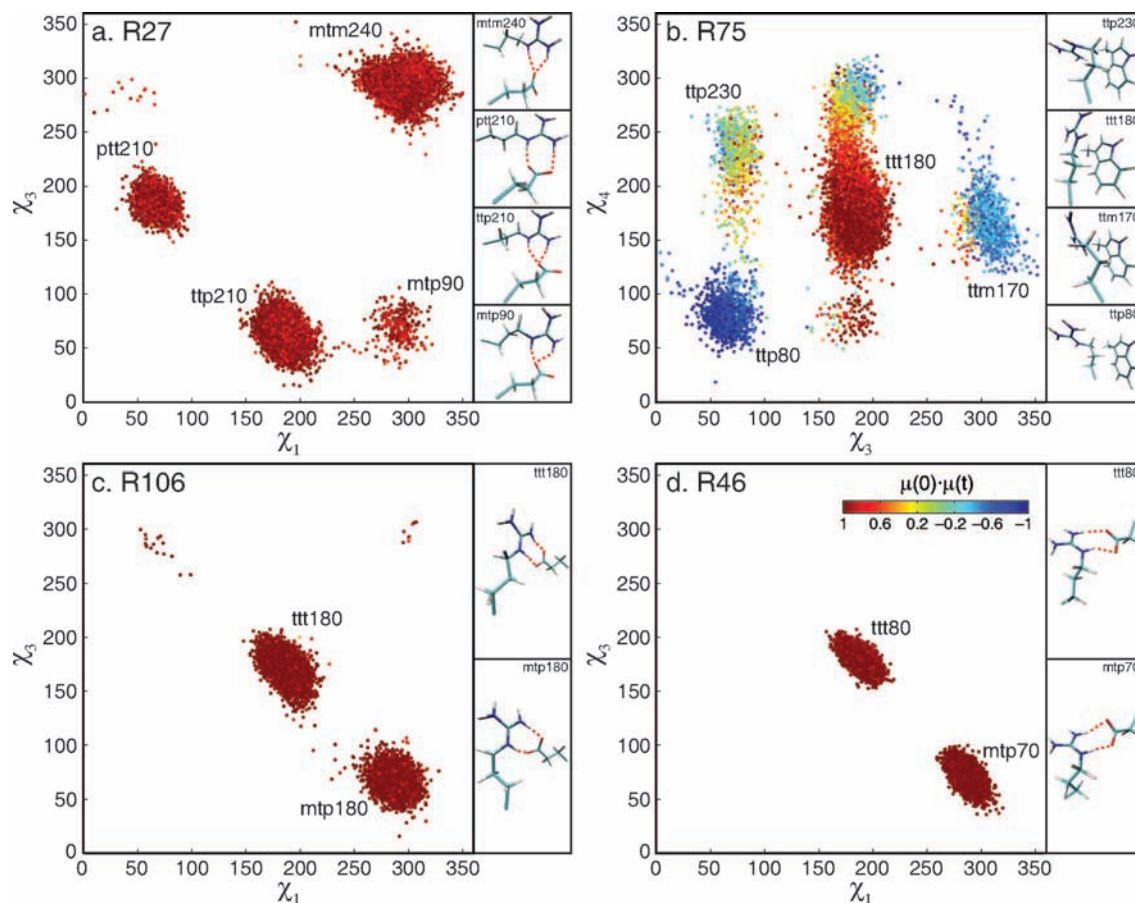
**Figure 5.** Two-dimensional χ-angle plots for select arginine residues. For R27, R75, and R106, every 10th conformation from all eight 20 ns trajectories is shown; for R46, every 7th conformation from the extended 117 ns trajectory is shown. Conformations are color-coded with respect to the scalar product of the orientation of the $N^\epsilon$–$H^\epsilon$ bond vector and its orientation in the first snapshot according to the color bar in the plot of R46. Representative structures of the respective arginine side chain and its interaction partner are shown, with salt bridges indicated by red dashed lines.
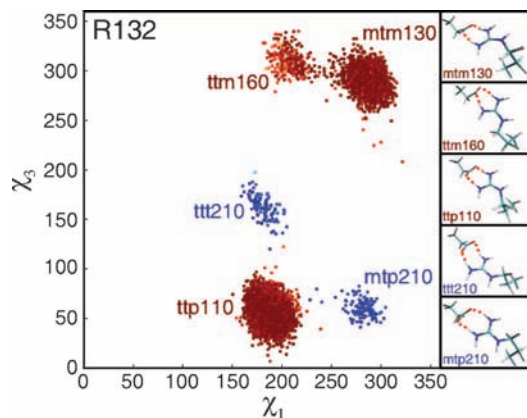


**Figure 6.** Two-dimensional χ-angle plot for R132. For clarity, every 10th conformation forming the salt bridge to E32 from all eight 20 ns trajectories is shown. Conformations are color-coded as in Figure 5. In addition, the salt-bridge configuration of each conformation is indicated by the color of the corresponding rotamer label (red and blue). Representative structures are shown, with salt bridges indicated by red dashed lines.

angle from the ideal rotamer in order to allow for $N^\epsilon$ salt-bridge participation. In all cases, the deviations are within the bounds observed in the raw data used to generate the library. Deviations from ideal rotamers were not considered in the clustering of the rotamer library.

