Dynamic ion pair behavior stabilizes single α-helices in proteins

🄳 **Matthew Batchelor**[‡1], 🄳 **Marcin Wolny**[‡1], **Emily G. Baker**[§2], **Emanuele Paci**[‡], **Arnout P. Kalverda**[‡], and 🄳 **Michelle Peckham**[‡3]

From the ‡*School of Molecular and Cellular Biology and the Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom and the* §*School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom*

Edited by Norma M. Allewell

Ion pairs are key stabilizing interactions between oppositely charged amino acid side chains in proteins. They are often depicted as single conformer salt bridges (hydrogen-bonded ion pairs) in crystal structures, but it is unclear how dynamic they are in solution. Ion pairs are thought to be particularly important in stabilizing single α-helix (SAH) domains in solution. These highly stable domains are rich in charged residues (such as Arg, Lys, and Glu) with potential ion pairs across adjacent turns of the helix. They provide a good model system to investigate how ion pairs can contribute to protein stability. Using NMR spectroscopy, small-angle X-ray light scattering (SAXS), and molecular dynamics simulations, we provide here experimental evidence that ion pairs exist in a SAH in murine myosin 7a (residues 858–935), but that they are not fixed or long lasting. *In silico* modeling revealed that the ion pairs within this α-helix exhibit dynamic behavior, rapidly forming and breaking and alternating between different partner residues. The low-energy helical state was compatible with a great variety of ion pair combinations. Flexible ion pair formation utilizing a subset of those available at any one time avoided the entropic penalty of fixing side chain conformations, which likely contributed to helix stability overall. These results indicate the dynamic nature of ion pairs in SAHs. More broadly, thermodynamic stability in other proteins is likely to benefit from the dynamic behavior of multi-option solvent-exposed ion pairs.

Ion pairs are interactions formed between the negatively charged residue side chains in Asp or Glu and positively charged residue side chains in Arg, His, or Lys. These interactions are thought to be important in stabilizing protein structure, and they often occur at catalytically active sites. In structural models, they are commonly depicted as fixed, despite the marginal energetic benefit they provide compared with individually solvated charged groups. Indeed, it is unclear how static or dynamic ion pairs are, as investigations into ion pair dynamics, especially when solvent-exposed, are lacking.

Ion pairs can be classified into different forms, depending on their geometry. A salt bridge is a hydrogen-bonded ion pair and constitutes the most tightly constrained or static picture of an ion pairing. Contact ion pairs have no intervening (water) molecules between the charged groups but do not necessarily fulfill the geometric requirements to form a hydrogen bond. Longer-range ion pairs with weaker electrostatic interactions can form across water molecules.

Ion pairs are thought to be particularly common in single α-helices (SAHs).[4] These are continuous α-helices that remain stable in the absence of any tertiary structure (1, 2). Their sequences are rich in Arg, Glu, and Lys, and the remarkable stability of these structures is thought to arise from a dense network of ion pairs that form between oppositely charged side chains across neighboring turns of the helix (3, 4). SAHs maintain their helicity over a wide range of salt and pH conditions and generally unfold in a noncooperative manner (2, 5–7). They are widespread, occurring in many different proteins (3, 8–12).

The high potential for ion pair formation in SAHs makes them a good model system to investigate how ion pairs contribute to protein stability. Molecular dynamics (MD) simulations of SAHs have suggested that the ion pairs exhibit dynamic behavior (4, 5). The variability in rotamer conformations for Glu-Arg and Glu-Lys ion pairs, from an analysis of helix crystal structures in the PDB, also provides evidence of some dynamic behavior (5, 7). Dynamic ion pairs, using multiple side chain rotamers and alternative pairings, are likely to lower the entropic cost compared with fixed ion pairs (or salt bridges), and this will contribute to their stabilizing properties in SAHs and other proteins. However, experimental evidence of the dynamic behavior of ion pair formation is very limited.

To determine the behavior of exposed ion pairs in solution, we analyzed the SAH from myosin 7a (M7A; mouse, residues 858–935) using a combination of approaches, including solution NMR spectroscopy and molecular dynamics simulations. We selected the M7A SAH, as its less repetitive sequence compared with that of other known SAHs was expected to facilitate

---

[4] The abbreviations used are: SAH, single α helix/helical; MD, molecular dynamics; AUC, analytical ultracentrifugation; MRE, mean residue ellipticity; SAXS, small-angle X-ray scattering; TROSY, transverse relaxation optimized spectroscopy; PDB, Protein Data Bank; HSQC, heteronuclear single quantum correlation; HiSQC, heteronuclear in-phase single quantum correlation; SUMO, small ubiquitin-like modifier; 3D, three-dimensional.

SASBMB

J. Biol. Chem. (2019) 294(9) 3219–3234 **3219**

a more complete assignment of the NMR spectra. A crystal structure recently obtained for mouse M7A (residues 866–935) (13) confirmed our original prediction that M7A contains a SAH between its lever and tail domains, in common with other myosin isoforms (2, 14–16). That study also showed that, whereas 35 Glu-Lys and Glu-Arg helix-compatible ion pairs can potentially form, only eight pairs of interactions were identified in their analysis (13), leaving the majority of Glu, Lys, and Arg residues in the sequence without an ion pair partner. It is likely that only some of the potential ion pairs are captured in a single crystal structure. Thus, the possibility remains that a larger and more diverse set of ion pairs is present in solution. We set out to test this hypothesis by using solution NMR spectroscopy. We also performed MD simulations, to further explore ion pair behavior in this sequence. Our experimental data show that the ion pairs are not static but instead are dynamic, and this is likely to be important for stabilizing this and other proteins.

Results

The full-length M7A SAH is highly helical, elongated, monomeric, and stable

An initial characterization of the full-length M7A SAH (Fig. S1) by CD, analytical ultracentrifugation (AUC), size-exclusion chromatography, and small-angle X-ray scattering (SAXS) demonstrated that our M7A SAH construct (residues 858–935) exhibited the properties of previously well-characterized SAHs in solution (2, 6, 10, 16). M7A SAH was ~90% helical at low temperature, melted noncooperatively, and remained highly helical over a wide range of pH and salt concentrations (Fig. S1, *A–D*). It was monomeric and had an elongated structure in solution, as demonstrated by analytical ultracentrifugation and size-exclusion chromatography, respectively (Fig. S1, *E* and *F*). Additional data from SAXS was again consistent with a continuous (extended) helix structure for M7A SAH (Fig. S1*G*). The Kratky plot of the scattering data (($q^2I(q)$ as a function of q, where the momentum transfer $q = 4\pi\sin\theta/\lambda$, where λ is the beam wavelength, 2θ is the scattering angle, and I is the scattering intensity) is not bell-shaped, as expected for a globular protein, but exhibits a plateau at high q, consistent with an extended structure (17). Using a collection of highly helical model structures with a range of radii of gyration (R_g), the data were best fit ($\chi^2 < 0.5$; Fig. S2) by nearly straight continuous helix model structures (18), and R_g was estimated to be 33.9 Å. This value agreed well with an R_g of 34.5 ± 0.6 Å (mean ± S.D.) estimated from molecular dynamics simulations (see below).

In NMR, the two-dimensional ^1H–^{15}N TROSY (transverse relaxation optimized spectroscopy) spectrum of M7A SAH showed a remarkable lack of dispersion (Fig. 1, *A* and *B*), more typical of a disordered protein than a globular protein of similar molecular mass. However, this is likely to result from the low complexity of the extended SAH, in which the chemical environment is diverse principally along only the helix axis rather than in all three dimensions as in a globular protein. Despite the lack of dispersion, we obtained a virtually complete backbone assignment (^1H_N, $^{15}N_H$, $^{13}C\alpha$, $^{13}C'$, and $^{13}C\beta$ based on standard BEST-TROSY style triple-resonance experiments (19–21)). A

combination of several different NMR experiments was used (Table S1) to obtain unambiguous assignment of shifts for Hα, Hβ, and, where possible, other side chain nuclei (Table S2).

As expected for an α-helical structure, all secondary Cα shifts away from the N and C termini (conservatively >10 residues from termini) were positive and typically fell within the range of 2–3 ppm (Fig. 1*C*) (22, 23). Importantly, there were no breaks in this trend across the length of the peptide, showing that there are no persistent unfolded regions away from the termini. This is a key finding; there is some variation in the local density of potential salt bridge interactions over the sequence (see below), and yet the protein is fully helical throughout. A fuller calculation of secondary structure populations from all of the measured backbone chemical shifts (24) corroborates this feature (Table S3). There is some fraying of the helix at the N- and C-terminal ends, as might be expected, and the overall helicity (at 23.4 °C) is estimated to be 84%. CS-Rosetta– based *de novo* structure generation, guided by chemical shift assignments of ^1H_N, $^1H\alpha$, $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$, and ^{15}N nuclei (25, 26), showed that the lowest-energy structures all exhibit extended helix conformations (Fig. 1*D*).

Next, we used heteronuclear NOE and relaxation experiments to interrogate the local and global dynamics of M7A SAH. ^1H–^{15}N heteronuclear NOE values were high and positive for residues 866–926 (Fig. 2*A*), demonstrating that these backbone amides (away from the termini) are ordered on the picosecond to nanosecond timescale and are thus stably folded. Relaxation NMR experiments showed that ^{15}N backbone nuclei in the more disordered residues at the N and C termini exhibited faster R_1 rates and slower R_2 rates compared with those from the more ordered central residues in the helix (Fig. 2, *B* and *C*). Longitudinal relaxation rates (R_1) of ^{15}N nuclei are sensitive to the fast (picosecond to nanosecond scale) dynamics, whereas transverse relaxation rates (R_2) also include a contribution from slower dynamic processes.

The global/overall rotational correlation time (τ_r) estimated using the R_1 and R_2 rates for the structurally ordered residues (27, 28), was high (13.8 ns at 950 MHz and 15.5 ns at 750 MHz) and would correspond to a mass in the range of 34–49 kDa for a model spherical protein (29). These data indicate that M7A SAH has a much slower rate of tumbling in solution compared with a globular protein of similar mass, consistent with the formation of a long continuous α-helix. The reason for the large apparent τ_r value is the highly anisotropic tumbling of the helix in solution. Backbone N–H bonds in a helix are oriented along the helical axis. Relaxation through reorientation of the N–H bond vector will therefore depend on the less dynamic end-over-end rotation of the helix as opposed to faster axial rotation. The SAH combines the lack of chemical shift dispersion seen in low-complexity intrinsically disordered proteins with the unfavorable relaxation properties of a globular protein with much larger mass, explaining the challenging nature of NMR approaches applied to this system.

NMR shows the presence of ion pairs in M7A SAH

Analysis of the side chains in the SAH, which have low sequence diversity, is much more challenging than in the backbone. Many side chain nuclei resonances, especially those of
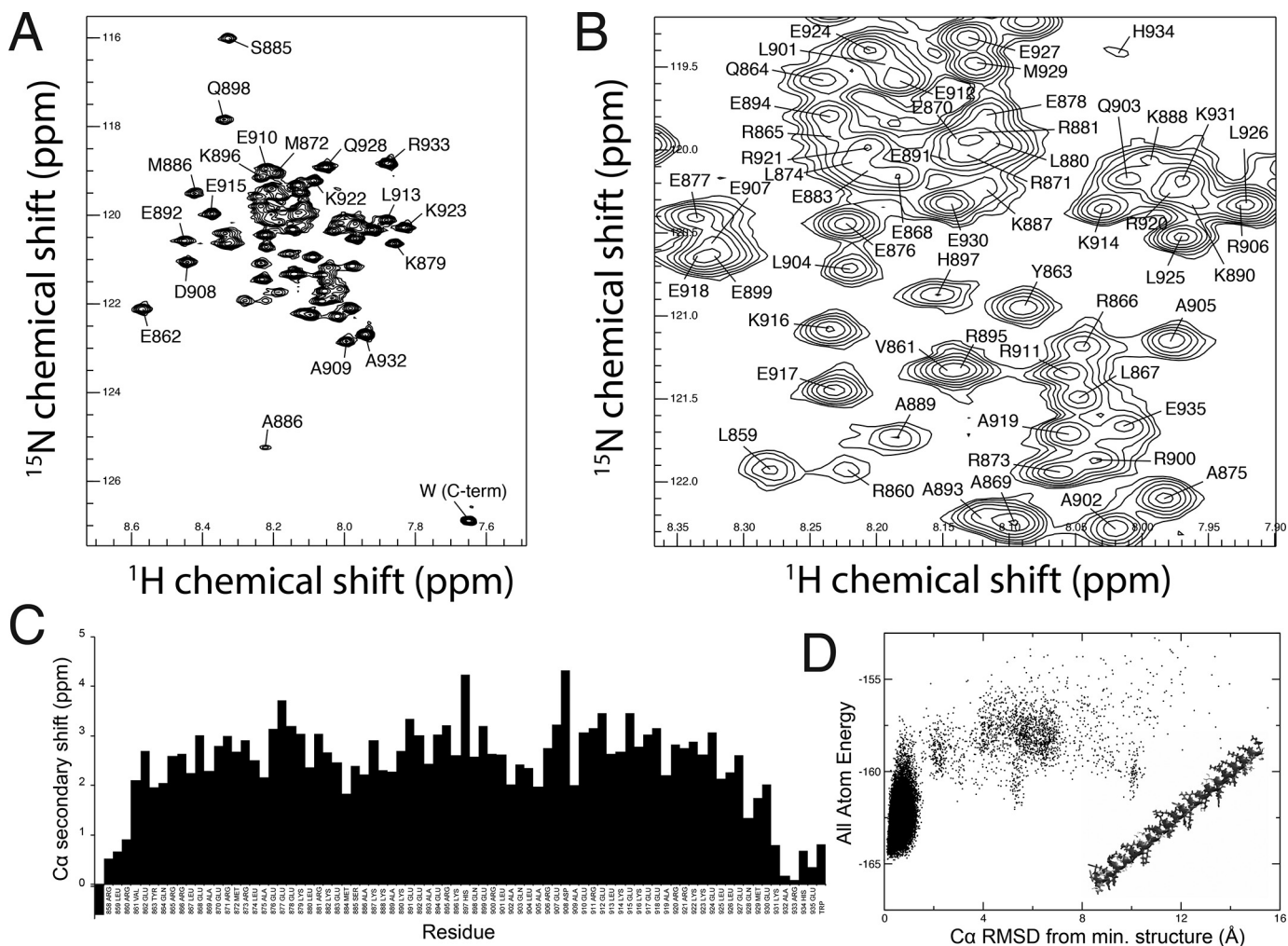
Figure 1. NMR TROSY spectrum for M7A SAH and chemical shift analysis. *A*, ^{1}H–^{15}N TROSY spectrum (750 MHz) for M7A SAH. Note the limited dispersion; the vast majority of peaks occur within a 4 ppm (^{15}N) and 0.4 ppm (^{1}H) region, and there is a high degree of peak overlap. *B*, a *zoomed-in* view of the spectrum shows the assignments of the central overlapping peaks. *C*, the secondary Cα shifts provide a good indication of the continuous helical nature of the peptide. *D*, CS-Rosetta results based on chemical shift (HN, N, CO, Cα, Cβ, Hα, and Hβ) assignments. Each point in the plot represents, for each of 30,000 chemical shift–compatible Rosetta-generated structures, the energy and root mean square deviation for Cα atom positions from those in the lowest energy structure. The lowest energy structure, which after exclusion of flexible tails ranges from Val[861] to Glu[930], is depicted on the plot.

Glu, Arg, and Lys side chains, are characterized by a significant degree of overlap. The level of degeneracy of resonance positions is illustrated for parts of the ^{13}C HSQC spectrum (Fig. 3A), which displays tight clusters of peaks for different side chain positions. The most extreme example of this is for the Lys Hϵ–Cϵ correlations, which appear effectively as a single peak for 11 residues (22 protons). This effectively puts an NOE-based NMR structure determination, using ARIA (30–32) or similar methods, beyond reach.

However, our main focus was to evaluate the behavior of the Glu, Arg, and Lys side chains and possible ion pairs between them. From studies on intrinsically disordered proteins, it has been established that N and CO nuclei retain the largest degree of chemical shift dispersion, and therefore experiments that relay magnetization to side chain (or backbone) N and CO hold the most promise in this regard (33, 34). Glu residue interactions can be inferred from chemical shift measurements of the side chain carbonyl carbon and nearby Hγ nuclei. Improved peak dispersion (over ^{13}C-HSQC spectra) was achieved for Glu

Hγ resonances by correlating with the adjacent carbonyl carbon (Cδ) nuclei, which are also the most relevant marker nuclei for ion pair interactions involving Glu (Fig. 3B). Specific assignment for all Glu Cδ nuclei and thereby a significant number of Glu Hγ pairs was achieved by linking Cδ shifts back to the backbone (see "Experimental procedures").

