Article:
Batchelor, M orcid.org/0000-0001-6338-5698, Wolny, M, Baker, EG et al. (3 more authors) (2019) Dynamic ion pair behavior stabilizes single alpha helices in proteins. Journal of Biological Chemistry, 294 (9). pp. 3219-3234. ISSN 0021-9258

https://doi.org/10.1074/jbc.ra118.006752

JBC Papers in Press. Published on December 28, 2018 as Manuscript RA118.006752
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.RA118.006752

Dynamic ion pair behavior stabilizes single alpha helices in proteins.

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

Running Title: *Ion pair dynamics in helices*

*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

Keywords: alpha helix, molecular dynamics, nuclear magnetic resonance, salt bridges

ABSTRACT

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 aspartate (Asp) or glutamate (Glu) and positively charged residue side chains in arginine (Arg), histidine (His) or lysine (Lys). These interactions are thought to be important in stabilizing protein structure, and often occur at catalytically active sites. In structural models they are commonly depicted as fixed, despite the marginal energetic benefit they provide compared to 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). 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 non-cooperative 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 provide 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 to 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 nuclear magnetic resonance (NMR) spectroscopy and molecular dynamics simulations. We selected the M7A SAH as its less repetitive sequence compared to that of other known SAHs was expected to facilitate 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, while 35 Glu–Lys and Glu–Arg helix-compatible ion pairs can potentially form, only 8 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 nuclear magnetic resonance (NMR) spectroscopy. We also performed molecular dynamics (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 circular dichroism (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 non-cooperatively, and remained highly helical over a wide range of pH and salt concentrations (Fig. S1A–D). It was monomeric and had an elongated structure in solution as demonstrated by analytical ultracentrifugation and size exclusion chromatography, respectively (Fig. S1E, F). Additional data from small angle X-ray scattering (SAXS) was again consistent with a continuous (extended) helix structure for M7A SAH (Fig. S1G). The Kratky plot of the scattering data ($(q^2I(q)$ as a function of q, where the momentum transfer $q = 4\pi \sin\theta / \lambda$, 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 was best fit ($\chi^2 < 0.5$, Fig. S2) by near-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 2D ^{1}H–^{15}N TROSY (transverse relaxation optimized spectroscopy) spectrum of M7A SAH showed a remarkable lack of dispersion (Fig. 1A–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 ($^{1}H_N$, $^{15}N_H$, $^{13}C\alpha$, $^{13}C'$ 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. 1C) (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 later) and yet the protein is fully helical throughout. A fuller calculation of secondary structure populations from all 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 $^{1}H_{N}$, $^{1}H\alpha$, $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$ and ^{15}N nuclei (25,26), showed that the lowest energy structures all exhibit extended helix conformations (Fig. 1D).

Next, we used heteronuclear NOE and relaxation experiments to interrogate the local and global dynamics of M7A SAH. ^{1}H–^{15}N heteronuclear NOE values were high and positive for residues 866 to 926 (Fig. 2A) demonstrating that these backbone amides (away from the termini) are ordered on the ps to ns 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 to those from the more ordered central residues in the helix (Fig. 2B–C). Longitudinal relaxation rates (R_1) of ^{15}N nuclei are sensitive to the fast (ps–ns 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 to 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 the backbone. Many side chain nuclei resonances, especially those of 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 eleven residues (22 protons). This effectively puts a NOE-based NMR structure determination, using ARIA (30-32) or similar, 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 Methods).

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 (Glu862 and Glu935) exhibited very small resolvable differences in chemical shift. This result indicates that side chains of the central Glu residues in M7A SAH are not 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 is not affected (Fig. S1B,C).

At pH 5.5 and temperature of 10 °C, the Lys Hζ–Nζ correlations in M7A SAH (from eleven Lys residues) are visible, albeit as a single peak that is only resolved into two peaks at high resolution (Fig. 4A). Poorly dispersed Hη–Nη correlations for the fourteen Arg residues were observed with Hη shifts only ranging from 6.7 to 7.1 ppm (Fig. 4B). 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. 4C). 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 $^{3}J_{N\varepsilon C\beta}$ couplings, and long-range inter-residue $^{3}J_{N\varepsilon C'}$ 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 Methods and Fig. S3).

The Arg Hε–Nε correlations fall into distinct groups that indicate their potential involvement in ion pairs (Fig. 4C). 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, Arg900, 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 previously seen 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. 4B). 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. 4D, Table 1) consistent with the side chains being relatively ordered compared to 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. While, 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 broadly similar pattern to that of the heteronuclear NOEs with higher values for the central residues compared to those at the termini (Table 1). This again is an indicator that central Arg side chains are more constrained than those at the termini (43). R_2 values are particularly high for Arg873, Arg920 and Arg895 (those with the most downshifted Hε) indicating that slower (>ns) 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 could we 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 Methods, Fig. S4).