**5. Other Side Chains.** Populations of several rotamers in a salt bridge also are observed for all glutamate residues engaging
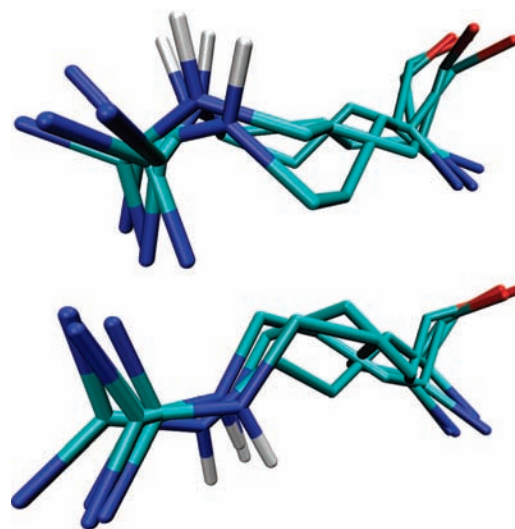


**Figure 7.** Two groups of arginine rotamer-library conformations clustered according to $N^\epsilon$–$H^\epsilon$ bond-vector orientation after joint heavy-atom superposition.

in such interactions. All of these form salt bridges with arginines in RNase H (Figure S4a–c, Supporting Information). The ability of glutamates to maintain rotameric disorder in salt bridges is predicted by the rotamer library. When the $\chi_2$ angle of the glutamate side chain is in the t rotamer, $\chi_3$ is largely unconstrained within a basin approximately 180° wide.[34] All three

conformations with $\chi_2$ in the t rotamer show similar $C^\gamma-C^\delta$ bond orientations and consequently can maintain the same salt bridge (Figure S4d). Most salt bridges examined show little association between side-chain dynamics of salt-bridge partners *in silico* (Figure S4a,b). The maximum association observed amounts to a Cramer's $V$ of 0.36, using products of individual dihedral-angle distributions as null hypotheses (Figure S4c).[51]

The behavior of glutamine side chains is analogous to that of glutamate side chains, despite the lack of symmetry in the carboxyamide group. Q76, the only glutamine residue engaged in a persistent hydrogen bond *in silico* (formed 51.9% of the time), populates several rotamers while maintaining the hydrogen bond (Figure S5a, Supporting Information). For glutamines, like glutamates, $\chi_3$ is unconstrained when $\chi_2$ is in the t rotamer. The corresponding cluster found in the rotamer library is shown in Figure S5b. The order parameter calculated over all hydrogen-bonded conformations of the relevant N–H bond is $0.76 \pm 0.02$.

Both lysine side chains forming persistent salt bridges *in silico* ($>50\%$ of the time), K86 and K96, populate several rotameric states while maintaining terminal salt bridges. K86 forms a salt bridge with D108 98.9% of the time *in silico*. Most of the conformational space sampled by K86 (red conformations in Figure S6a, Supporting Information) corresponds to a cluster of rotamer-library conformations with similar $C^\epsilon-N^\zeta$ bond orientations (Figure S6b(i)). Three additional conformations, however, are able to form the salt bridge with altered $C^\epsilon-N^\zeta$ bond orientations (blue conformations in Figure S6a), reflecting the lack of orientational constraint on the amino group in a salt bridge due to the three equivalent amino hydrogen atoms. The simulated order parameter for the $C^\epsilon-N^\zeta$ bond of K86 is $0.57 \pm 0.04$.