Two key findings are demonstrated by these experiments. First, the measured Glu Cδ shifts (>183.1 ppm) indicate that all Glu residues in M7A SAH are in the deprotonated COO^{-} form rather than the protonated COOH form. For comparison, in a random coil model peptide, deprotonated Glu Cδ shifts were 183.8 ppm, whereas protonated Glu Cδ had significantly lower chemical shifts (179.7 ppm) (35). Second, most of the accurately assigned Hγ pairs from the central part of the protein displayed two distinct peaks that result from restricted rotational motion between rotameric states. In contrast, Hγ pairs in residues near the protein termini (Glu[862] and Glu[935]) exhibited very small resolvable differences in chemical shift. This result indicates that side chains of the central Glu residues in M7A SAH are not
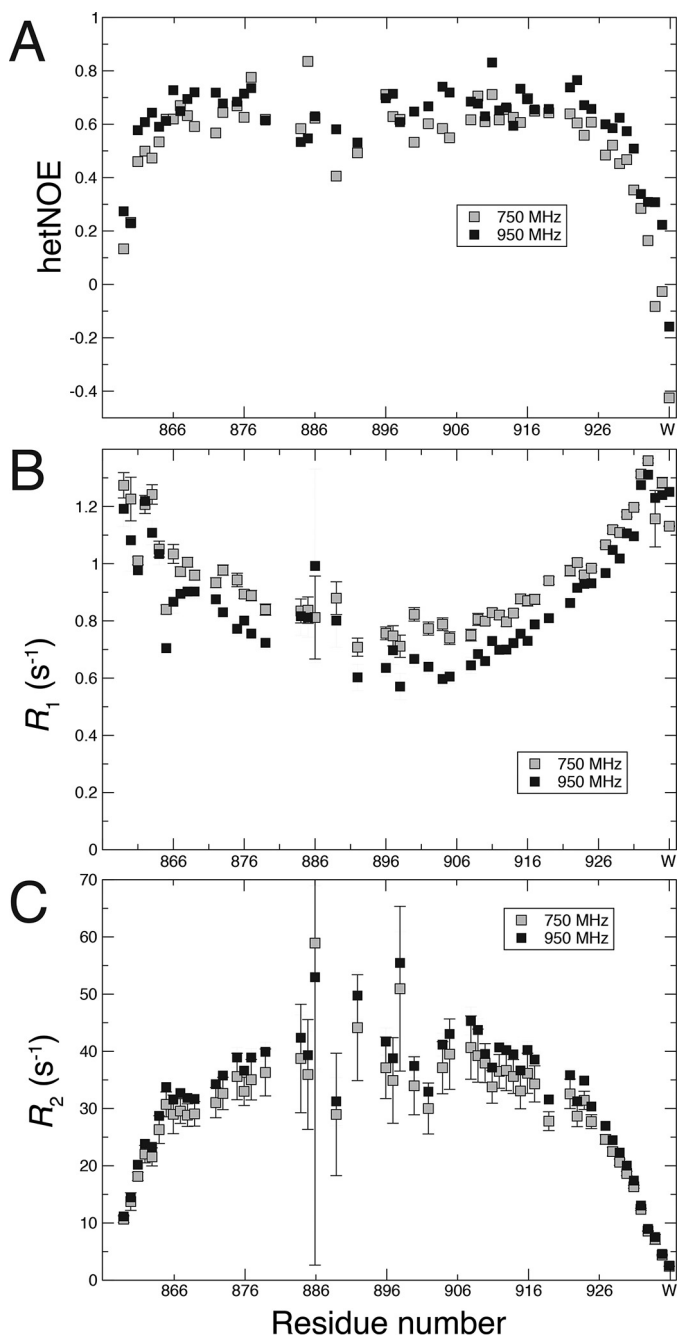
Figure 2. Dynamic properties from [1]H–[15]N NOE and ^{15}N relaxation studies of M7A SAH. Shown are plots of [1]H–[15]N NOE (*A*), longitudinal (R_1) (*B*), and transverse (R_2) (*C*) ^{15}N relaxation rates, as a function of residue number. Relaxation data were not recorded for a number of residues, most particularly between residues 876 and 896; many of the associated peaks were heavily overlapped with higher intensity neighbors. *Error bars* in *B* and *C*, error estimation from Monte Carlo–based fitting.

able to freely rotate in the structure and are thus interacting with other residues in the protein.

The most appropriate nuclei to report on salt bridge formation from the positively charged residues (Lys and Arg) are the side chain nitrogens (36–42). Under our original conditions of pH 7.4 and 23.4 °C, the ionizable side chain H–N correlations in Lys and Arg were not accessible; HSQC and HiSQC spectra (41) lack peaks for these groups due to rapid hydrogen exchange with water. However, reducing the pH to 5.5 and temperature

to 10 °C slowed proton exchange and enabled us to observe Lys Hζ–Nζ, Arg Hη–Nη, and Arg Hε–Nε correlations for M7A SAH. The overall helicity changes little from the original conditions, supporting the assumption that the general structure and mode of stabilization are not affected (Fig. S1, *B* and *C*).

At pH 5.5 and a temperature of 10 °C, the Lys Hζ–Nζ correlations in M7A SAH (from 11 Lys residues) are visible, albeit as a single peak that is only resolved into two peaks at high resolution (Fig. 4*A*). Poorly dispersed Hη–Nη correlations for the 14 Arg residues were observed with Hη shifts only ranging from 6.7 to 7.1 ppm (Fig. 4*B*). In contrast, it was possible to independently interrogate all Arg Hε–Nε correlations, which appear as sharper peaks across a broader range of ^{1}H chemical shifts (Fig. 4*C*). As the Arg Hε–Nε correlations are very well-resolved, we utilized NMR techniques to investigate their participation in ion pairs. We used existing protocols or designed new experiments to measure the chemical shifts, dynamics (43, 44), intra-residue $^3J_{NεCβ}$ couplings, and long-range interresidue $^3J_{NεC'}$ couplings between Arg Nε and carbonyl carbons (C′) for the Glu side chains (Cδ). The Hε–Nε correlations were assigned to each Arg residue by linking back to the backbone (see "Experimental procedures" and Fig. S3).

The Arg Hε–Nε correlations fall into distinct groups that indicate their potential involvement in ion pairs (Fig. 4*C*). Peaks from Arg residues that can potentially make multiple interactions are generally found on the far left (Hε >7.6 ppm). These match the six Arg residues that form ion pairs in chain A of the crystal structure (13). Peaks from Arg residues close to the N and C termini are found on the far right (Hε <7.3 ppm). The peak for the central Arg residue, Arg[900], which has no potential ion pair partners, appears close to the terminal Arg residues. Participation of an NH group in a hydrogen bond (or salt bridge) should be accompanied by a downfield shift of the proton (45). Although the Hε shifts observed here are good indicators of ion pair participation, more significant chemical shift changes (Hε ~9 ppm) have been seen previously in examples where Arg side chains participate in a hydrogen bond (43, 46). It is worth noting that the Hη–Nη correlations all appear close together (Fig. 4*B*). There are no downshifted Hη or Nη resonances that would indicate salt bridge formation through these groups, unlike the Hη–Nη correlations seen in spectra of complexes formed through Arg–carboxylate (47) or Arg–phosphotyrosine (43) interactions.

Further evidence points to the majority of central Arg residues being involved to some degree in ion pair formation. The Hε–Nε peaks for central Arg residues show small positive heteronuclear NOEs (Fig. 4*D* and Table 1) consistent with the side chains being relatively ordered compared with those of the N and C termini. Those located close to the N or C terminus have clear negative [1]H–[15]N heteronuclear NOEs. Whereas the comparatively high values of the NOEs indicate a restricted or slower timescale of local motion, the values do not approach those measured for the backbone residues, nor those of Arg Nε involved in binding to phosphotyrosine (43). R_1 and R_2 values for Arg Nε show a pattern broadly similar to that of the heteronuclear NOEs with higher values for the central residues compared with those at the termini (Table 1). This again is an indicator that central Arg side chains are more constrained than
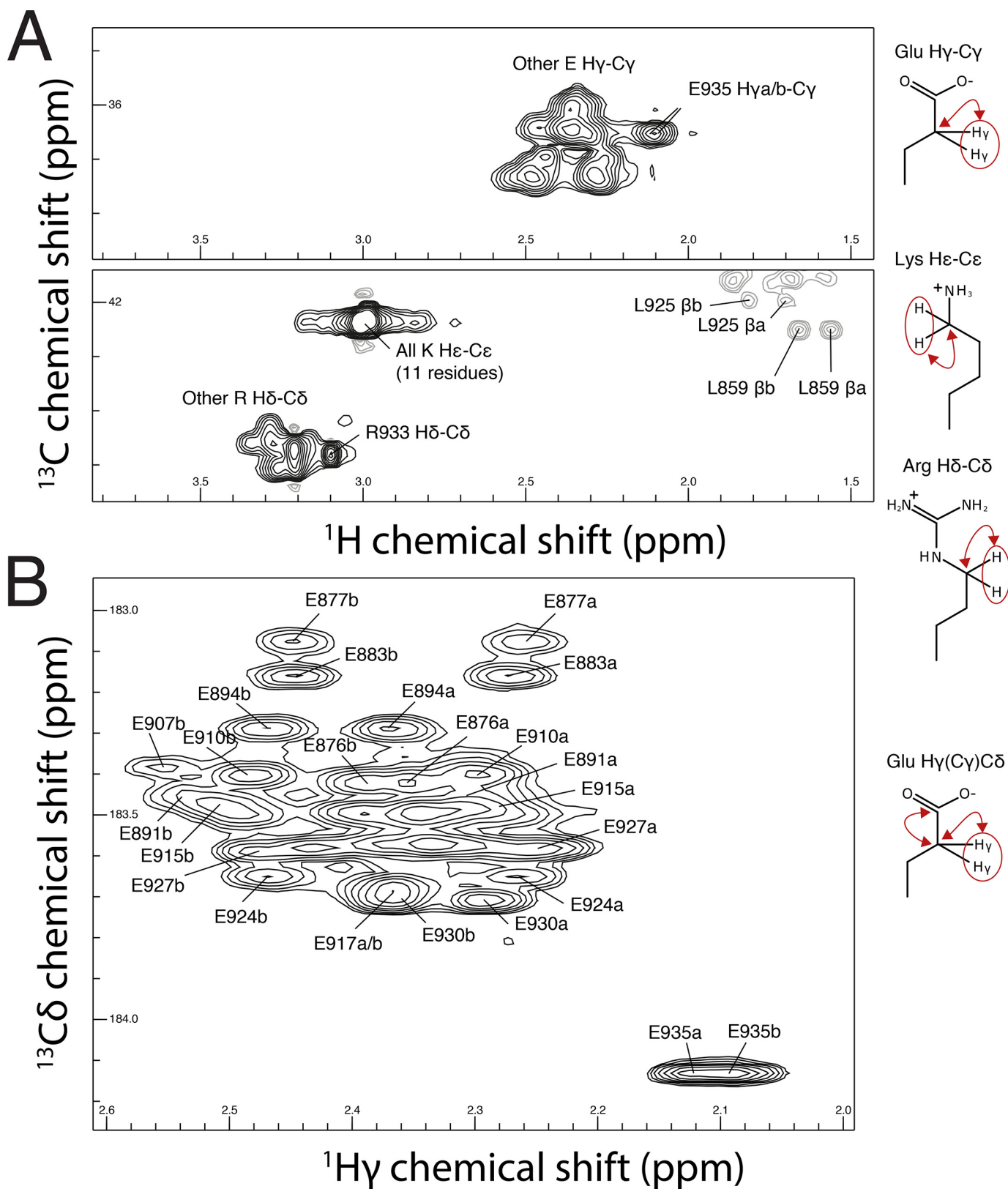
ASBMB

Figure 3. NMR spectra to interrogate side chain nuclei in Glu, Lys, and Arg. *A*, sections of the ^{13}C HSQC spectrum for M7A highlighting the most distal H-C correlation from the side chain of Glu (Hγ–Cγ), Lys (Hε–Cε), and Arg (Hδ–Cδ). *B*, the Hγ(Cγ)Cδ Glu-specific spectrum exhibits improved peak dispersion over the Hγ–Cγ correlations in the ^{13}C HSQC.
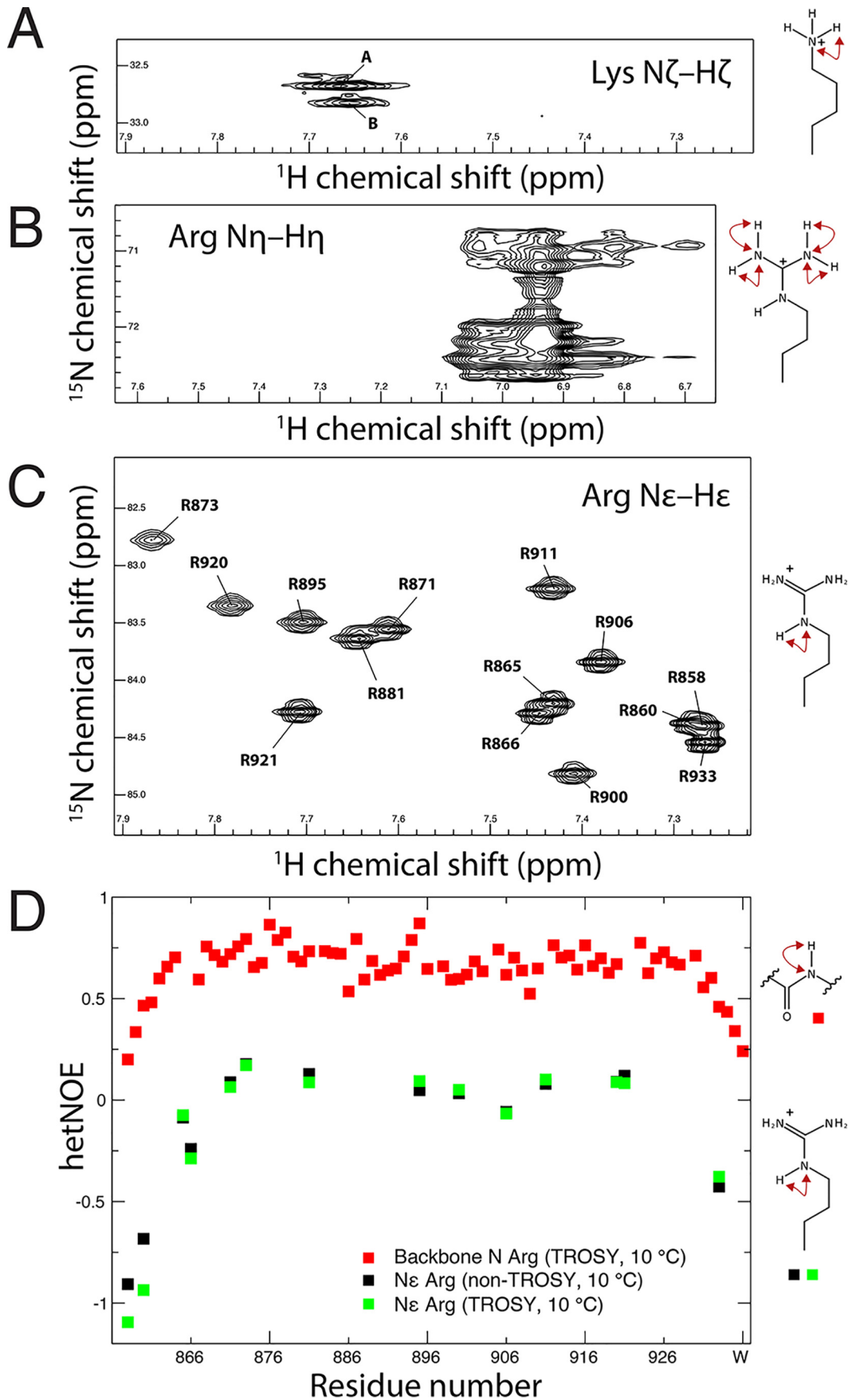
Table 1
Measured $^3J_{NC}$ values, heteronuclear NOE values, and relaxation rates for Arg Nε

Arginine	$^3J_{N\epsilon C\beta}{}^a$	hetNOE[b]	$R_1{}^c$	$R_2{}^d$	No. of Glu neighbors[e]
	Hz		s^{-1}	s^{-1}	
858[f]	1.1	−1.00 ± 0.09	0.53	2.3	1
860[f]	1.2	−0.81 ± 0.13	0.60	2.8	0
865	1.3	−0.08 ± 0.01	0.68	4.3	2
866	1.1	−0.26 ± 0.02	0.69	5.0	2
871	0.9 ± 0.1	0.08 ± 0.01	0.74	6.7	1
873	NR	0.17 ± 0.003	0.63	16.2 ± 0.7	3
881	0.7 ± 0.1	0.11 ± 0.02	0.74	6.0	2
895	0.6 ± 0.2	0.07 ± 0.02	0.74	9.2 ± 0.1	3
900	1.7 ± 0.1	0.04 ± 0.009	0.76	6.9	0
906	0.8 ± 0.1	−0.06 ± 0.004	0.72	5.0	1
911	0.6 ± 0.1	0.09 ± 0.01	0.77	7.8	2
920	NR	0.09 ± 0.001	0.72	12.5 ± 0.2	2
921	1.2 ± 0.1	0.10 ± 0.018	0.76	5.5	3
933	1.1	−0.40 ± 0.02	0.64	2.9	1

^{a} Errors for $^3J_{NC}$ were calculated by finding the range in $^3J_{NC}$ values through adding and subtracting the baseline noise level to the peak intensities (<0.1 Hz where not shown). NR, peaks for Arg[873] and Arg[920] were barely visible in the sub-spectra, on the level of baseline noise; the error associated with the ratio in peak intensities was too high to give a meaningful $^3J_{NC}$ value.
[b] hetNOE errors are S.D. values from the two experiments (TROSY and non-TROSY).
^{c} R_1 errors are all <0.01 s⁻¹.
^{d} Other than those shown, R_2 errors are all <0.1 s⁻¹.
[e] The last column shows the number of Glu neighbors compatible with forming an ion pair within the helix context for each Arg residue.
^{f} Overlapping peaks.

those at the termini (43). R_2 values are particularly high for Arg[873], Arg[920], and Arg[895] (those with the most downshifted Hε), indicating that slower (greater than nanosecond) dynamic relaxation processes are also important for these residues.

NMR shows that ion pairs in M7A SAH are dynamic

The above data shows evidence for ion pair formation in M7A SAH. The next question was whether we could detect salt bridges (or hydrogen-bonded ion pairs). Fixed and stable salt bridges should show through-hydrogen-bond couplings between the associated nitrogen and carbonyl carbon nuclei (48, 49). Very weak coupling (~0.2 Hz) for mobile hydrogen bonding interactions between Lys Nζ and C′ nuclei in ubiquitin has also been observed using this approach (42). In addition, there is an example of a weak coupling reported for a side chain–side chain interaction (between Arg and Asp), although no value was reported (46). To this end, we adapted a spin-echo difference experiment (42) to measure $^3J_{NC'}$ couplings between Arg Nε and Glu Cδ (see "Experimental procedures" and Fig. S4).

First, measurements of small $^3J_{NC}$ couplings *within* each Arg residue (*i.e.* $^3J_{N\epsilon C\beta}$) were used to validate the spin-echo difference experiment and report on the Arg side chain conformational ensemble (42). The experiment could readily measure $^3J_{N\epsilon C\beta}$ coupling along the Arg side chain, as shown by a drop in peak intensity between the two component spectra (Fig. 5A). The values ranged from 0.6 Hz for the downfield-shifted Arg[895], to 1.7 Hz for Arg[900] (Table 1). The $^3J_{N\epsilon C\beta}$ coupling constant reports on the distribution of χ_3 (Nε–Cδ–Cγ–Cβ) angles occupied by each Arg residue. By analogy with other

N-C–C-C dihedrals (42), larger $^3J_{N\epsilon C\beta}$ values are indicative of larger *trans* χ_3 contributions, and smaller $^3J_{N\epsilon C\beta}$ values indicate larger *gauche* contributions (see Fig. S5A for a qualitative description). Arg[900], which has no potential ion pair Glu neighbors, spends more time in the less sterically demanding *trans* conformation. Arg[895], on the other hand, favors *gauche* conformations, presumably to more readily interact with one or more of its three potential Glu ion pair neighbors. Intermediate $^3J_{N\epsilon C\beta}$ values point toward variation in the χ_3 dihedral angle, with side chains occupying both *trans* and *gauche* rotamers.