First, measurements of small $^3J_{NC}$ couplings *within* each Arg residue (*i.e.* $^3J_{N\varepsilon C\beta}$) were used to validate the spin-echo difference experiment and report on the Arg side chain conformational ensemble (42). The experiment could readily measure $^3J_{N\varepsilon C\beta}$ 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 Arg895, to 1.7 Hz for Arg900 (Table 1). The $^3J_{N\varepsilon C\beta}$ 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 $^3J_{N\varepsilon C\beta}$ values are indicative of larger *trans* χ_3 contributions and smaller $^3J_{N\varepsilon C\beta}$ values indicate larger *gauche* contributions (see Fig. S5A for a qualitative description). Arg900, which has no potential ion pair Glu neighbors, spends more time in the less sterically demanding *trans* conformation. Arg895, 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 $^3J_{N\varepsilon C\beta}$ values point towards 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 ($^3J_{N\varepsilon C'}$) to Glu Cδ. As there was no drop in peak intensity between the component spectra of the experiment, values of $^3J_{N\varepsilon C'}$ 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 sub-spectra for Arg873 and Arg920, due to their fast relaxation properties, estimation of $^3J_{N\varepsilon C'}$ 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, while the very small size or absence of through-hydrogen-bond couplings argues against specific salt bridges (50). These data either indicate 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 while 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 observed during the simulation, albeit with a wide range of occupancies (1–91%, Table S4). For example, the Glu892–Arg895 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 Glu892 (Glu892–Lys888 and Glu892–Lys896) were also observed albeit for a lower fraction of the overall simulation time (9% and 4%, respectively). The pairing Arg895–Glu899 (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, while 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 Arg873–Glu870 and Arg873–Glu877 were formed simultaneously for 45% of the simulation, whereas Arg873–Glu876 and Arg873–Glu877 were simultaneously occupied for 8% (see snapshot images in Fig. 6B–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: 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 N atoms (Lys or Arg) and a side chain O 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 Lys882–Glu878. On the other hand, the ion pair reported to form between Glu907 and Arg911 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 Arg895–Glu892 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 Glu878 in chain B, displaying two side chain conformers that form an ion pair to either Arg881 or Lys882. In chain A, Glu878 occupies a single position where it forms an ion pair with both Arg881 and Lys882 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. 7B, C), and yet the number of observed ion pairs was still small given the potential number of ion pairs that potentially could exist.

We estimated $^3J_{N\varepsilon C\beta}$ for Arg residues using dihedral angles measured in the crystal structure (Fig. S5B, red lines), and compared them to experimental $^3J_{N\varepsilon C\beta}$ 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 $^3J_{N\varepsilon C\beta}$ values were estimated (Fig. S5D). The correlation of simulation with experimental $^3J_{N\varepsilon C\beta}$ 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 (52). Our combination of experimental data and simulations support 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 non-cooperative 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 shows 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 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 $^{3}J_{N\varepsilon C\beta}$ values, that most of the central Arg residues are forming ion pairs. The one exception, Arg900, is not expected to form an ion pair. Inspection of the helical net plot (Fig. 7A) shows it has no potential ion pair partners compatible with the helix structure (Fig. 7A). Correspondingly, it appears in the same part of the HiSQC spectrum as the terminal Arg residues (Fig. 4C) and has a high $^{3}J_{N\varepsilon C\beta}$ 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).

While 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. $^{3}J_{NC'}$ values of 0.23 and 0.17 Hz were measured for mobile H-bonds in ubiquitin (42). Their simulated average lifetimes were 45 ps and 30 ps, and H-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 over-stabilize salt bridges compared to 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. 7A), 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. 7A) 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 Ni-NTA 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 to enzyme ratio 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: 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_2HPO_4/NaH_2PO_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 (TB) medium (4 x 400 ml). After reaching OD_{600} ~ 0.6 the cultures were spun down, combined pellets were washed with 100–200 ml of minimal media (M9 salts, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 10 ml of basal vitamins) and resuspended in 250 ml of minimal media. After 45 min in 37 ℃, 220 RPM the culture was supplemented with isotopes (100 mM $^{15}NH_4Cl$, 50 mM ^{13}C D-glucose, Goss Scientific) and grown in the same conditions for additional 35 min after which culture was induced with 0.5 mM IPTG and grown overnight at 21 ℃, 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 1 kDa mwco) overnight against 50 mM ammonium acetate, pH 7.4 and analyzed by TOF MS analysis (The 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 in the *Protein expression and purification* section. 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 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 previously described (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), ribonuclease 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 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/) (60), and 99% confidence limits 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 with small-angle X-ray scattering

(SEC-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 m from the sample. Data were analyzed using ScÅtter (v3.0a) (Robert P. Rambo, http://www.bioisis.net/welcome). The Kratky plot of the scattering data is $q^2I(q)$ *versus* q, while 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.