No analogous behavior is observed for methionine side chains *in silico*. Interestingly, the methionine conformations found in the rotamer library can be clustered with respect to $S^\delta-C^\epsilon$ bond orientation, indicating that the methionine side chain is able to keep its terminal methylsulfanyl group fixed while retaining rotameric disorder (Figure S7, Supporting Information). Although methionine methyl order parameters are usually low, in calmodulin methionines account for some of the highest methyl order parameters and exhibit the most pronounced increases in methyl order parameters upon ligand binding.[2]

Neither the simulations nor the rotamer library contains any evidence that amino acids with less than three side-chain dihedral angles have similar capabilities, in agreement with the established correlation between methyl order parameters and rotamer jumps for all methyl-bearing side chains except methionine.[2,52]

## Discussion

The results presented herein indicate that arginine $N^\epsilon$ dynamics are governed primarily by salt bridges of the guanidinium group and secondarily by packing interactions. Furthermore, the data show that the degrees of freedom of the arginine side chain enable it to retain substantial disorder while maintaining specific interactions through its charged terminus. Dynamic decoupling in arginines can occur even when the side chain is completely buried, as in the case of R46 (2.2% rsa). A significant fraction of salt bridges are buried in proteins, and the majority of these buried salt bridges appear functionally important.[48]

The effect of terminal interactions on $N^\epsilon$ spin relaxation depends on salt bridge geometry: $N^\epsilon$ salt bridges necessarily restrict $N^\epsilon-H^\epsilon$ orientation, but $N^\eta$ salt bridges may or may not do so. Interconversion between $N^\eta$ salt-bridge configurations, which reorients the $N^\epsilon-H^\epsilon$ bond vector, must be faster than overall tumbling in order to affect $N^\epsilon$ spin relaxation. Convergence of the internal correlation functions for R132 and R138 is incomplete and indicates that the reconfiguration time scales must be at least on the order of overall tumbling for both residues. Thus, larger $S^2_{MD}$ values for R132 and R138 may stem from excessive energy barriers for transiently breaking the salt bridges *in silico*. $N^\epsilon$ order parameters calculated from the full conformational ensemble of the eight 20 ns simulations are reduced to 0.62 and 0.51 for R132 and R138, respectively, yielding significantly better agreement with experiment. Thus, a lower barrier and consequently faster reconfiguration would result in a higher contribution of the associated $N^\epsilon-H^\epsilon$ reorientation to spin relaxation and would reduce the order parameter. A detailed analysis of the discrepancy, however, is outside the scope of this work, not only due to sampling limitations but also due to the large number of factors possibly affecting the stability of salt bridges *in silico* (including protein parametrization, solvent model, ionic strength, and deviations from classical electrostatics). On the experimental side, fast reconfiguration of $N^\eta$ salt bridges could explain low $N^\epsilon$ order parameters measured previously for arginines with putatively specific substrate interactions in complexes, without the requirement of alternate hydrogen-bonding patterns or concerted motions of protein and ligand.[3,12,13]

Both R132 and R138 show decoupling of $N^\epsilon$ dynamics from rotameric disorder within each of two $N^\eta$ salt-bridged configurations. In analogy to arginines forming $N^\epsilon$ salt bridges, local interactions of the guanidinium group restrict $N^\epsilon-H^\epsilon$ orientation, while the aliphatic part can populate several rotameric states. Motional independence of the terminus from the rest of the side chain suggests that the aliphatic broadening in arginine residues with high $N^\epsilon$ order parameters observed by Forman-Kay and co-workers can be explained by stable salt bridges without the need for coupled motions at the complex interface.[11] Instead, interactions of the terminal group introduce coupling between dihedral angles of the arginine side chain. The strongest dihedral-angle association observed *in silico* across a salt bridge, i.e., between an arginine and a glutamate side chain, amounts to a Cramer's $V$ of 0.36 (R27 $\chi_3$ – E6 $\chi_1$), whereas the strongest association within an arginine side chain amounts to a Cramer's $V$ of 1.00 (R46 $\chi_1-\chi_3$ from the extended simulation). The majority of strong dihedral-angle associations are observed between the $\chi_1$ and $\chi_3$ angles of an arginine side chain, reflecting the aforementioned concerted $\chi_1-\chi_3$ flips. In most of these cases, $\chi_2$ is in the t rotamer, making the $\chi_1$ and $\chi_3$ axes of rotation ($C^\alpha-C^\beta$ and $C^\gamma-C^\delta$ bonds, respectively) nearly parallel. In this geometry, concerted rotations in opposite directions around $\chi_1$ and $\chi_3$ preserve the orientation of the guanidinium group, and thus its salt bridges. Analogous main-chain–side-chain torsional correlations in both polar and nonpolar amino acids have been observed by LeMaster and co-workers (also termed "crankshaft" motions).[53,54]

In addition to revealing dynamic decoupling in arginines, the simulations also indicate that all polar side chains with more than two dihedral angles are able to reconcile local interactions

(51) Agresti, A. *Categorical Data Analysis*, 2nd ed.; Wiley-Interscience: Hoboken, NJ, 2002.