The same spin-echo difference experiment was then used to examine through-hydrogen-bond coupling ($^3J_{N\epsilon C'}$) to Glu Cδ. As there was no drop in peak intensity between the component spectra of the experiment, values of $^3J_{N\epsilon C'}$ remained below the detection limit (of ~0.2 Hz) for at least 12 of the 14 Arg residues (Fig. 5B). As peaks were not observed in the subspectra for Arg[873] and Arg[920], due to their fast relaxation properties, estimation of $^3J_{N\epsilon C'}$ was not possible for these residues. Very small coupling constants indicate limited orbital overlap from Arg Nε to Glu Cδ.

Taken together, the chemical shifts and dynamics of the Arg Nε reflect the participation of the central Arg side chains in ion pair interactions with neighboring groups, whereas the very small size or absence of through-hydrogen-bond couplings argues against specific salt bridges (50). These data indicate either that the SAH is composed of ion pairs in a dynamic network or that there is a preference for weaker solvent-bridged ion pairs. Salt bridges and contact ion pairs, if they form, can only exhibit short lifetimes, perhaps as part of a dynamic ensemble of bonded, unbonded, and solvent-bridged states.

Molecular dynamics simulation for M7A SAH indicates dynamic ion pair interactions between charged side chain residues

The NMR data indicate that whereas the M7A SAH is helical, formation of salt bridges is transient, or looser ion pairs are formed (*e.g.* bridged by an intervening water molecule). To explore this further, we ran equilibrium simulations of M7A SAH using a perfect α-helix as a starting structure. During the simulation, the structure flexes but remains a continuous single helix throughout, with an average overall helicity of 90–92%. As indicated above, the average R_g of the protein from simulation (34.5 Å) matched the experimental value from SAXS (33.9 Å).

Simulations show the presence of ion pairs not observed in the crystal structure and highlight transitions between different accessible ion pairings. Plots of the distance between charge centers in Glu-Arg and Glu-Lys pairs show many large- and small-scale transitions (Fig. 6A). The ion pair occupancy (fraction of time that charged residues were close enough to form a "contact" ion pair, calculated using a threshold distance of 4 Å (5)) showed which ion pairs are most likely. All of the ion pairs compatible with an extended helical conformation were ob-

Figure 4. ^{15}N HiSQC NMR spectra for side chain Lys or Arg N–H groups, and heteronuclear NOE data for Arg Nε–Hε at pH 5.5, 10 °C. Sections of ^{15}N HiSQC spectra for M7A SAH highlighting the lysine NH₃ region ~33 ppm (*A*), Arg Nη–Hη region ~72 ppm (*B*), and Arg Hε–Nε region ~80 ppm (*C*). *D*, heteronuclear NOEs for backbone ^{15}N nuclei (*red*) and for Nε nuclei in all 14 Arg residues (*black/green*). A TROSY version of the heteronuclear NOE experiment was required to resolve the backbone amide correlations at this temperature.
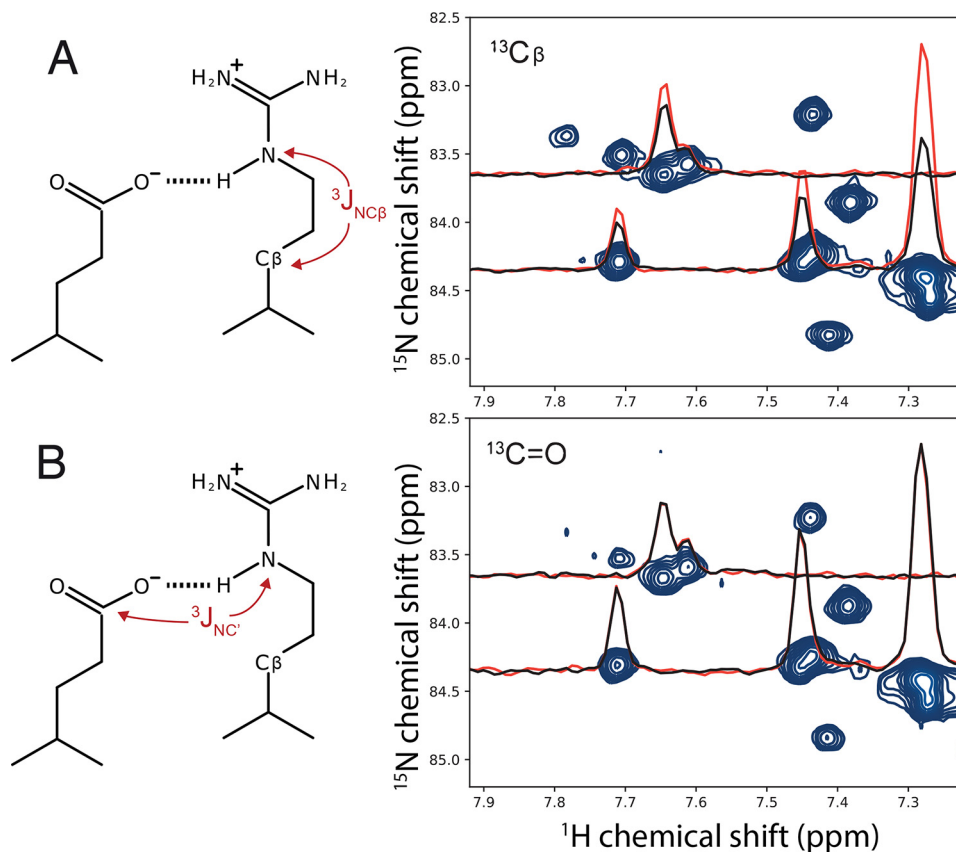
Figure 5. Spin-echo difference ^1H–^{15}N correlation experiments to measure $^3J_{NC}$. *A*, reference sub-spectrum for the Arg NϵCβ experiment. *Overlaid* are two example line scans from the reference (*red*) and the attenuated sub-spectrum (*black*). A reduction in peak intensity between the two sub-spectra indicates a measurable $^3J_{NC}$ coupling for each peak. Peaks for Arg[873] (and Arg[920]) were barely visible in the sub-spectra, on the level of baseline noise; the error associated with the ratio in peak intensities was too high to give a meaningful $^3J_{NC}$ value. *B*, reference sub-spectrum for the through-hydrogen bond $^3J_{NC}$ experiment to measure Arg Nϵ coupling to Glu Cδ. The example line scans show that the peak intensities do not change between the two sub-spectra, indicating no measurable $^3J_{NC}$ for each peak. Peaks for Arg[873] and Arg[920] were not observed in the sub-spectra due to their fast rate of relaxation and the long evolution time required. A maximum coupling constant, based on the level of the baseline noise and the intensity of visible peaks in the sub-spectra, was conservatively 0.2 Hz. Plots for *A* and *B* were prepared using nmrglue (80).

served during the simulation, albeit with a wide range of occupancies (1–91%; Table S4). For example, the Glu[892]-Arg[895] ion pair is highly occupied, being observed for 61% of the time during the simulation. This pair was also found to be present in the crystal structure (13). Alternative ion pairs involving Glu[892] (Glu[892]-Lys[888] and Glu[892]-Lys[896]) were also observed, albeit for a lower fraction of the overall simulation time (9 and 4%, respectively). The pairing Arg[895]-Glu[899] (35% occupancy) exemplifies one of many cases where ion pairs that were absent from the crystal structure were present in the simulation. On average, potential Glu-Lys ion pairs were occupied for 18% of the simulation, whereas Glu-Arg ion pairs were occupied for 37%. The total number of ion pairs observed at any one time was 11.8 ± 2.2 (mean ± S.D.); there are 41 potential helix-compatible ion pairs.

We also found that charged residues formed ion pairs simultaneously with two neighboring partners without breaking the helix structure, in agreement with our earlier findings (5). For example, the ion pairs Arg[873]-Glu[870] and Arg[873]-Glu[877] were formed simultaneously for 45% of the simulation, whereas Arg[873]-Glu[876] and Arg[873]-Glu[877] were simultaneously occupied for 8% (see *snapshot images* in Fig. 6 (*B–E*) and Movie S1).

Contact ion pairs form and break on a rapid timescale. Estimates of the average ion pair lifetimes were position-dependent, but overall, the lifetimes were longer for Glu-Lys (38 ps)

than for Glu-Arg (21 ps), in line with our previous simulations of SAHs (5). Additional simulations, which specifically used the starting conformations of one or other of the two chains in the unit cell from the crystal structure, showed similar dynamic behavior in terms of ion pair lifetimes and pairings that were absent from the crystal structure. Overall, the simulations of M7A SAH show a more complicated and dynamic picture of ion pair interactions than one of a small number of specific ion pairs implied by the crystal structure (13).

Analysis of the M7A(866–935) crystal structure reveals evidence for variable and alternative ion pairs

Our experimental findings prompted us to reanalyze the ion pairs made within the deposited crystal structure of M7A(866–935) (PDB code 5WST). There are two SAH chains in the unit cell, and the analysis is performed for both. The results of this new analysis are consistent with the potential dynamic behavior of ion pairs within it. A helical net plot demonstrates the large number of potential ion pairs that could form in M7A SAH (Fig. 7A).

Using a threshold distance of 4 Å between one of the side chain nitrogen atoms (Lys or Arg) and a side chain oxygen atom (Glu) to define a contact ion pair (5), 11 ion pairs were defined in chain A (Fig. 7B) and 8 in chain B (Fig. 7C), some of which were overlooked in the original paper, such as that between
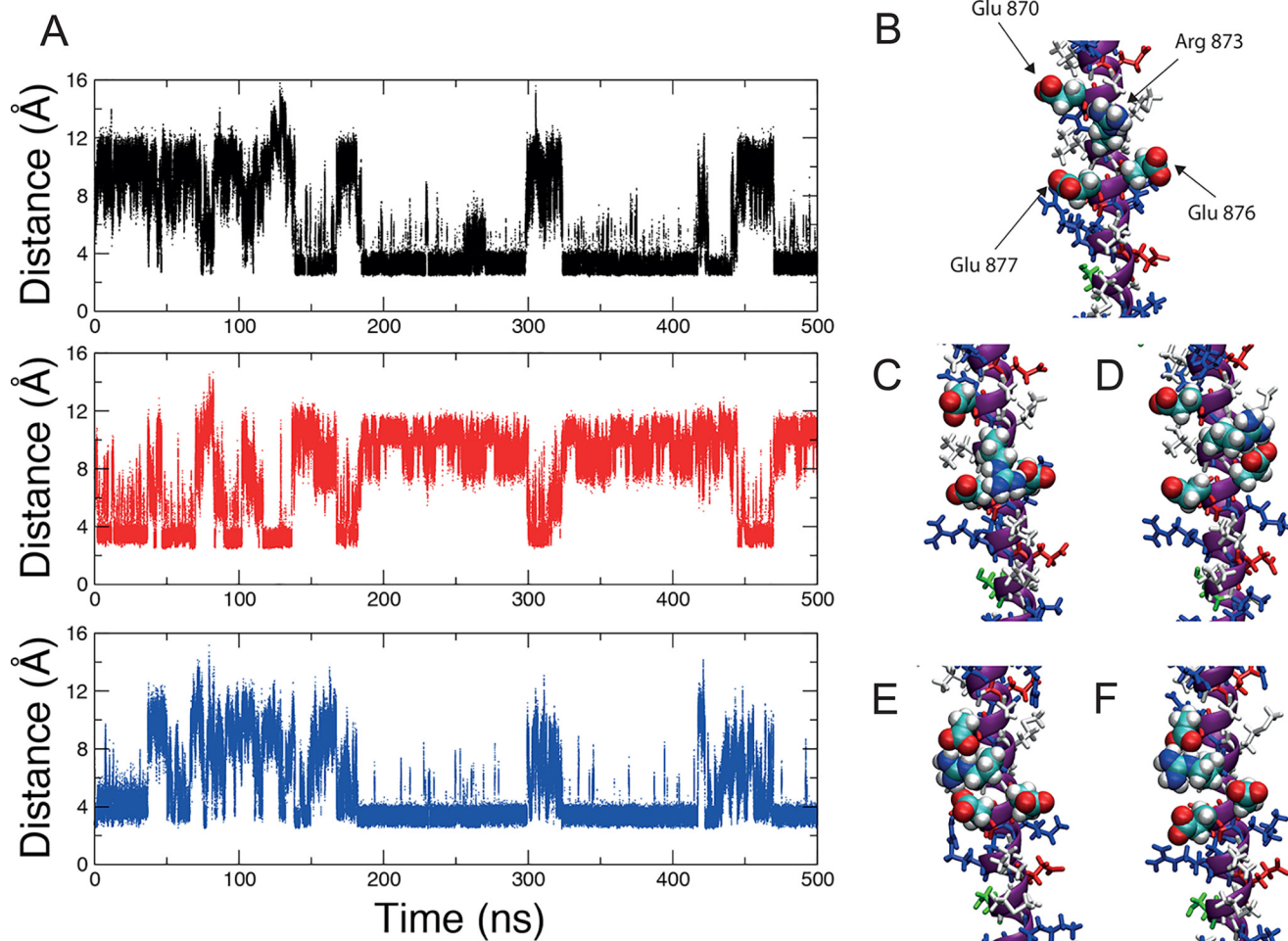
Figure 6. Simulation results highlighting the variable nature of ion pairs for an example residue Arg[873]. *A*, distances between charged residue pairs throughout the simulation. Arg[873]-Glu[870] (*black*), Arg[873]-Glu[876] (*red*), and Arg[873]-Glu[877] (*blue*). The distance criterion chosen to mark the formation of an ion pair was 4 Å (5). *B–F*, snapshots from the simulation indicating many interaction modes. A movie of the full simulation trajectory centered on Arg[873] is provided (Movie S1). The helix is oriented with the N terminus to the *top* of the image. Arg[873], Glu[870], Glu[876], and Glu[877] are shown in *space-fill*, the nitrogen atoms of Arg[873] are *blue*, and oxygen atoms of the Glu residues are *red*. *B*, no ion pair formation. *C*, simultaneous Arg[873]-Glu[876] and Arg[873]-Glu[877] ion pair formation. *D*, Arg[873]-Glu[876] only. *E*, simultaneous Arg[873]–Glu[870] and Arg[873]-Glu[877] ion pair formation. *F*, Arg[873]-Glu[870] only.

Lys[882] and Glu[878]. On the other hand, the ion pair reported to form between Glu[907] and Arg[911] is unlikely to do so as this exceeds the threshold distance (>4.7 Å).

Specific ion pairs were observed in one of the two helices in the unit cell, but not in the other, suggesting that the ion pairs may be dynamic. For example, the shortest N–O distance in the Arg[895]-Glu[892] pair was 2.8 Å in one chain but 7.3 Å in the other. Explicit dynamic behavior is additionally shown by the placement of alternate side chain conformers into the electron density (51) for Glu[878] in chain B, displaying two side chain conformers that form an ion pair to either Arg[881] or Lys[882]. In chain A, Glu[878] occupies a single position where it forms an ion pair with both Arg[881] and Lys[882] simultaneously. In addition, side chains forming intermolecular (crystal contact) ion pairs in the crystal may instead form intramolecular bridges in the monomer in solution. This analysis demonstrated a higher variability in ion pairs than originally described (Fig. 7, *B* and *C*), and yet the number of observed ion pairs was still small given the potential number of ion pairs that could exist.

We estimated $^3J_{N\epsilon C\beta}$ for Arg residues using dihedral angles measured in the crystal structure (Fig. S5B, *red lines*) and com-

pared them with experimental $^3J_{N\epsilon C\beta}$ values (Fig. S5C). The correlation between crystal structure and NMR values was poor, indicating substantial deviation from the crystal structure dihedral angles in solution. Distributions of χ_3 dihedral angles were also calculated for each Arg residue from the MD simulation (Fig. S5B, *blue*), and ensemble-averaged $^3J_{N\epsilon C\beta}$ values were estimated (Fig. S5D). The correlation of simulation with experimental $^3J_{N\epsilon C\beta}$ values, albeit not perfect, was improved over those of the crystal structure analysis. Some dihedral undersampling during the simulation may contribute to any differences observed. Variation of the χ_3 dihedral, and with it the ability of Arg residues to form ion pairs, is a feature described by both NMR and MD results. These analyses support our experimental findings that ion pairs exhibit dynamic behavior in solution for M7A SAH.

Discussion

Here we have addressed the question of whether the ion pairs formed between specific residues in the M7A SAH are strong and persistent, whether the ion pairs change, and if ion pairs are even important for stability at all, which has been questioned
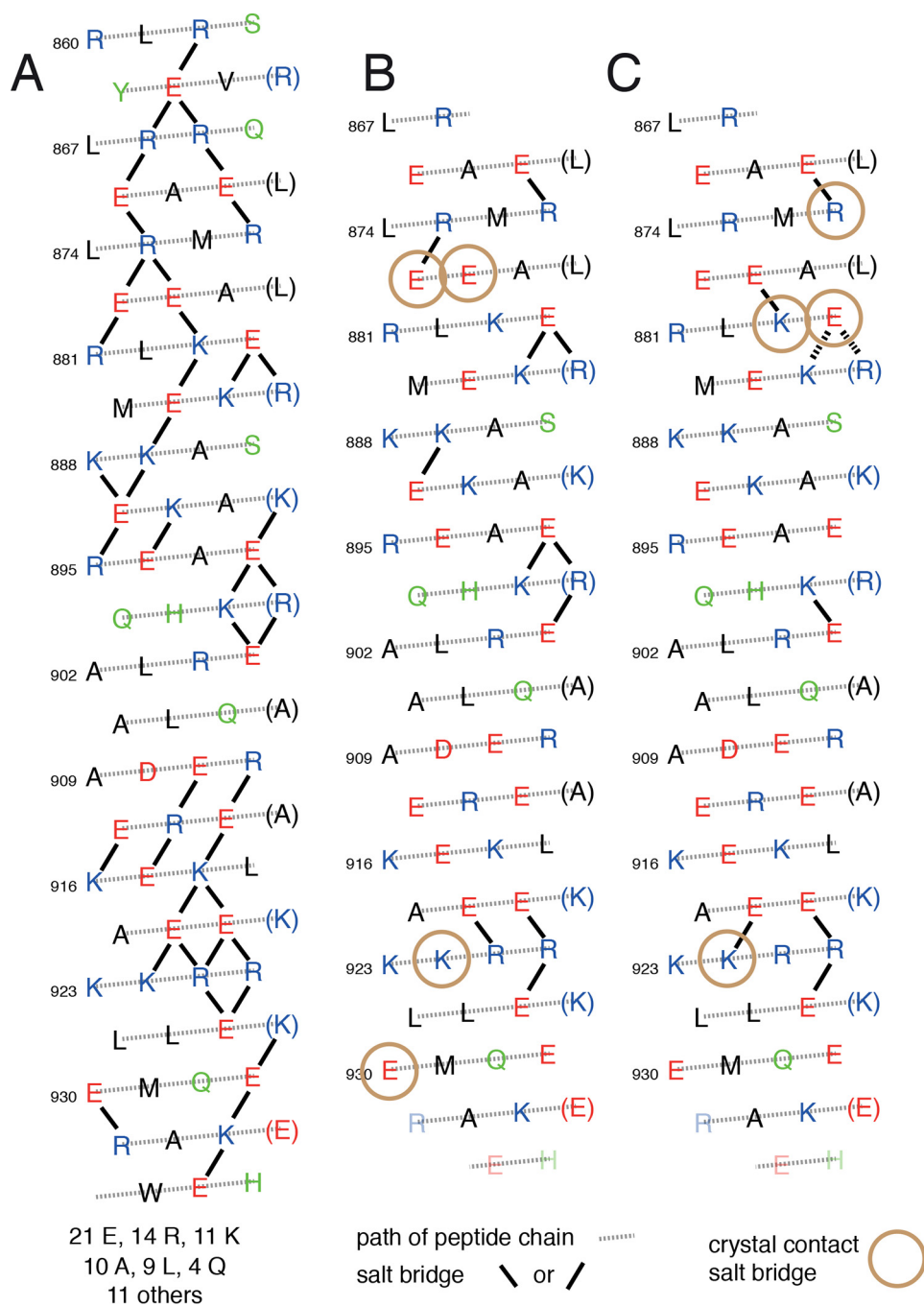
Figure 7. Helical net plot representation highlighting potential pairs in M7A(858–935) (M7A SAH) and observed ion pairs between side chains in the crystal structure of M7A(866–935). *A*, sequence of M7A SAH including the additional N-terminal Ser and C-terminal Trp shown as a *helical net plot*. All potential K/R$_i$–E$_{i+4}$, E$_i$–K/R$_{i+4}$, and K/R$_i$–E$_{i+3}$ ion pairs and the number of Glu, Arg, Lys, and other residues are shown. *B*, in chain A from PDB entry 5WST, 11 ion pairs are observed. *C*, in chain B from PDB entry 5WST, eight ion pairs are observed. The side chain of Glu[878] displays two conformers (a and b) in the crystal structure that form an ion pair to either Arg[881] or Lys[882] (*dashed lines*). The three C-terminal residues (shown *faded*) were not observed in the crystal structure. Residues involved in an ion pair between different helices within the crystal are marked with *circles* (crystal contact ion pair). These were Glu[876](A)–Lys[879](B) (N–O 3.7 Å), Glu[877](A)–Lys[922](B) (N–O 2.7 Å), Lys[922](A)–Glu[878b](B) (N–O 2.2 Å), and Glu[930](A)–Arg[871](B) (N–O 3.3 Å). If only those residues visible in the crystal structure are considered, the average total number of ion pairs from MD simulation was 10.3 ± 2.1 (mean ± S.D.).

(52). Our combination of experimental data and simulations supports the idea that the stability of SAHs is not driven by strong, hydrogen-bonded, and persistent salt bridges, but via a network of fluctuating ion pairs that continually form and break. This, together with the high thermodynamic stability and noncooperative unfolding of these domains, supports the idea that the extended helical state has high conformational entropy. This agrees with and extends our recent molecular dynamics simulations on artificial SAH sequences that showed that the interplay between oppositely charged residues is highly dynamic, with salt bridges being formed and broken on the 10-ps timescale in MD (5).

The NMR data show that the M7A SAH forms a long α-helical structure in solution. However, specific (fixed or long lasting) ion pair interactions could not be identified. Together with the supporting data from MD simulations, we show that this is

because the charged residues have a choice of interactions and can exchange on a very rapid timescale. The NMR data indicate, by more indirect measures, such as the chemical shifts of the Arg Hε protons and the dynamics of the Nε and $^3J_{N\epsilon C\beta}$ values, that most of the central Arg residues are forming ion pairs. The one exception, Arg[900], is not expected to form an ion pair. Inspection of the helical net plot (Fig. 7*A*) shows that it has no potential ion pair partners compatible with the helix structure (Fig. 7*A*). Correspondingly, it appears in the same part of the HiSQC spectrum as the terminal Arg residues (Fig. 4*C*) and has a high $^3J_{N\epsilon C\beta}$ value. It should be noted that some alternative experimental strategies exist to probe the presence of ionic interactions that would be complementary to our work (53, 54).

Whereas the data do not support the existence of specific long lifetime salt bridges, they are consistent with the dynamic behavior of ion pairs observed in MD simulations. $^3J_{NC'}$ values of 0.23 and 0.17 Hz were measured for mobile hydrogen bonds in ubiquitin (42). Their simulated average lifetimes were 45 ps and 30 ps, and hydrogen bond occupancies were 80 and 51%, respectively, using a salt bridge distance criterion of 3.5 Å. If we apply this more stringent criterion to our simulations, the average Arg-Glu (and Lys-Glu) lifetimes are much shorter ($<$10 ps), and Hε-mitigated ion pairs average 25% occupancy. Thus, smaller coupling constants, below the level dictated by the noise, are in keeping with the MD results. The general consensus would appear to be that, if anything, MD simulations still overstabilize salt bridges compared with experiment (55–57). We think that the continuous and rapid exchange of ion pairs is important for providing SAHs with their stability. The choice of multiple interactions together with their density is likely to give this system its remarkable properties.

Interestingly, despite the local variation in the density of potential ion pairs in M7A SAH (Fig. 7*A*), its behavior is consistent with that of a SAH along its entire length. This suggests that the density of interactions may not relate directly to the propensity of being helical. Both crystallography and MD simulations indicate that the presence of around 10 ion pairs or 30% of the total possible number at any one time is enough to keep the system fully helical. The subtle balance of interactions allows for a reversible response to external pressures such as the mechanical stresses (6) encountered in the function of these proteins.

In summary, it has been unclear exactly how potential ion pair formation can contribute to the stability of SAHs. It has been speculated that their stability arises from the flexibility of the ion pair network, providing it with an entropic advantage (9, 58). Our combined NMR data and MD simulations support this idea, suggesting that a rapidly fluctuating network of ion pairs is key to SAH stability. More broadly, it might be expected that solvent-exposed ion pairs in many other proteins are also likely to demonstrate this dynamic behavior in solution.

Experimental procedures

Expression constructs

The DNA sequence for residues 858–935 from myosin 7a (mouse, Uniprot P97479; Fig. 7*A*) was synthesized (GeneArt; GenScript) and subcloned into the pET28a SUMO vector to introduce an N-terminal His tag and SUMO fusion protein for increased expression and solubility as described (6). To facilitate concentration measurements of purified protein, a single C-terminal Trp was also incorporated. The purified protein segment M7A SAH is 80 residues in length, including the additional N-terminal Ser and C-terminal Trp residues.

Protein expression and purification

Protein expression was carried out in *E. coli* BL21 Rosetta 2 (Novagen), and purification used a nickel-nitrilotriacetic acid affinity chromatography column. Expressed protein was dialyzed against 150 mM NaCl, 20 mM Tris, 1 mM DTT, pH 8.0, and proteolyzed for 2 h at room temperature, using ULP1 recombinant SUMO protease in a substrate/enzyme ratio of 100:1. Cleavage of SUMO carries forward a single, N-terminal Ser residue into M7A SAH. M7A SAH was purified on a 5-ml SP Sepharose column using an AKTA system. Buffers used were as follows: 20 mM Tris-HCl, pH 7.5, 0.03% NaN$_3$ (Buffer A); 1 M NaCl, 20 mM Tris-HCl, pH 7.5, 0.03% NaN$_3$ (Buffer B); salt gradient: 100–500 mM. The purest fractions were combined and concentrated, resulting in a 2–10 mg/ml protein solution. Purified protein was dialyzed against 100 mM NaCl, 10 mM sodium phosphate (Na$_2$HPO$_4$/NaH$_2$PO$_4$), pH 7.4, and snap-frozen in liquid nitrogen for long-term storage at -80 °C. Protein concentration was measured by absorption at 280 nm. Absorption coefficients were obtained from ProtParam software.

Protein expression for NMR analysis was initially carried out in Terrific Broth medium (4 × 400 ml). After reaching A_{600} ~0.6, the cultures were spun down, and combined pellets were washed with 100–200 ml of minimal medium (M9 salts, 2 mM MgSO$_4$, 0.1 mM CaCl$_2$, 10 ml of basal vitamins) and resuspended in 250 ml of minimal medium. After 45 min at 37 °C at 220 rpm, the culture was supplemented with isotopes (100 mM ^{15}NH$_4$Cl, 50 mM ^{13}C-labeled D-glucose; Goss Scientific) and grown in the same conditions for an additional 35 min, after which culture was induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside and grown overnight at 21 °C and 180 rpm. Purification was carried out as described above. Two ^{15}N-^{13}C–labeled M7A SAH samples were prepared for the NMR experiments. The first was prepared in 100 mM NaCl, 10 mM sodium (hydrogen) phosphate, pH 7.4, at a concentration of ~4 mg/ml (0.4 mM). The second was prepared in 100 mM NaCl, 10 mM phosphate, at the lower pH of 5.5 and at a concentration of 10 mg/ml (1 mM). Between experiments, samples were stored at 5 °C.

Mass spectrometry

Samples of M7A SAH (0.5 ml; 15–20 μM) were dialyzed (G-Biosciences dialyzers, molecular weight cutoff 1000) overnight against 50 mM ammonium acetate, pH 7.4, and analyzed by TOF MS analysis (University of Leeds Mass Spectrometry Facility). The peptide masses from MS were as expected: 9.97 kDa for unlabeled and 10.54 kDa for ^{15}N-^{13}C–labeled.

CD spectroscopy

CD measurements were performed on an Applied Photo Physics Chirascan CD spectropolarimeter with a 0.1-cm path length quartz cuvette in buffers as specified under "Protein

expression and purification." Protein concentrations used were in the range 10–20 μM. Data were collected every 1 nm with a scan rate of 120 nm/min, each measurement being presented as averaged from at least two separate measurements of different protein preparations. Thermal measurements were performed over a temperature range of 10–85 °C with data acquisition every 1 °C using a 0.7 °C/min heating rate. The sample cooling rate prior to measurement of refolded protein was ~2 °C/min. The mean residue molar ellipticity (MRE) of proteins was calculated as described (59). The helical content of proteins was calculated from values of the amide nπ* transition at 222 nm ([MRE$_{222}$]), as described previously (59).

Size-exclusion chromatography

A GE Healthcare Tricorn 10/20 column was packed with Superdex 75 resin and calibrated using the GE Healthcare gel filtration calibration kit, which comprises albumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), RNase A (13.7 kDa), and aprotinin (6.5 kDa). The elution profile of M7A SAH was obtained by injecting 200 μl of protein sample (40 μM) in column buffer (150 mM NaCl, 10 mM sodium phosphate, 0.02% NaN$_3$, pH 7.4) onto the gel filtration column at a flow rate of 0.5 ml/min, using an AKTA system.

Analytical ultracentrifugation

Analytical ultracentrifugation sedimentation-equilibrium experiments were performed in triplicate as described previously (5, 7). A Beckman Optima XL-A analytical ultracentrifuge was used at a temperature of 20 °C with an AN50 8-place rotor and cells comprising epon 6-channel centerpieces and quartz windows. A 1 mg/ml sample of M7A SAH was prepared in 100 mM NaCl, 10 mM sodium (di)hydrogen phosphate, 0.03% NaN$_3$, pH 7.4 (110 μl). The reference channel contained sample buffer only (120 μl). Samples were centrifuged at speeds of 18,000, 22,000, 26,000, 30,000, 34,000, 38,000, and 42,000 rpm. Absorbance data at radial distances of 5.8–7.3 cm were recorded after 8 h at each speed and then again after a further 1 h before moving onto the next speed, to retrospectively confirm that the sample had reached equilibrium. Absorbance and residuals were plotted against $R^2 - R_0^2$, where R is the radial distance from the center of the rotor, and R_0 is an arbitrary reference radius. Data were fitted to a single ideal species using Ultrascan II (http://ultrascan2.uthscsa.edu/)[5] (60), and 99% confidence limits were determined by Monte Carlo analyses of the fits. Representative data for one channel are shown.

Small-angle X-ray scattering

Experiments combining size-exclusion chromatography SAXS were performed at Diamond light source beamline B21. A 2.4-ml Superdex 200 Increase 3.2/300 column (GE Healthcare) connected to an Agilent 1200 HPLC system was equilibrated at room temperature with 100 mM NaCl, 10 mM sodium phosphate, pH 7.4. 45 μl of 2 mg/ml M7a SAH construct was loaded onto the column. The output flow from the HPLC was directed through a quartz capillary cell held in vacuum. The

flow rate through the capillary cell was 0.075 ml/min. A fixed wavelength of 1.0 Å (12.4 keV) was used with the X-ray detector (PILATUS 2M) placed at a distance of 4 meters from the sample. Data were analyzed using Scatter (version 3.0a) (http://www.bioisis.net) (81).[5] The Kratky plot of the scattering data is $q^2I(q)$ *versus* q, whereas the Guinier plot is $\ln(I(q))$ *versus* q^2. I is the scattering intensity, q is the momentum transfer ($q = 4\pi\sin\theta/\lambda$, where λ is the beam wavelength and 2θ is the scattering angle). Scattering intensity curves were generated for high-helicity model structures (generated from simulation or CS-Rosetta; see below) using CRYSOL (18); comparison with the experimental data was evaluated using χ^2 values.

NMR spectroscopy

NMR spectra were recorded on 600- and 750-MHz Bruker Avance spectrometers or a 950-MHz Bruker Ascend Aeon spectrometer, each equipped with a cryoprobe. After testing the pH 7.4 sample with ^{1}H–^{15}N TROSY spectra, a temperature of 23.4 °C was initially used to balance maximum helical content (low temperature) with peak width (high temperature). The ^{1}H–^{15}N TROSY was repeated periodically to check for sample integrity; the sample remained stable in solution at 5 °C over a period of several months. A list of NMR experiments and details of their use are given (Table S1). Spectra were processed using NMRPipe/NMRDraw (61), and then peak assignments were carried out using CCPNmr Analysis (62).

The ^{1}H–^{15}N TROSY spectrum of M7A SAH (pH 7.4) only resolves about 55 backbone amide peaks (of a total of 79 backbone NH groups). Despite this, we were able to observe almost all backbone correlations in the 3D HNCO, allowing us to obtain an almost complete backbone assignment (96% of backbone NH correlations) through three pairs of standard BEST-TROSY style triple-resonance experiments (19–21) (acquired at 750 MHz). Assignment was supported by linking (i, $i + 1$) backbone amide peaks using an NOESY–^{15}N HSQC spectrum (63, 64). Further backbone and side chain shift assignments (Table S2) were carried out using a number of additional NMR experiments (Table S1). The chemical shifts have been deposited in the Biological Magnetic Resonance Bank (entry number 27626). Many assignments, particularly for the large numbers of Arg, Lys, and Glu side chain nuclei, could only be made by relaying the magnetization back to the backbone amides (by means of 3D H(CCCO)NH and (H)C(CCO)NH experiments (65, 66)).

The HG(CG)CD "Hγ(Cγ)Cδ" experiment (Fig. 3B), used to interrogate Glu side chains, was based on a modified HCACO experiment (67) with the "Cα" position shifted to 36 ppm, resulting in the selective transfer of magnetization from Glu Hγ via Cγ to Cδ (68). This spectrum also shows Gln Hγ(Cγ)Cδ and Asp Hβ(Cβ)Cγ correlations, which act as useful shift markers to link with other spectra (*e.g.* ^{13}C HSQC). Specific assignment for all Glu Cδ nuclei and thereby a significant number of Glu Hγ pairs was achieved by linking Cδ shifts back to the backbone using a 3D (HγCγ)Cδ(CγCβCαCO)NH experiment (69). Assignment was assisted in some cases through linking Glu Hγ pairs using a NOESY-HG(CG)CD experiment: an NOE transfer period was added prior to the pulse sequence for the HG(CG)CD experiment described above. For some correla-

---

[5] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

tions in the highly overlapped central region of the Hγ(Cγ)Cδ spectrum, unambiguous Glu Hγ assignment was not possible by this method, so assignment relied on the much less well-resolved H(CCCO)NH spectrum, resulting in those Hγ pairs appearing to have degenerate shifts.

The backbone assignments were transferred from the pH 7.4 sample to the pH 5.5 sample, by comparing HNCO/HN-(CA)CO pairs of triple-resonance experiments (recorded at 23.4 and 10 °C, pH 5.5), and supplemented by following peak movements in a set of TROSY spectra recorded at ∼3 °C intervals (from 23.4 °C down to 10 °C). It was then possible to assign each of the Hε–Nε correlations to specific Arg residues by matching side chain proton resonances in TOCSY-Nε-HSQC and TROSY H(CCCO)NH spectra at 23.4 °C (see Fig. S3). Additional support was given by links to local distinctive proton resonances from an Nε-NOESY-HSQC spectrum at 10 °C.

Secondary chemical shifts (*i.e.* the difference between the measured Cα shift and a reference Cα shift for that residue in a coil configuration) were calculated using reference values from Schwarzinger *et al.* (70) as implemented in the CCPNmr Analysis software.

Longitudinal (T_1) and transverse (T_2) relaxation rates (R_1 and R_2, respectively) and heteronuclear NOEs for backbone ^{15}N nuclei were measured using 750- and 950-MHz instruments. Equivalent experiments for Arg Nε were carried out using the 950-MHz instrument. The recycle delay in T_1 and T_2 experiments was 2.5 s (one test T_1 experiment using a recycle delay of 5.0 s gave no measurable difference in results from those using 2.5 s). The relaxation periods ranged from 20 ms to 1.6 s for T_1 experiments and from ∼16 to 237 ms for T_2 experiments. Eight different relaxation periods were used for each experiment, two of which were duplicated to enable error estimations. The heteronuclear NOE experiments used a recycle delay of 5 s. Measurements were limited to those residues exhibiting resolvable peaks in the HSQC spectrum; those broader peaks that often appear partially overlapped with sharper peaks were omitted (several examples between residues 876 and 896). Peak intensities were measured using NMRView (71). Analysis of relaxation rates was performed using in-house Python scripts, with Monte Carlo–based fitting and error estimation. Using a much-simplified isotropic tumbling model, estimates of the rotational correlation time (τ_r) for the protein were calculated from the average R_2/R_1 ratio for stably folded residues (NOE values > 0.6) with R_2 and R_1 errors <10% (27, 28).