(52) Best, R. B.; Clarke, J.; Karplus, M. *J. Am. Chem. Soc.* **2004**, *126*, 7734–7735.

(53) LeMaster, D. M. *J. Am. Chem. Soc.* **1999**, *121*, 1726–1742.

(54) LeMaster, D. M.; Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255–9264.

of the terminus with substantial residual conformational entropy. The resulting thermodynamic benefit for glutamates, glutamines, and lysines forming salt bridges or hydrogen bonds is analogous to that for arginines. The behavior observed *in silico* is in agreement with a rotamer library for all four amino acids. Importantly for estimates of conformational entropy from spin relaxation, the simulations indicate that the orientations of the relevant bond vectors of glutamine and lysine side chains can be restricted by salt bridges or hydrogen bonds, and therefore might become insensitive to aliphatic disorder, like the $N^\epsilon - H^\epsilon$ bond vectors of arginines. Finally, although not observed *in silico*, analysis of the rotamer library indicates that methionine side chains also have the potential to restrict their side-chain terminus while retaining rotameric disorder. This might afford some thermodynamic benefit due to the large surface area of the methylsulfanyl group.

Taken together, the findings of this study illuminate the difficulties of inferring side-chain conformational entropy from NMR spin-relaxation probes of local dynamics. For example, an arginine side chain could engage in intramolecular hydrophobic packing in the apo state, with its guanidinium group solvent-exposed, and form an $N^\epsilon$ salt bridge with the substrate. The side chain would show a substantial increase in $N^\epsilon$ order parameter upon binding, although its conformational entropy would largely remain unchanged. On the basis of a difference in order parameter of 0.3 (corresponding to the difference in order parameters between R75 and R27), the estimated difference in side-chain entropy at 300 K would amount to approximately 0.6 kcal/mol, or 1 kT, per residue.[55] Resolution of these difficulties requires further development of methods for direct detection of methylene conformational dynamics, which hitherto have been limited to relatively small proteins (<15 kDa).[56−60]

**Conclusion**

Comparison of simulated and experimental arginine side-chain $N^\epsilon$ order parameters for RNase H has shown that modern MD simulations can provide detailed biophysical interpretations of NMR spin-relaxation data. $N^\epsilon$ dynamics are determined primarily by guanidinium salt bridges and secondarily by packing interactions. Residual disagreement between simulation and experiment may indicate that energy barriers for breaking salt bridges are too large in current MD force fields. Nevertheless, the strikingly good overall agreement between the two methods has revealed that local interactions can dynamically decouple the $N^\epsilon - H^\epsilon$ bond vector from the aliphatic part of the arginine side chain. Guanidinium salt bridges can restrict the $N^\epsilon - H^\epsilon$ bond vector while the aliphatic part remains relatively disordered, and hydrophobic packing can restrict the aliphatic part while the guanidinium group remains relatively flexible. The simulations in concert with the analysis of a rotamer library reveal that all amino acids with more than two side-chain dihedral angles, including present (R, Q, M) and future (E, K) NMR probes of side-chain dynamics, may have similar capabilities. Thus, determination of side-chain conformational entropy from NMR spin relaxation suffers from insensitivity of side-chain termini to rotameric disorder, in addition to other well-known limitations due to overall tumbling and coupling of local motions.[22] The generality of the results for long side chains suggests that dynamic decoupling of the functional terminus from the rest of the side chain may be a general biophysical strategy for preserving residual conformational entropy in proteins and their complexes.

**Supporting Information Available:** Tables containing measured $R_1$ (S1), $R_2$ (S2), and NOE (S3) rates; figures depicting dependence of computed side-chain conformational entropies on the bin size of the integration mesh (S1), as well as dynamic decoupling in R138 (S2), R41 (S3), glutamate (S4), glutamine (S5), lysine (S6), and methionine (S7) side chains. This material is available free of charge via the Internet at http://pubs.acs.org.

JA806475K

(55) Yang, D.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369–382.
(56) Idiyatullin, D.; Daragan, V. A.; Mayo, K. H. *J. Magn. Reson.* **2004**, *171*, 4–9.
(57) Zheng, Y.; Yang, D. *J. Biomol. NMR* **2004**, *28*, 103–116.
(58) Engelke, J.; Rüterjans, H. *J. Biomol. NMR* **1998**, *11*, 165–183.
(59) Jin, C.; Prompers, J. J.; Brüschweiler, R. *J. Biomol. NMR* **2003**, *26*, 241–247.
(60) Yang, D.; Mittermaier, A.; Mok, Y. K.; Kay, L. E. *J. Mol. Biol.* **1998**, *276*, 939–954.