The CS-Rosetta server (25, 26) was used to generate an ensemble of structures from chemical shift assignments of ^{1}H$_N$, ^{1}Hα, ^{13}Cα, ^{13}Cβ, ^{13}C′, and ^{15}N nuclei using the default method. Different runs yielding 3000, 10,000, or 30,000 structures each gave the same result. Structures with C_α root mean square deviations >2 Å away from the lowest-energy structure all have noticeably higher energies.

The pulse program used for measuring $^3J_{NC}$ coupling constants (Fig. S4) is based on the experiment used by Zandarashvili *et al.* (42). The carrier positions and pulses were altered to be appropriate for Arg Nε rather than LysNζ. Dephasing, resulting from $^1J_{N\epsilon C\delta}$ and $^1J_{N\epsilon C\zeta}$, was refocused by means of two pairs of shaped pulses. Carrier positions were as follows: ^{1}H, the position of the water resonance; Arg ^{15}Nε, 84 ppm; Arg ^{13}Cδ, 42

or 41 ppm; Arg ^{13}Cβ, 28 ppm; Arg ^{13}Cζ, 158 ppm; and Glu ^{13}Cδ (C′), 181 ppm. IBURP-2 180° shaped pulses were used (pulse lengths for 950 MHz): Arg ^{13}Cγ (2.37 ms), Arg ^{13}Cζ (2.37 ms), Arg ^{13}Cβ (2.37 ms), and Glu ^{13}Cδ (1.185 ms). Two sub-spectra ("reference" and "attenuated") were recorded for each experiment that differ only in the position of the shaped pulses for Arg Cβ or Glu Cδ (see Fig. S4). $^3J_{NC}$ coupling constants were calculated using the expression $J = (1/\pi T)\arccos(I_{att}/I_{ref})$, where I_{att} and I_{ref} are the peak intensities in the attenuated and reference sub-spectra (42). The net evolution time, T, was 208 ms in the $^3J_{N\epsilon C\beta}$ experiment and 424 ms in the $^3J_{N\epsilon C'}$ experiment. Peak intensities were measured using PINT (72, 73).

Modeling

Simulations were performed using the CHARMM36 force field parameters with TIP3P water. The initial M7A SAH structure was built as a perfect α-helix (internal dihedrals Φ = −57° and Ψ = −47°). N-terminal Ser and C-terminal Trp residues were included, and the termini were uncapped. Structures were energy-minimized for 1000 steepest decent steps in vacuum using CHARMM (74). Using VMD (75), a 1.5-nm surround of water molecules (11,906 water molecules) and Na$^+$ and Cl$^-$ ions were added to neutralize the peptide and give a NaCl concentration of ∼150 mM. A further minimization (10,000 steps), 0–300 K heating protocol, and short pre-equilibration (100,000 steps) was performed using NAMD (76). Data are taken from a 500-ns simulation run using NAMD at 300 K (27 °C). The time step used was 2 fs, and trajectory frames were recorded every 500 steps. We also ran simulations starting from the two chains from the PDB structure 5WST (kindly supplied by Professor Mingjie Zhang's group). These structures contain residues 862–932 (chain A) or residues 863–932 (chain B). N and C termini were capped with acetyl and methyl amino groups, respectively. Structures were solvated, ions were added, and simulations were initiated in the manner described above and ran for 200 ns.

Analysis of the simulation trajectories made extensive use of Wordom (77). The helicity (or average helical fraction) of the peptide was calculated using the DSSPcont criteria (78). A Φ- and Ψ-angle–based method for the helicity calculation was also used with similar results (79). For analysis of ion pairs, the distance between lysine Nζ atoms and the centroid of glutamate Oε1 and Oε2 atoms was calculated for each potential E$_i$–K$_{i+3}$, K$_i$–E$_{i+3}$, E$_i$–K$_{i+4}$, and K$_i$–E$_{i+4}$ pair; the distances between each of the three arginine NH1/NH2/Nε atoms and the centroid of glutamate Oε1 and Oε2 atoms were calculated for each potential E$_i$–R$_{i+3}$, R$_i$–E$_{i+3}$, E$_i$–R$_{i+4}$, and R$_i$–E$_{i+4}$ pair. The definition of a contact ion pair at any frame of the trajectory required any of the resulting distances described to be <4 Å. VMD was used to produce snapshot images (75).

Ion pair dynamics in helices

References

1. Swanson, C. J., and Sivaramakrishnan, S. (2014) Harnessing the unique structural properties of isolated α-helices. *J. Biol. Chem.* **289,** 25460–25467 CrossRef Medline
2. Knight, P. J., Thirumurugan, K., Xu, Y., Wang, F., Kalverda, A. P., Stafford W. F 3rd, Sellers, J. R., and Peckham, M. (2005) The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. *J. Biol. Chem.* **280,** 34702–34708 CrossRef Medline
3. Peckham, M., and Knight, P. J. (2009) When a predicted coil is really a single α-helix, in myosins and other proteins. *Soft Matter* **5,** 2493–2503
4. Sivaramakrishnan, S., Spink, B. J., Sim, A. Y., Doniach, S., and Spudich, J. A. (2008) Dynamic charge interactions create surprising rigidity in the ER/K α-helical protein motif. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 13356–13361 CrossRef Medline
5. Wolny, M., Batchelor, M., Bartlett, G. J., Baker, E. G., Kurzawa, M., Knight, P. J., Dougan, L., Woolfson, D. N., Paci, E., and Peckham, M. (2017) Characterization of long and stable *de novo* single α-helix domains provides novel insight into their stability. *Sci. Rep.* **7,** 44341 CrossRef Medline
6. Wolny, M., Batchelor, M., Knight, P. J., Paci, E., Dougan, L., and Peckham, M. (2014) Stable single α-helices are constant force springs in proteins. *J. Biol. Chem.* **289,** 27825–27835 CrossRef Medline
7. Baker, E. G., Bartlett, G. J., Crump, M. P., Sessions, R. B., Linden, N., Faul, C. F., and Woolfson, D. N. (2015) Local and macroscopic electrostatic interactions in single α-helices. *Nat. Chem. Biol.* **11,** 221–228 CrossRef Medline
8. Batchelor, M., Wolny, M., Dougan, L., Paci, E., Knight, P. J., and Peckham, M. (2015) Myosin tails and single α-helical domains. *Biochem. Soc. Trans.* **43,** 58–63 CrossRef Medline
9. Süveges, D., Gáspári, Z., Tóth, G., and Nyitray, L. (2009) Charged single α-helix: a versatile protein structural motif. *Proteins* **74,** 905–916 CrossRef Medline
10. Gáspári, Z., Süveges, D., Perczel, A., Nyitray, L., and Tóth, G. (2012) Charged single α-helices in proteomes revealed by a consensus prediction approach. *Biochim. Biophys. Acta* **1824,** 637–646 CrossRef Medline
11. Simm, D., Hatje, K., and Kollmar, M. (2015) Waggawagga: comparative visualization of coiled-coil predictions and detection of stable single α-helices (SAH domains). *Bioinformatics* **31,** 767–769 CrossRef Medline
12. Simm, D., Hatje, K., and Kollmar, M. (2017) Distribution and evolution of stable single α-helices (SAH domains) in myosin motor proteins. *PLoS One* **12,** e0174639 CrossRef Medline
13. Li, J., Chen, Y., Deng, Y., Unarta, I. C., Lu, Q., Huang, X., and Zhang, M. (2017) Ca²⁺-induced rigidity change of the myosin VIIa IQ motif-single α helix lever arm extension. *Structure* **25,** 579–591.e4 CrossRef Medline
14. Lu, Q., Ye, F., Wei, Z., Wen, Z., and Zhang, M. (2012) Antiparallel coiled-coil-mediated dimerization of myosin X. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 17388–17393 CrossRef Medline
15. Ropars, V., Yang, Z., Isabet, T., Blanc, F., Zhou, K., Lin, T., Liu, X., Hissier, P., Samazan, F., Amigues, B., Yang, E. D., Park, H., Pylypenko, O., Cecchini, M., Sindelar, C. V., Sweeney, H. L., and Houdusse, A. (2016) The myosin X motor is optimized for movement on actin bundles. *Nat. Commun.* **7,** 12456 CrossRef Medline
16. Spink, B. J., Sivaramakrishnan, S., Lipfert, J., Doniach, S., and Spudich, J. A. (2008) Long single α-helical tail domains bridge the gap between structure and function of myosin VI. *Nat. Struct. Mol. Biol.* **15,** 591–597 CrossRef Medline
17. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40,** 191–285 Medline
18. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28,** 768–773 CrossRef
19. Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D., and Brutscher, B. (2013) BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J. Biomol. NMR* **55,** 311–321 CrossRef Medline
20. Lescop, E., Schanda, P., and Brutscher, B. (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187,** 163–169 CrossRef Medline
21. Favier, A., and Brutscher, B. (2011) Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J. Biomol. NMR* **49,** 9–15 CrossRef Medline
22. Jie, J., Löhr, F., and Barbar, E. (2015) Interactions of yeast dynein with dynein light chain and dynactin: general implications for intrinsically disordered duplex scaffolds in multiprotein assemblies. *J. Biol. Chem.* **290,** 23863–23874 CrossRef Medline
23. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222,** 311–333 CrossRef Medline
24. Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* **51,** 2224–2231 CrossRef Medline
25. Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 4685–4690 CrossRef Medline
26. Shen, Y., Vernon, R., Baker, D., and Bax, A. (2009) *De novo* protein structure generation from incomplete chemical shift assignments. *J. Biomol. NMR* **43,** 63–78 CrossRef Medline
27. Worrall, J. A., Liu, Y., Crowley, P. B., Nocek, J. M., Hoffman, B. M., and Ubbink, M. (2002) Myoglobin and cytochrome b_5: a nuclear magnetic resonance study of a highly dynamic protein complex. *Biochemistry* **41,** 11721–11730 CrossRef Medline
28. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28,** 8972–8979 CrossRef Medline
29. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., 3rd, Rance, M., and Skelton, N. J. (2007) *Protein NMR Spectroscopy: Principles and Practice*, 2nd Ed., p. 21, Academic Press, Inc., Amsterdam
30. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2003) ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19,** 315–316 CrossRef Medline
31. Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339,** 71–90 CrossRef Medline
32. Nilges, M., Macias, M. J., O'Donoghue, S. I., and Oschkinat, H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β-spectrin. *J. Mol. Biol.* **269,** 408–422 CrossRef Medline
33. Bermel, W., Bertini, I., Felli, I. C., Lee, Y. M., Luchinat, C., and Pierattelli, R. (2006) Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. *J. Am. Chem. Soc.* **128,** 3918–3919 CrossRef Medline
34. Bermel, W., Bruix, M., Felli, I. C., Kumar, M. V. V., Pierattelli, R., and Serrano, S. (2013) Improving the chemical shift dispersion of multidimen-

sional NMR spectra of intrinsically disordered proteins. *J. Biomol. NMR* **55,** 231–237 CrossRef Medline

35. Platzer, G., Okon, M., and McIntosh, L. P. (2014) pH-dependent random coil ^{1}H, ^{13}C, and ^{15}N chemical shifts of the ionizable amino acids: a guide for protein pKa measurements. *J. Biomol. NMR* **60,** 109–129 CrossRef Medline

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

37. Mackenzie, H. W., and Hansen, D. F. (2017) A ^{13}C-detected ^{15}N double-quantum NMR experiment to probe arginine side-chain guanidinium ^{15}N$^\eta$ chemical shifts. *J. Biomol. NMR* **69,** 123–132 CrossRef Medline

38. Nguyen, D., Hoffpauir, Z. A., and Iwahara, J. (2017) Internal motions of basic side chains of the Antennapedia homeodomain in the free and DNA-bound states. *Biochemistry* **56,** 5866–5869 CrossRef Medline

39. Werbeck, N. D., Kirkpatrick, J., and Hansen, D. F. (2013) Probing arginine side-chains and their dynamics with carbon-detected NMR spectroscopy: application to the 42 kDa human histone deacetylase 8 at high pH. *Angew. Chem. Int. Ed. Engl.* **52,** 3145–3147 CrossRef Medline

40. Yoshimura, Y., Oktaviani, N. A., Yonezawa, K., Kamikubo, H., and Mulder, F. A. (2017) Unambiguous determination of protein arginine ionization states in solution by NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **56,** 239–242 CrossRef Medline

41. Iwahara, J., Jung, Y. S., and Clore, G. M. (2007) Heteronuclear NMR spectroscopy for lysine NH$_3$ groups in proteins: unique effect of water exchange on ^{15}N transverse relaxation. *J. Am. Chem. Soc.* **129,** 2971–2980 CrossRef Medline

42. Zandarashvili, L., Li, D. W., Wang, T., Brüschweiler, R., and Iwahara, J. (2011) Signature of mobile hydrogen bonding of lysine side chains from long-range ^{15}N-^{13}C scalar J-couplings and computation. *J. Am. Chem. Soc.* **133,** 9192–9195 CrossRef Medline

43. Pascal, S. M., Yamazaki, T., Singer, A. U., Kay, L. E., and Forman-Kay, J. D. (1995) Structural and dynamic characterization of the phosphotyrosine binding region of a Src homology 2 domain–phosphopeptide complex by NMR relaxation, proton exchange, and chemical shift approaches. *Biochemistry* **34,** 11353–11362 CrossRef Medline

44. Trbovic, N., Cho, J. H., Abel, R., Friesner, R. A., Rance, M., and Palmer, A. G., 3rd (2009) Protein side-chain dynamics and residual conformational entropy. *J. Am. Chem. Soc.* **131,** 615–622 CrossRef Medline

45. Wagner, G., Pardi, A., and Wuethrich, K. (1983) Hydrogen bond length and proton NMR chemical shifts in proteins. *J. Am. Chem. Soc.* **105,** 5948–5949 CrossRef

46. Liu, A., Hu, W., Majumdar, A., Rosen, M. K., and Patel, D. J. (2000) NMR detection of side chain-side chain hydrogen bonding interactions in ^{13}C/^{15}N-labeled proteins. *J. Biomol. NMR* **17,** 305–310 CrossRef Medline

47. Birdsall, B., Polshakov, V. I., and Feeney, J. (2000) NMR studies of ligand carboxylate group interactions with arginine residues in complexes of *Lactobacillus casei* dihydrofolate reductase with substrates and substrate analogues. *Biochemistry* **39,** 9819–9825 CrossRef Medline

48. Cordier, F., and Grzesiek, S. (1999) Direct observation of hydrogen bonds in proteins by interresidue 3hJNC′ scalar couplings. *J. Am. Chem. Soc.* **121,** 1601–1602 CrossRef

49. Cordier, F., Nisius, L., Dingley, A. J., and Grzesiek, S. (2008) Direct detection of N–H[. . .]O=C hydrogen bonds in biomolecules by NMR spectroscopy. *Nat. Protoc.* **3,** 235–241 CrossRef Medline

50. Zandarashvili, L., Esadze, A., Kemme, C. A., Chattopadhyay, A., Nguyen, D., and Iwahara, J. (2016) Residence times of molecular complexes in solution from NMR data of intermolecular hydrogen-bond scalar coupling. *J. Phys. Chem. Lett.* **7,** 820–824 CrossRef Medline

51. Lang, P. T., Ng, H. L., Fraser, J. S., Corn, J. E., Echols, N., Sales, M., Holton, J. M., and Alber, T. (2010) Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein Sci.* **19,** 1420–1431 CrossRef Medline

52. Ghosh, T., Garde, S., and García, A. E. (2003) Role of backbone hydration and salt-bridge formation in stability of α-helix in solution. *Biophys. J.* **85,** 3187–3193 CrossRef Medline

53. Tomlinson, J. H., Ullah, S., Hansen, P. E., and Williamson, M. P. (2009) Characterization of salt bridges to lysines in the protein G B1 domain. *J. Am. Chem. Soc.* **131,** 4674–4684 CrossRef Medline

54. Williamson, M. P., Hounslow, A. M., Ford, J., Fowler, K., Hebditch, M., and Hansen, P. E. (2013) Detection of salt bridges to lysines in solution in barnase. *Chem. Commun.* **49,** 9824–9826 CrossRef Medline

55. Ahmed, M. C., Papaleo, E., and Lindorff-Larsen, K. (2018) How well do force fields capture the strength of salt bridges in proteins? *PeerJ* **6,** e4967 CrossRef Medline

56. Debiec, K. T., Gronenborn, A. M., and Chong, L. T. (2014) Evaluating the strength of salt bridges: a comparison of current biomolecular force fields. *J. Phys. Chem. B* **118,** 6561–6569 CrossRef Medline

57. Piana, S., Lindorff-Larsen, K., and Shaw, D. E. (2011) How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **100,** L47–L49 CrossRef Medline

58. Ulrich, A. K. C., Seeger, M., Schütze, T., Bartlick, N., and Wahl, M. C. (2016) Scaffolding in the spliceosome via single α helices. *Structure* **24,** 1972–1983 CrossRef Medline

59. Greenfield, N., and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8,** 4108–4116 CrossRef Medline

60. Gorbet, G., Devlin, T., Hernandez Uribe, B. I., Demeler, A. K., Lindsey Z. L., Ganji, S., Breton, S., Weise-Cross, L., Lafer, E. M., Brookes, E. H., and Demeler, B. (2014) A parametrically constrained optimization method for fitting sedimentation velocity experiments. *Biophys. J.* **106,** 1741–1750 CrossRef Medline

61. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293 Medline

62. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59,** 687–696 CrossRef Medline

63. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J. Am. Chem. Soc.* **111,** 1515–1517 CrossRef

64. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114,** 10663–10665 CrossRef

65. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of backbone amide and aliphatic side-chain resonances in ^{13}C/^{15}N-enriched proteins by isotropic mixing of ^{13}C magnetization. *J. Magn. Reson. Ser. B* **101,** 114–119 CrossRef

66. Montelione, G. T., Lyons, B. A., Emerson, S. D., and Tashiro, M. (1992) An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *J. Am. Chem. Soc.* **114,** 10974–10975 CrossRef

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

68. Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y., and Nakamura, H. (1994) Individual ionization constants of all the carboxyl groups in ribonuclease HI from *Escherichia coli* determined by NMR. *Biochemistry* **33,** 5275–5284 CrossRef Medline

69. Tollinger, M., Forman-Kay, J. D., and Kay, L. E. (2002) Measurement of side-chain carboxyl pK_a values of glutamate and aspartate residues in an unfolded protein by multinuclear NMR spectroscopy. *J. Am. Chem. Soc.* **124,** 5714–5717 CrossRef Medline

70. Schwarzinger, S., Kroon, G. J., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* **123,** 2970–2978 CrossRef Medline

71. Johnson, B. A., and Blevins, R. A. (1994) NMR View: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4,** 603–614 CrossRef Medline

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

ASBMB

73. Niklasson, M., Otten, R., Ahlner, A., Andresen, C., Schlagnitweit, J., Petzold, K., and Lundström, P. (2017) Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* **69,** 93–99 CrossRef Medline

74. Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., *et al.* (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30,** 1545–1614 CrossRef Medline

75. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38, 27–28 CrossRef Medline

76. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 CrossRef Medline

77. Seeber, M., Felline, A., Raimondi, F., Muff, S., Friedman, R., Rao, F., Caflisch, A., and Fanelli, F. (2011) Wordom: a user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. *J. Comput. Chem.* **32,** 1183–1194 CrossRef Medline

78. Carter, P., Andersen, C. A., and Rost, B. (2003) DSSPcont: continuous secondary structure assignments for proteins. *Nucleic Acids Res.* **31,** 3293–3295 CrossRef Medline

79. Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., and Mackerell, A. D., Jr. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* **8,** 3257–3273 CrossRef Medline

80. Helmus, J. J., and Jaroniec, C. P. (2013) Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *J. Biomol. NMR* **55,** 355–367 CrossRef Medline

81. Rambo, R. P., and Tainer, J. A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496,** 477–481 CrossRef Medline

ASBMB

Dynamic ion pair behavior stabilizes single α-helices in proteins

Matthew Batchelor, Marcin Wolny, Emily G. Baker, Emanuele Paci, Arnout P. Kalverda and Michelle Peckham

Access the most updated version of this article at doi: 10.1074/jbc.RA118.006752

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 80 references, 8 of which can be accessed free at
http://www.jbc.org/content/294/9/3219.full.html#ref-list-1

Supporting information for *Dynamic ion pair behavior stabilizes single α-helices in proteins*.

Matthew Batchelor,[1] Marcin Wolny,[1] Emily G. Baker,[2] Emanuele Paci,[1] Arnout P. Kalverda,[1] and Michelle Peckham*[1]

From the [1] School of Molecular and Cellular Biology and the Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK; [2] School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK

* To whom correspondence should be addressed: Michelle Peckham, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, UK, LS2 9JT, m.peckham@leeds.ac.uk, Tel. 44 (0)1133434348, Fax. 44 (0)1133434228

Supporting Information. Fig. S1: Initial experimental evidence that M7A SAH forms a stable α-helix in solution. Fig. S2: Fitting model helical structures to SAXS data. Fig. S3: Strategy for assignment of Nε–Hε correlations. Fig. S4: Pulseprogram for $^3J_{NC}$ coupling experiment. Fig. S5: Qualitative relationship between χ_3 dihedral and $^3J_{NC}$, selective MD results for χ_3. Table S1: NMR experiment list. Table S2: List of chemical shifts. Table S3: δ2D output file for M7A SAH backbone chemical shift data. Table S4: Simulation Arg–Glu and Lys–Glu ion pair occupancies. Movie S1: mpg movie related to Fig. 6.
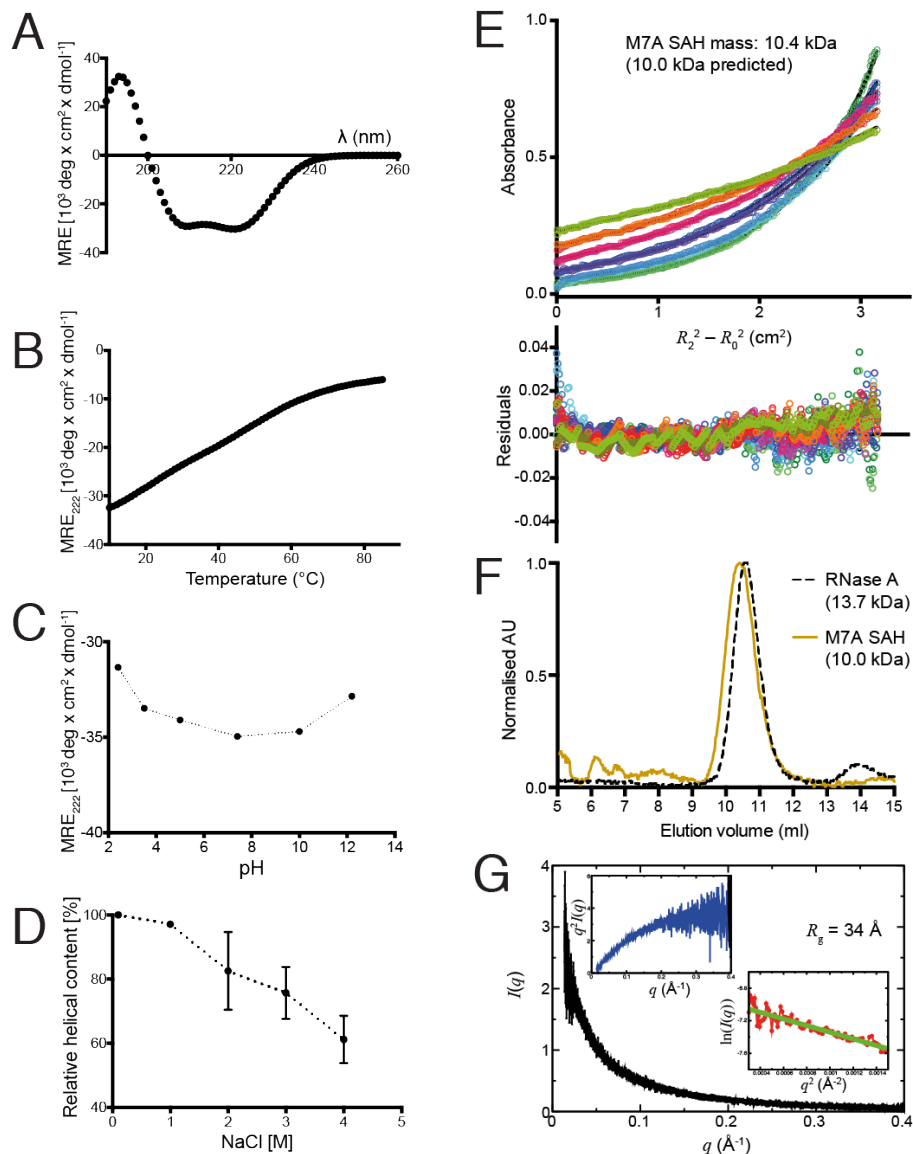
Figure S1. Initial experimental evidence that M7A SAH forms a stable α-helix in solution. (A) CD spectrum (pH 7.4, 10 °C) of M7A SAH shows the highly helical nature of the construct. **(B)** Thermal denaturation of M7A SAH. Mean residue ellipticity at $\lambda = 222$ nm (MRE_{222}) values at each temperature. On heating, the MRE_{222} reduced non-cooperatively with increase in temperature from 10 to 65 °C. **(C)** MRE_{222} remains high over a wide range of pH. **(D)** Salt stability of M7A SAH. The relative helical content (MRE_{222} values) is plotted as a percentage relative to the value measured at 0.1 M NaCl. Helical content remains high up to 4 M NaCl. **(E)** Results from equilibrium analytical ultracentrifugation for M7A SAH. Duplicated results are shown for rotor speeds: 22 krpm (pale greens), 26 krpm (orange/red), 30 krpm (pinks), 34 krpm (purple/blue), 38 krpm (pale blues), and 42 krpm (greens). Fitting the data (see methods) gave the M7A SAH mass as 10.390 kDa (with 99% confidence limits of 10.306–10.482 kDa) *cf.* the predicted mass of 10.0 kDa for a monomer. **(F)** Size exclusion chromatograph of M7A SAH. For comparison, the result for a globular protein (ribonuclease A/RNAse A, 13.7 kDa) is also shown. The faster elution of the M7A SAH than this reference globular protein with larger mass is consistent with an elongated structure. **(G)** Small angle X-ray scattering (SAXS) data: raw intensity plot (black), Kratky plot (blue) and Guinier plot (red). The Guinier plot gradient (green) is used to estimate the radius of gyration ($R_g = 34$ Å) under the Guinier approximation.
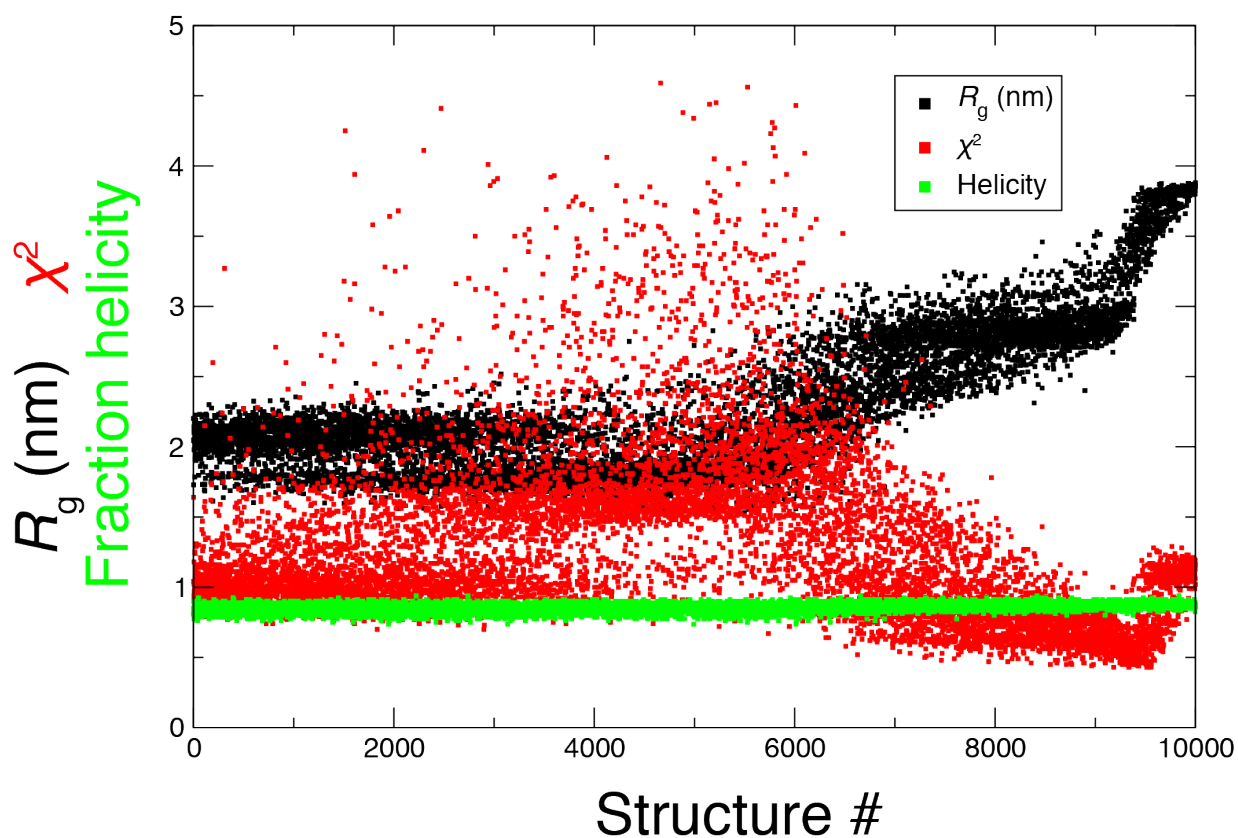
Figure S2. Fitting model helical structures to SAXS data. Helix-rich structures generated from CS-Rosetta analysis (and from simulation) were used to generate predicted SAXS data that were then compared with the experimental data using CRYSOL. Results are shown for 10,000 structures, which ranged from collapsed helical bundles (with low R_g) to fully extended and linear single helices. All have a high helicities: 0.85 ± 0.03 (mean $\pm$ S.D.). χ^2 values for $I(q)$ fits to the experimental data are lowest for the structures that are nearly (but not quite) perfect single helices.
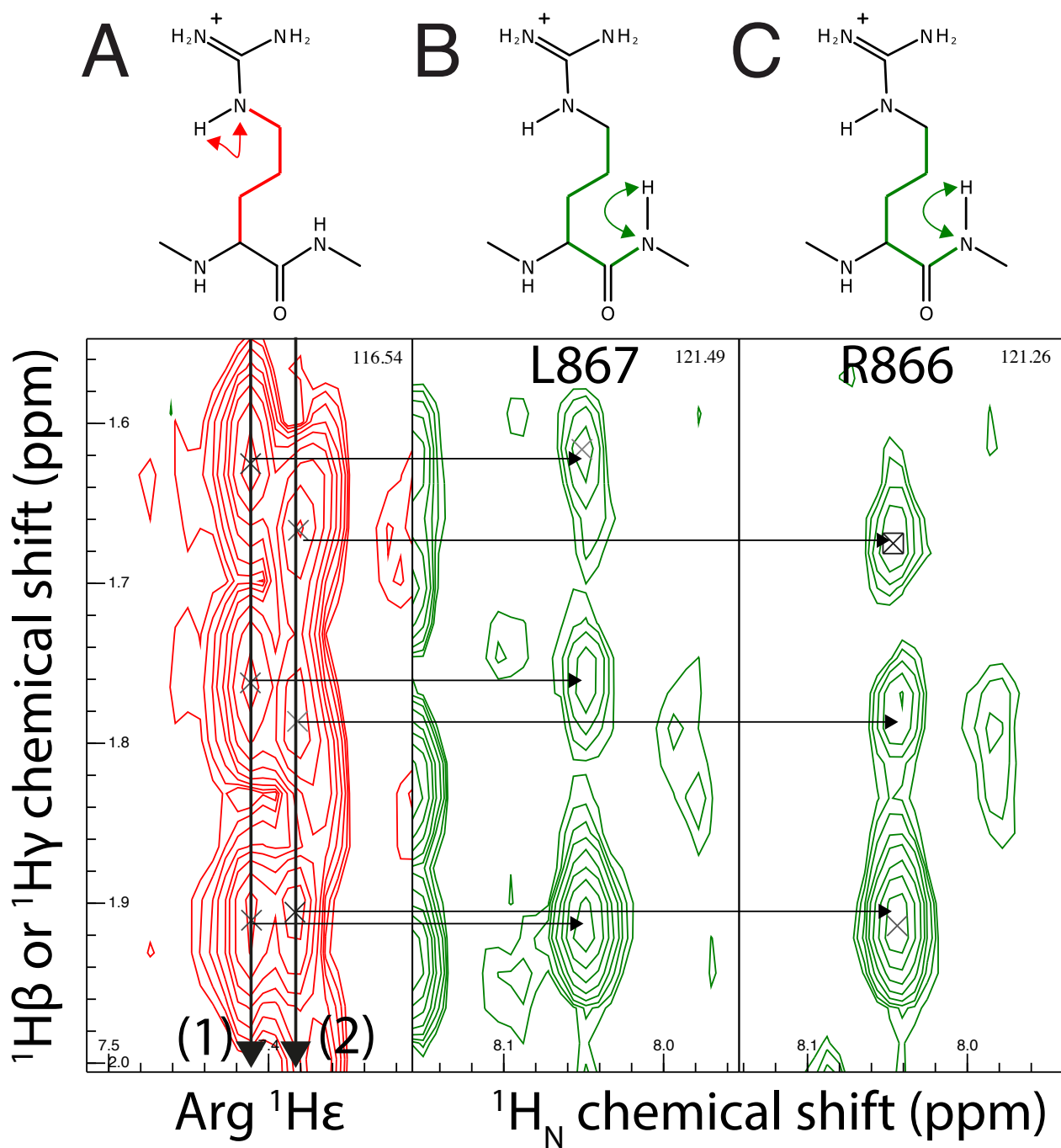
Figure S3. Residue assignment of Nε–Hε correlations. The TOCSY-N-HSQC spectrum (A, red) links the Arg aliphatic proton shifts along the side chain (Hα, Hβ, Hγ, Hδ) to Nε–Hε peaks in the ^{15}N-HiSQC NMR spectra (Fig. 4C). The TROSY H(CCCO)NH spectrum (B, C, green) links the same aliphatic proton shifts to the assigned N–H backbone correlation for the (i + 1) residue in the TROSY spectrum (Fig. 1A and 1B). The third dimension enumerated at the top right of each panel is the ^{15}N chemical shift of the Arg Nε (A), or backbone ^{15}N shift from (B, C). Panel A shows the Hβ and Hγ region from the TOCSY-N-HSQC spectrum for two neighboring Nε–Hε correlations, which appear as two sets of 3 resolved peaks (1) and (2). All the Hβ and Hγ chemical shifts (and Hα and Hδ shifts, not shown) must match with those from the (i + 1) backbone NH in the TROSY H(CCCO)NH spectrum (panel B and C). In the example shown the matching sets of proton shift are from L867 and R866 on the backbone, leading to assignment of the Nε–Hε peaks as R866 (1) and R865 (2), respectively.

4
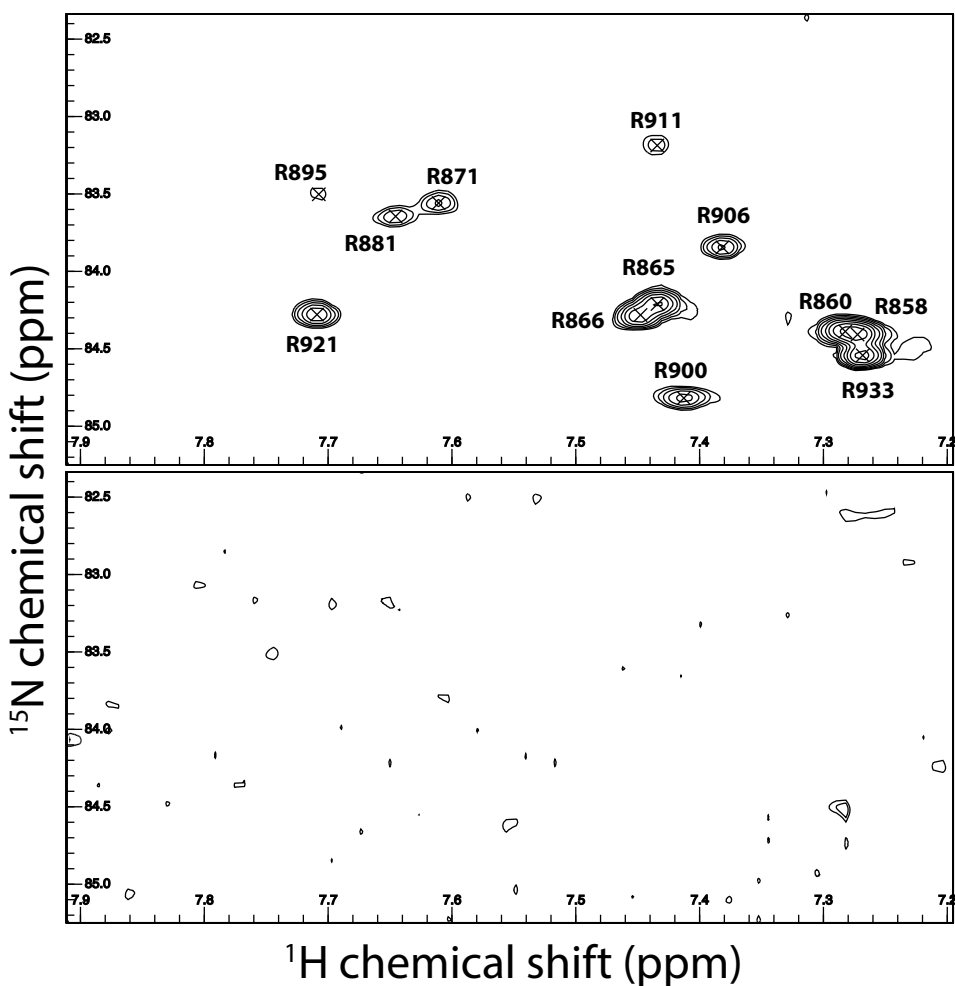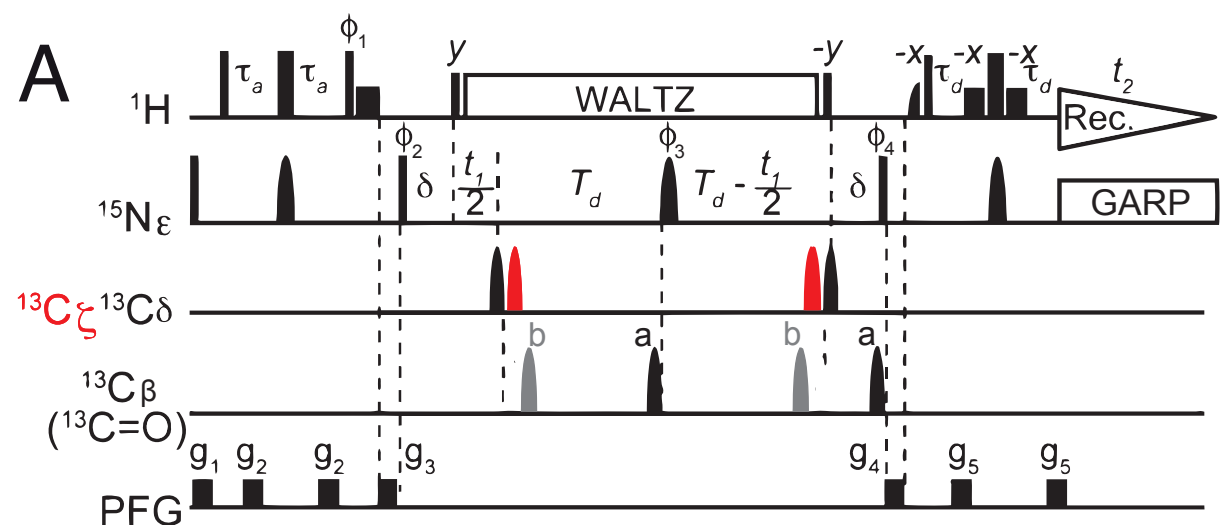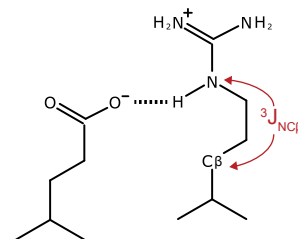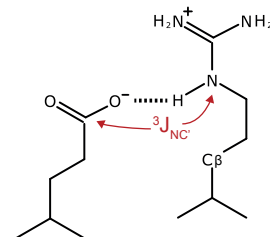
Figure S4. Pulseprogram and difference spectra for the constant time spin-echo difference experiment to measure $^3J_{NC}$ couplings from Nε–Hε correlations in arginine. The pulseprogram **(A)** is based on the experiment used by Zandarashvili *et al*. (1), with the carbon pulses altered to be appropriate for Arg (rather than Lys) and an additional pair of shaped pulses included to account for the Arg Cζ nuclei (red). Carrier positions were: ^{1}H, the position of the water resonance; Arg ^{15}Nε, 84 ppm; Arg ^{13}Cδ, 42 or 41 ppm; Arg ^{13}Cβ, 28 ppm; Arg ^{13}Cζ, 158 ppm and Glu ^{13}Cδ, 181 ppm. IBURP-2 180° shaped pulses were used (pulselengths at 950 MHz): Arg ^{13}Cδ (2.37 ms), Arg ^{13}Cζ (2.37 ms), Arg ^{13}Cβ (2.37 ms) and

Glu ^{13}Cδ (1.185 ms). 1 ms ^{1}H sinc shaped pulses were used for waterflipback and Watergate. ^{15}N r-SNOB pulses (2.02 ms) were used for selective inversion of ^{15}Nε. Delays: τ_a, τ_d = 2.7 ms, δ= 2.6 ms. T_d depends on the balance between making small couplings observable and the relaxation rate of ^{15}Nε. ^{1}H decoupling: 3.125 kHz WALTZ64. ^{15}N decoupling: 1.3 kHz GARP. Pulsed field gradients are applied with strengths of g_1: 9.7 G cm^{-1}, g_2: 15.6 G cm^{-1}, g_3: 22.7 G cm^{-1}, g_4: 19.5 G cm^{-1}, g_5: 42.2 G cm^{-1}. The gradient length was 0.3 ms for g_2 and 1 ms for the others. Phase cycles: ϕ_1 = {y, $-y$}, ϕ_2 = {x, x, $-x$, $-x$}, ϕ_3 = {$8x$, $8y$, 8($-x$), 8($-y$)}, ϕ_4 = ($4x$, $4(-x)$), receiver = {x, $-x$, $-x$, x, 2($-x$, x, x, $-x$), x, $-x$, $-x$, x}. The net evolution time, T, is 2(T_d+δ). Two sub-spectra were recorded with shaped pulses for Arg Cβ or Glu Cδ) located at either positions 'a' (black) or 'b' (grey). Figure adapted from ref. (1). **(B)** Difference spectrum (reference spectrum – attenuated spectrum) to measure intra side chain coupling constants for Arg Nε to Arg Cβ. The presence of a peak in the difference spectrum indicates a measurable coupling constant. **(C)** Difference spectrum to measure through-hydrogen-bond coupling constants for Arg Nε to Glu Cδ. The absence of peaks in the difference spectrum sets a maximum coupling constant based on the level of the baseline noise and the intensity of visible peaks in the sub-spectra; this was conservatively 0.2 Hz.
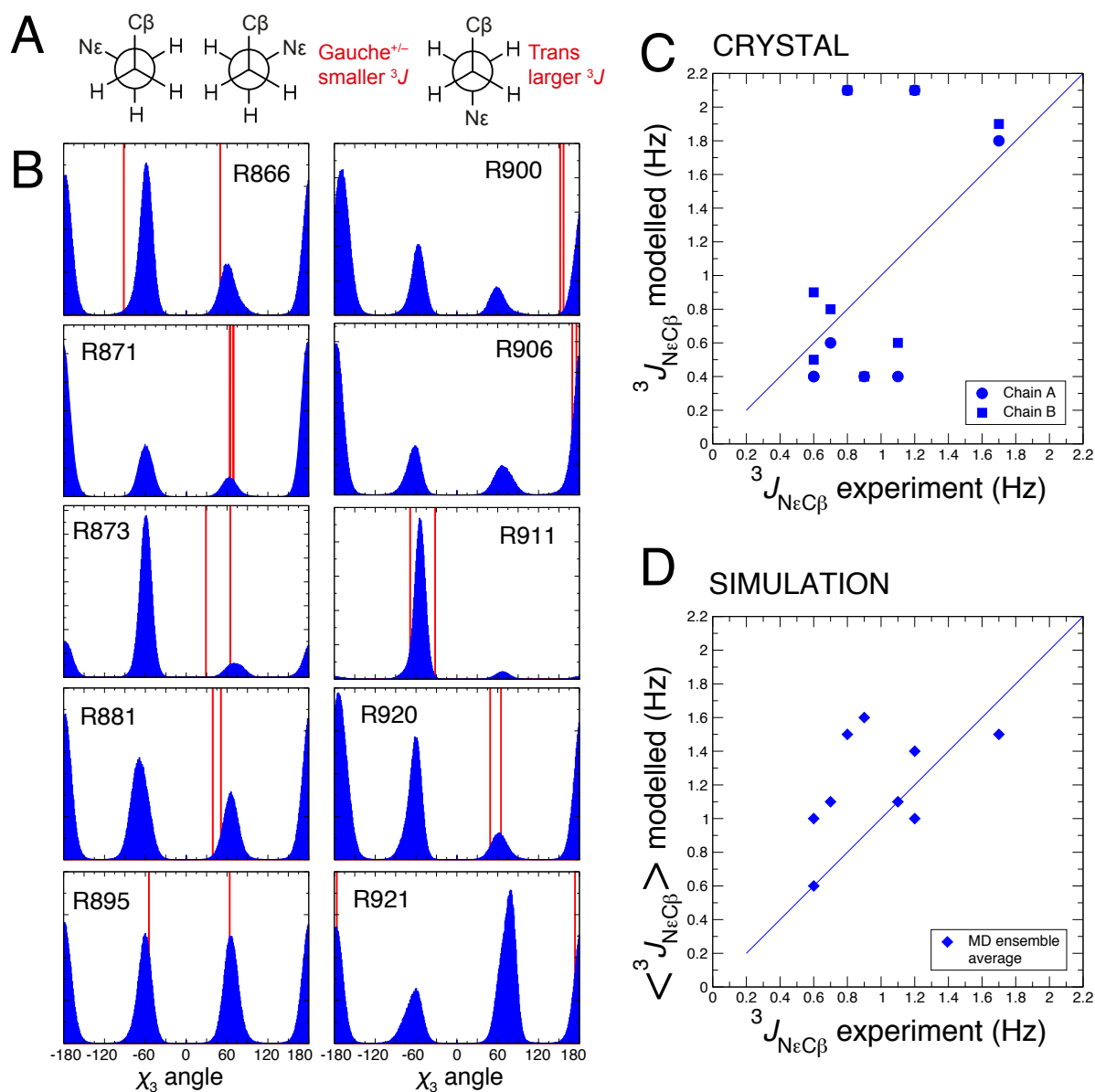
Figure S5. Comparison of modeled and experimentally determined $^3J_{NC}$ values across the Arg χ_3 dihedral. (A) Qualitative relationship between χ_3 dihedral and $^3J_{NC}$. The three rotamers for Arg χ_3: *gauche*$^+$ (60°), *gauche*$^-$ (–60°) and *trans* (~180°). *gauche* rotamers will give rise to small $^3J_{NC}$ whilst *trans* rotamers will give rise to larger $^3J_{NC}$. **(B)** Arg χ_3 angle distributions from MD simulation (blue). The two red lines mark the dihedral angles from the two chains in the crystal structure. Arg895 and Arg900 had the smallest and largest $^3J_{NC}$, respectively, in NMR measurements. For these two examples, Arg895 had an enhanced contribution from *gauche* rotamers compared to Arg900, giving rise to a smaller coupling constant. **(C)** $^3J_{NC}$ values calculated for the two chains in the crystal structure 5WST compared with experiment. The root mean square difference (RMSD) values were 0.7 and 0.6 Hz for chain A and B, respectively. **(D)** Ensemble averaged $^3J_{NC}$ values calculated from MD simulation of M7A SAH. The RMSD value was 0.4 Hz. For the crystal structure: $^3J = A \cos^2 \chi + B \cos \chi + C$, and for simulation: $<^3J> = A <\cos^2 \chi> + B <\cos \chi> + C$, where $< >$ is the ensemble average. The Karplus parameters have not been described for Arg χ_3. However, parameters used for calculating coupling constants are mostly dependent on the nuclei involved in the dihedral, so using the known N-C-C-C dihedral relation for $^3J_{NC}$ values should be a fairly reasonable estimate for Arg χ_3. The $^3J_{NC}$ Karplus parameters known for χ_1 (2) were used, which have been shown to be accurate for χ_4 in Lys (1): $A = 1.29$, $B = -0.49$, $C = 0.37$. Changing Karplus parameters to $A = 1.83$, $B = 0.09$, $C = 0.29$ (values 'optimised' by Zanderashvili *et al*. for Lys)

altered ^3J values by 0.2 Hz or less, and correlation of experiment with MD was still better than with crystal structure values.

Table S1. List of NMR experiments used for assignment

Experiment	Exp. detail	Instrument	pH	Temp (°C)	Use
HN-based					
N-HSQC	Various	Various	5.5/7.4	10–23.4	Protein fingerprint
^{1}H–^{15}N TROSY	Various	Various	5.5/7.4	10–23.4	Protein fingerprint
HNCO	BEST TROSY	750 MHz	7.4	23.4	Standard triple resonance
HNCO	BEST TROSY	950 MHz	5.5	23.4	Transfer assignments
HNCO	BEST TROSY	600 MHz	5.5	10	Transfer assignments
HN(CA)CO	BEST TROSY	750 MHz	7.4	23.4	Standard triple resonance
HN(CA)CO	BEST TROSY	600 MHz	5.5	23.4	Transfer assignments
HN(CA)CO	BEST TROSY	600 MHz	5.5	10	Transfer assignments
HN(CO)CA	BEST TROSY	750 MHz	7.4	23.4	Standard triple resonance
HNCA	BEST TROSY	750 MHz	7.4	23.4	Standard triple resonance
HN(COCA)CB	BEST	750 MHz	7.4	23.4	Standard triple resonance
HN(CA)CB	BEST TROSY	750 MHz	7.4	23.4	Standard triple resonance
NOESY-N-HSQC	Ref. (3,4)	950 MHz	7.4	23.4	Assist assignment through HN_i–HN_{i+1} links
HBHA(CO)NH	Ref. (5)	600 MHz (Varian)	7.4	23.4	Some Hβ and Hα assignments
H(CCCO)NH	Ref. (6,7)	750 MHz	7.4	23.4	Assign aliphatic side chain ^{1}H for residues with clean (i + 1) peak in ^{15}N-HSQC
(H)CC(CO)NH	Ref. (6)	750 MHz	7.4	23.4	Assign aliphatic side chain ^{13}C for residues with clean (i + 1) peak in ^{15}N-HSQC
(HACA)CO(CACO)NH	Ref. (8)	750 MHz	7.4	23.4	For specific assignment of Glu Cδ (and thereby Glu Hγ) resonances via known backbone N–H correlations at (i + 1) positions
HiSQC (Lys)	Ref. (9)	950 MHz	5.5	10	Examine Hζ–Nζ correlations
HiSQC (Arg)		950 MHz	5.5	10, 23.4	Examine Hε–Nε and

					Hη–Nη correlations
TOCSY-Nε-HSQC	Ref. (10)	950 MHz	5.5	23.4	Assign Hε–Nε correlations
TROSY H(CCCO)NH	Ref. (7)	950 MHz	5.5	23.4	Assign Hε–Nε correlations
Nε-NOESY-HSQC	Ref. (3,4)	950 MHz	5.5	10	Assign Hε–Nε correlations
HC-based					
C-HSQC	Constant time	600 MHz (Varian)	7.4	23.4	Some sp^3 ^{13}C H–C correlations assigned
C-HSQC	Constant time	950 MHz	5.5	10	Compare correlations with those at 23.4 °C, pH 7.4 (very similar in most cases)
C-HSQC (aromatics)	Constant time	600 MHz (Varian)	7.4	23.4	All eleven sp^2 ^{13}C H–C correlations assigned (for Y863, H897, H934, W at C-terminus)
H(C)CH TOCSY	Ref. (11)	950 MHz (Crick Institute, London)	7.4	23.4	Partial side chain assignment (^1H)
(H)CCH TOCSY	Ref. (11)	950 MHz (Crick Institute, London)	7.4	23.4	Partial side chain assignment (^{13}C)
NOESY-C-HSQC	Ref. (10)	950 MHz	7.4	23.4	Assist assignment
NOESY-C-HSQC (aromatics)	Ref. (10) (HH) 2D only	950 MHz	7.4	23.4	Assist assignment
HA(CA)CO	Ref. (12)	950 MHz	7.4	23.4	Some Ha assignments
HG(CG)CD	Ref. (13)	950 MHz	7.4	23.4	Glu specific, also highlights Gln HG(CG)CD and Asp HB(CB)CG correlations
NOESY-HA(CA)CO		950 MHz	7.4	23.4	Assist with some Hα assignment
NOESY-HG(CG)CD		950 MHz	7.4	23.4	Assist with Glu Hγ pair assignment, look for nearby Lys, Arg nuclei
CON	Ref. (14)	950 MHz	7.4	23.4	Check backbone assignment

Table S2. List of assigned chemical shifts for M7A SAH at pH 7.4, 23.4 °C.

		H	N	H_α	H_β	H_γ	H_δ	H_ε	C	C_α	C_β	C_γ	C_δ	C_ε	
N-ter	Ser			3.79	3.78 3.90				176.49	57.97	65.95				
858	Arg		123.73	4.31	1.86 1.86	1.64 1.64	3.21 3.21		176.87	57.03	30.44	27.14	43.33		
859	Leu	8.28	121.90	4.28	1.56 1.66	1.64	0.87 0.93		178.04	55.99	42.19	27.05	23.52 24.90		
860	Arg	8.23	121.92	4.31	1.93 1.93	1.69 1.69	3.21 3.21		177.47	57.21	30.31	27.16	43.31		
861	Val	8.15	121.32	3.93	2.11	1.02 0.95			177.45	64.58	32.25	21.06 21.51			
862	Glu	8.57	122.10	4.15	2.05 2.05	2.33 2.31			178.05	58.56	29.34	36.23	183.58		
863	Tyr	8.09	120.95	4.35	3.17 3.13		7.13	6.83	177.19	60.25	38.32		133.03	118.33	
864	Gln	8.24	119.58	3.98	2.17 2.17	2.46 2.46		7.63 6.88	178.14	58.27	28.33	33.74	179.87		$N^{\varepsilon 2}$ 111.96
865	Arg	8.24	119.90	4.13	1.92 1.92	1.68 1.68	3.21 3.21		178.85	59.03	30.23	27.40	43.46		
866	Arg	8.05	121.18	4.13	1.92 1.92	1.68 1.68	3.21 3.21		178.54	59.00	29.74	27.43	43.51		
867	Leu	8.05	121.49	4.09	1.57 1.68	1.52	0.83 0.83		179.75	57.62	41.77	26.83	23.52 24.79		
868	Glu	8.19	120.14	4.12	2.08 2.18	2.36 2.36			178.54	58.95	29.46	36.10	183.57		
869	Ala	8.10	122.23	4.19	1.55				180.75	55.09	17.96				
870	Glu	8.14	119.95	4.22	2.07 2.14	2.34 2.34			178.34	58.85	29.56	35.85	183.57		
871	Arg	8.13	120.03	4.04	1.99 1.99	1.66 1.66	3.23 3.23		179.77	59.63	29.89	27.31	43.40		
872	Met	8.20	119.06	4.26	2.19 2.19	2.76 2.60		2.11	178.26	58.41	32.49	32.27		16.96	
873	Arg	8.07	121.93	4.20			3.25 3.25		179.02	59.21	29.92	27.03	43.10		
874	Leu	8.22	120.09	4.10	1.84 1.59	1.78	0.91 0.93		179.69	57.78	41.70	26.95	23.39 25.05		
875	Ala	7.99	122.11	4.23	1.56				180.93	54.94	18.11				
876	Glu	8.22	120.43	4.12	2.20 2.20	2.36 2.39			178.91	59.22	29.47	36.24	183.42		
877	Glu	8.34	120.41	4.01	2.17 2.17	2.45 2.25			179.16	59.75	29.48	36.53	183.08		
878	Glu	8.12	119.80	4.10	2.13 2.13	2.34 2.34			178.41	59.23	29.41	36.01	183.49		
879	Lys	7.86	120.63	4.04	1.99 1.99	1.39 1.59	1.72 1.72	2.98 2.98	179.32	59.74	32.61	24.89	29.67	42.06	
880	Leu	8.11	119.97	4.18	1.85 1.69	1.81	0.93 0.97		179.50	57.74	41.76	26.98	24.94 23.68		
881	Arg	8.12	119.91						179.30	59.45					
882	Lys			4.12					179.22	59.30	32.57	25.11	29.13	42.14	
883	Glu	8.21	120.13	4.22	2.14 2.14	2.27 2.45			178.79	58.68	30.05	36.36	183.16		
884	Met	8.43	119.52	4.41	2.21 2.22	2.71 2.72		2.12	178.20	57.57	32.57	31.86		17.00	
885	Ser	8.33	116.02	4.30	4.10 4.02				176.43	60.83	63.05				
886	Ala	8.22	125.21	4.22	1.53				180.30	55.05	18.22				
887	Lys	8.12	120.22		1.95 1.95				178.74	59.55	32.76	25.08		42.11	
888	Lys	7.99	120.07	4.21	1.94 1.94	1.51 1.51	1.69 1.69	2.96 2.96	178.75	58.82	32.26	24.97	29.04	42.39	
889	Ala	8.19	121.73	4.23	1.52				180.59	54.98	18.21				
890	Lys	7.97	120.36					3.02 3.02	178.86	59.39		24.86		42.06	
891	Glu	8.15	120.06	4.12	2.19 2.19	2.54 2.31			179.71	59.32	29.61	36.55	183.45		
892	Glu	8.45	120.59	4.20	2.02 2.17	2.35 2.35			178.51	59.00	29.55		183.51		
893	Ala	8.12	122.20	4.16	1.58				181.03	55.23	17.89				
894	Glu	8.24	119.78	4.21	2.17 2.17	2.47 2.37			178.76	59.11	29.14	36.18	183.29		
895	Arg	8.14	121.31	4.12			3.29 3.29		179.36	59.58	29.91	27.30	43.32		
896	Lys	8.24	119.14	4.09	1.97 1.97	1.67 1.67		2.97 2.97	179.34	59.20	32.34	25.45	28.80	42.13	
897	His	8.16	120.87	4.49	3.19 3.40		7.06	7.75	177.31	59.51	30.86		119.37	138.46	

11

	H	N	H_α	H_β	H_γ	H_δ	H_ε	C	C_α	C_β	C_γ	C_δ	C_ε	
898 Gln	8.34	117.83	3.89	2.24 2.23	2.61 2.61		6.87 7.76	179.53	58.76	27.84	33.72	179.96		N^ε 112.58
899 Glu	8.32	120.62	4.12					179.12	59.26	29.53	36.39	183.51		
900 Arg	8.04	121.89	4.13			3.29 3.29		178.91	59.06	29.59	27.33	43.26		
901 Leu	8.19	119.48	3.98	1.78 1.44	1.47	0.74 0.78		180.14	57.91	41.34	26.54	22.55 25.12		
902 Ala	8.02	122.26	4.21	1.53				180.39	54.82	17.80				
903 Gln	8.01	120.19	4.08	2.25 2.25	2.41 2.48		6.86 7.86	178.32	58.60	27.86	33.10	180.12		N^ε 112.67
904 Leu	8.22	120.70	4.10	1.61 1.84	1.78	0.90 0.90		179.12	57.70	41.71	27.00	23.21 25.32		
905 Ala	7.98	121.14	4.27	1.53				180.79	54.76	17.94				
906 Arg	7.93	120.34			1.63 1.63	3.26 3.26		178.80	59.21	30.02	27.56	43.53		
907 Glu	8.32	120.55	4.09	2.12 2.12	2.29 2.55			179.57	59.31	29.48	36.65	183.39		
908 Asp	8.45	121.05	4.50	2.81 2.72				178.15	57.20	40.74	179.23			
909 Ala	8.00	122.83	4.22	1.56				180.76	54.99	18.10				
910 Glu	8.22	118.95	4.09	2.13 2.13	2.30 2.49			178.88	59.13	29.59	36.47	183.41		
911 Arg	8.06	121.35	4.09	2.03 2.03		3.27 3.27		178.63	59.60	29.77	27.27	43.18		
912 Glu	8.19	119.57	4.10	2.10 2.10	2.42 2.23			178.94	59.44	29.55	36.45	183.58		
913 Leu	7.88	120.11	4.14	1.74 1.86	1.76	0.93 0.97		179.61	58.06	41.80	26.81	24.76 24.09		
914 Lys	8.03	120.34	4.13	2.00 2.00	1.55 1.55	1.70 1.70	3.02 3.02	179.50	59.37	32.26	25.04	29.28	42.19	
915 Glu	8.37	119.98	4.11	2.12 2.12	2.50 2.29			179.88	59.43	29.46	36.62	183.47		
916 Lys	8.24	121.10	4.15	1.99 1.99	1.51 1.70	1.71 1.71	2.98 2.98	179.65	59.49	32.47	25.58	29.36	42.25	
917 Glu	8.23	121.44	4.24	2.03 2.20	2.37 2.37			178.88	58.96	29.20	36.06	183.68		
918 Glu	8.33	120.58	4.18	2.12 2.12	2.36 2.36			178.91	59.15	29.37	36.23	183.50		
919 Ala	8.06	121.69	4.16	1.54				180.60	55.02	18.02				
920 Arg	7.98	120.26	4.13	1.99 1.99	1.71 1.71	3.25 3.25		178.55	59.29	29.83	26.99	43.13		
921 Arg	8.21	119.98	4.09	1.97 1.97	1.74 1.74	3.23 3.23		178.97	59.10	30.07	27.45	43.27		
922 Lys	8.09	119.21	4.04	1.91 1.91	1.64 1.45	1.69 1.69	2.99 2.99	178.57	59.47	32.52	25.38	29.29	42.13	
923 Lys	7.83	120.28	4.09	1.99 1.99	1.43 1.56	1.70 1.70	2.98 2.98	178.61	59.23	32.60	24.86	29.65	42.23	
924 Glu	8.21	119.40	4.12	2.12 2.12	2.26 2.47			178.92	59.04	29.56	36.51	183.65		
925 Leu	7.97	120.51	4.18	1.82 1.70	1.74	0.92 0.94		179.32	57.56	41.95	27.00	24.73 23.93		
926 Leu	7.92	120.32	4.17	1.70 1.86	1.77	0.92 0.94		179.60	57.67	41.78	26.90	24.96 23.65		
927 Glu	8.13	119.34	4.13	2.11 2.11	2.26 2.48			178.66	58.70	29.60	36.65	183.59		
928 Gln	8.06	118.92	4.15	2.20 2.20	2.53 2.46		6.83 7.53	178.02	57.74	28.61	33.84	180.28		N^ε 111.84
929 Met	8.13	119.49	4.29	2.18 2.18	2.60 2.75		2.11	177.78	57.46	32.77	32.11		17.06	
930 Glu	8.15	120.29	4.17	2.09 2.09	2.36 2.29			177.73	57.92	29.82	36.25	183.71		
931 Lys	7.97	120.19	4.18	1.83 1.88	1.45 1.54	1.67 1.67	2.98 2.98	177.23	57.39	32.76	25.04	29.21	42.14	
932 Ala	7.94	122.67	4.26	1.40				177.95	52.87	18.94				
933 Arg	7.88	118.82	4.09	1.70 1.72	1.47 1.56	3.10 3.10		176.19	56.54	30.57	27.06	43.37		
934 His	8.02	119.41	4.55	2.92 2.95		6.88	7.85	174.83	55.98	30.61		120.18	138.02	
935 Glu	8.01	121.65	4.26	1.98 1.81	2.12 2.09			175.00	56.30	30.43	36.25	184.13		
C-ter Trp	7.65	126.89	4.50	3.16 3.34		7.18	10.01 7.66	180.97	58.46	30.14		126.92	121.21	$N^{\varepsilon 1}$ 128.58, $H^{\zeta 2}$ 7.42, $H^{\zeta 3}$ 7.13, $H^{\eta 2}$ 7.18, $C^{\zeta 2}$ 114.47, $C^{\zeta 3}$ 121.93, $C^{\eta 2}$ 124.46

Table S3. Output file from δ2D (determination of secondary structure populations from chemical shifts)

```
#SQ:  SRLRVEYQRRLEAERMRLAEEEKLRKEMSAKKAKEEAERKHQERLAQLAREDAERELKEKEEARRKKELLEQMEKARHEW
#SS:  CCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCC
```

#Total Populations:
#Helix(H): 84.3%
#Extended-Beta(E): 0.1%
#Polyproline II (PPII)(P): 1.9%
#Coil(C): 13.8%

#Populations per residue (residues marked with a * are less reliable):

#num	res	Helix	Beta	Coil	PPII	SS
#1	S					
2	R	0.137	0.002	0.646	0.214	C
3	L	0.186	0.002	0.674	0.139	C
4	R	0.287	0.001	0.631	0.082	C
5	V	0.502	0.001	0.459	0.038	H
6	E	0.719	0.001	0.263	0.017	H
7	Y	0.850	0.001	0.143	0.006	H
8	Q	0.916	0.001	0.081	0.002	H
9	R	0.948	0.000	0.051	0.001	H
10	R	0.953	0.000	0.045	0.001	H
11	L	0.939	0.000	0.059	0.002	H
12	E	0.901	0.000	0.095	0.004	H
13	A	0.873	0.000	0.121	0.006	H
14	E	0.850	0.000	0.143	0.008	H
15	R	0.906	0.000	0.090	0.005	H
16	M	0.950	0.000	0.048	0.002	H
17	R	0.980	0.000	0.019	0.001	H
18	L	0.980	0.000	0.019	0.001	H
19	A	0.972	0.000	0.027	0.001	H
20	E	0.954	0.000	0.044	0.002	H
21	E	0.951	0.000	0.047	0.001	H
22	E	0.950	0.000	0.049	0.001	H
23	K	0.970	0.000	0.030	0.000	H
24	L	0.981	0.000	0.019	0.000	H
25	R	0.980	0.000	0.020	0.000	H
26	K	0.954	0.000	0.046	0.000	H
27	E	0.920	0.000	0.080	0.001	H
28	M	0.909	0.000	0.090	0.001	H
29	S	0.922	0.000	0.077	0.001	H
30	A	0.931	0.000	0.068	0.001	H
31	K	0.936	0.000	0.062	0.002	H
32	K	0.940	0.000	0.058	0.002	H
33	A	0.960	0.000	0.039	0.001	H
34	K	0.961	0.000	0.038	0.001	H
35	E	0.945	0.000	0.053	0.002	H
36	E	0.907	0.000	0.089	0.003	H
37	A	0.902	0.000	0.095	0.003	H
38	E	0.897	0.000	0.100	0.003	H
39	R	0.938	0.000	0.060	0.002	H
40	K	0.965	0.000	0.034	0.001	H
41	H	0.971	0.000	0.027	0.001	H
42	Q	0.965	0.000	0.033	0.001	H
43	E	0.958	0.000	0.041	0.001	H
44	R	0.971	0.000	0.028	0.001	H
45	L	0.984	0.000	0.016	0.000	H
46	A	0.989	0.000	0.011	0.000	H
47	Q	0.986	0.000	0.014	0.000	H
48	L	0.980	0.000	0.020	0.000	H
49	A	0.975	0.000	0.025	0.000	H
50	R	0.964	0.000	0.035	0.001	H
51	E	0.961	0.000	0.038	0.001	H
52	D	0.939	0.000	0.059	0.002	H
53	A	0.917	0.000	0.080	0.003	H
54	E	0.903	0.000	0.094	0.003	H
55	R	0.941	0.000	0.057	0.002	H
56	E	0.973	0.000	0.026	0.001	H
57	L	0.987	0.000	0.013	0.000	H
58	K	0.984	0.000	0.015	0.000	H
59	E	0.976	0.000	0.024	0.001	H
60	K	0.962	0.000	0.036	0.002	H
61	E	0.947	0.000	0.050	0.003	H
62	E	0.952	0.000	0.046	0.002	H
63	A	0.965	0.000	0.034	0.001	H
64	R	0.964	0.000	0.035	0.001	H
65	R	0.954	0.000	0.046	0.000	H
66	K	0.941	0.000	0.059	0.000	H
67	K	0.946	0.000	0.053	0.001	H
68	E	0.955	0.000	0.044	0.001	H
69	L	0.946	0.000	0.053	0.001	H
70	L	0.929	0.000	0.070	0.001	H
71	E	0.893	0.000	0.104	0.003	H
72	Q	0.827	0.000	0.164	0.009	H
73	M	0.669	0.000	0.303	0.028	H
74	E	0.450	0.001	0.483	0.067	C
75	K	0.272	0.001	0.618	0.109	C
76	A	0.138	0.002	0.719	0.141	C
77	R	0.065	0.006	0.778	0.152	C
78	H	0.026	0.012	0.795	0.167	C
79	E	0.011	0.019	0.790	0.180	C
#80	W					

Table S4. List of Arg–Glu and Lys–Glu ion pair occupancies from simulation of M7A SAH.

Arg	Glu	Occupancy (%)		Lys	Glu	Occupancy (%)
858	862	49.5		879	876	32.2
865	868	90.7		879	883	20.5
865	862	2.0		882	878	22.8
866	870	34.9		887	883	7.0
866	862	5.0		887	891	21.9
871	868	2.1		888	891	65.3
873	870	52.7		888	892	4.1
873	876	24.4		890	894	21.9
873	877	60.3		896	892	9.2
881	877	19.3		896	899	43.6
881	878	57.3		914	910	3.0
895	891	9.1		914	917	17.4
895	892	60.9		914	918	15.6
895	899	34.6		916	912	10.8
906	910	41.8		922	918	4.7
911	907	0.8		923	927	18.7
911	915	84.5		931	927	8.3
920	917	61.6		931	935	2.4
920	924	31.8				
921	917	30.4				
921	918	46.2				
921	924	48.1				
933	930	6.4				

Movie S1. A full simulation trajectory (500 ns) centered on Arg873 to accompany Fig. 6. Movie frames were recorded every 1 ns. The helix is oriented with the N-terminus to the top of the image. Arg873, Glu870, Glu876 and Glu877 are shown in space-fill, the nitrogen atoms of Arg873 are blue and oxygen atoms of the Glu residues are red.

References

1. Zandarashvili, L., Li, D. W., Wang, T., Bruschweiler, R., and Iwahara, J. (2011) Signature of mobile hydrogen bonding of lysine side chains from long-range 15N-13C scalar J-couplings and computation. *J. Am. Chem. Soc.* **133**, 9192-9195

2. Perez, C., Lohr, F., Ruterjans, H., and Schmidt, J. M. (2001) Self-consistent Karplus parametrization of 3J couplings depending on the polypeptide side-chain torsion chi1. *J. Am. Chem. Soc.* **123**, 7081-7093

3. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665

4. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J. Am. Chem. Soc.* **111**, 1515-1517

5. Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J. Biomol. NMR* **3**, 185-204

6. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in 13C/15N-Enriched Proteins by Isotropic Mixing of 13C Magnetization. *J. Magn. Reson., Ser. B* **101**, 114-119

7. Montelione, G. T., Lyons, B. A., Emerson, S. D., and Tashiro, M. (1992) An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *J. Am. Chem. Soc.* **114**, 10974-10975

8. Tollinger, M., Forman-Kay, J. D., and Kay, L. E. (2002) Measurement of side-chain carboxyl pK(a) values of glutamate and aspartate residues in an unfolded protein by multinuclear NMR spectroscopy. *J. Am. Chem. Soc.* **124**, 5714-5717

9. Iwahara, J., Jung, Y. S., and Clore, G. M. (2007) Heteronuclear NMR spectroscopy for lysine NH(3) groups in proteins: unique effect of water exchange on (15)N transverse relaxation. *J. Am. Chem. Soc.* **129**, 2971-2980

10. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1.beta. *Biochemistry* **28**, 6150-6156

11. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) 1H–1H correlation via isotropic mixing of 13C magnetization, a new three-dimensional approach for assigning 1H and 13C spectra of 13C-enriched proteins. *J. Magn. Reson. (1969)* **88**, 425-431

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

13. Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y., and Nakamura, H. (1994) Individual ionization constants of all the carboxyl groups in ribonuclease HI from Escherichia coli determined by NMR. *Biochemistry* **33**, 5275-5284

14. Kostic, M., Pochapsky, S. S., and Pochapsky, T. C. (2002) Rapid Recycle 13C',15N and 13C,13C' Heteronuclear and Homonuclear Multiple Quantum Coherence Detection for Resonance Assignments in Paramagnetic Proteins: Example of Ni2+-Containing Acireductone Dioxygenase. *J. Am. Chem. Soc.* **124**, 9054-9055