Nuclear magnetic resonance spectroscopy

NMR spectra were recorded on 600 MHz and 750 MHz Bruker Avance spectrometers or a 950 MHz Bruker Aeon Ascend 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 is given (Table S1). Spectra were processed using NMRPipe/NMRDraw (61) and then peak assignments 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 (out 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 a 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 correlations 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 °C and 10 °C, pH 5.5), and supplemented by following peak movements in a set of TROSY spectra recorded at ~3 °C intervals (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 residue 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 ^1H_N, $^1H\alpha$, $^{13}C\alpha$, $^{13}C\beta$, $^{13}C'$ 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 RMSDs >2 Å away from the lowest energy structure all have noticeably higher energies.

The pulseprogram 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 Lys Nε. Dephasing, resulting from $^1J_{N\varepsilon C\delta}$ and $^1J_{N\varepsilon C\zeta}$, was refocused by means of two pairs of shaped pulses. Carrier positions were: ^{1}H, the position of the water resonance; Arg ^{15}Nε, 84 ppm; Arg $^{13}C\delta$, 42 or 41 ppm; Arg $^{13}C\beta$, 28 ppm; Arg $^{13}C\zeta$, 158 ppm and Glu $^{13}C\delta$ (C'), 181 ppm. IBURP-2 180° shaped pulses were used (pulselengths for 950 MHz): Arg $^{13}C\gamma$ (2.37 ms), Arg $^{13}C\zeta$ (2.37 ms), Arg $^{13}C\beta$ (2.37 ms) and Glu $^{13}C\delta$ (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 $J = \left(\frac{1}{\pi T}\right)\arccos\left(\frac{I_{att}}{I_{ref}}\right)$ 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 $^3J_{N\varepsilon C\beta}$ experiment and 424 ms in the $^3J_{N\varepsilon C'}$ 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 timestep 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 added and simulations 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 less than 4 Å. VMD was used to produce snapshot images (75).

Acknowledgements
Chi Trinh and Sam Hickman (Leeds) collected SAXS data. Thanks to Derek N. Woolfson (Bristol) for facilitating AUC measurements and to Peter J. Knight (Leeds) for useful discussions on the paper. This work was undertaken on ARC2, part of the High Performance Computing facilities at the University of Leeds, UK. MB and MW were supported by BBSRC grant (BB/M009114/1 to MP and APK). The Wellcome Trust (WT094232) funded the CD spectropolarimeter. EGB was supported by a BBSRC/ERASynBio grant (BB/M005615/1). For access to the 950 MHz spectrometers we acknowledge the Astbury BioStructure Laboratory BioNMR Facility, which was funded by the University of Leeds, and the Francis Crick Institute through provision of access to the MRC Biomedical NMR Centre. The Francis Crick Institute receives its core funding from Cancer Research UK (FC001029), the UK Medical Research Council (FC001029) and the Wellcome Trust (FC001029).

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Swanson, C. J., and Sivaramakrishnan, S. (2014) Harnessing the Unique Structural Properties of Isolated alpha-Helices. *The Journal of biological chemistry* **289**, 25460-25467
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. *The Journal of biological chemistry* **280**, 34702-34708
3. Peckham, M., and Knight, P. J. (2009) When a predicted coiled 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 alpha-helical protein motif. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 13356-13361
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 alpha-helix domains provides novel insight into their stability. *Scientific reports* **7**, 44341
6. Wolny, M., Batchelor, M., Knight, P. J., Paci, E., Dougan, L., and Peckham, M. (2014) Stable Single alpha-Helices Are Constant Force Springs in Proteins. *The Journal of biological chemistry* **289**, 27825-27835
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 alpha-helices. *Nat Chem Biol* **11**, 221-228
8. Batchelor, M., Wolny, M., Dougan, L., Paci, E., Knight, P. J., and Peckham, M. (2015) Myosin tails and single alpha-helical domains. *Biochem Soc Trans* **43**, 58-63
9. Suveges, D., Gaspari, Z., Toth, G., and Nyitray, L. (2009) Charged single alpha-helix: a versatile protein structural motif. *Proteins* **74**, 905-916
10. Gaspari, Z., Suveges, D., Perczel, A., Nyitray, L., and Toth, G. (2012) Charged single alpha-helices in proteomes revealed by a consensus prediction approach. *Biochimica et biophysica acta* **1824**, 637-646
11. Simm, D., Hatje, K., and Kollmar, M. (2015) Waggawagga: comparative visualization of coiled-coil predictions and detection of stable single alpha-helices (SAH domains). *Bioinformatics* **31**, 767-769
12. Simm, D., Hatje, K., and Kollmar, M. (2017) Distribution and evolution of stable single alpha-helices (SAH domains) in myosin motor proteins. *PLoS One* **12**, e0174639
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 alpha Helix Lever Arm Extension. *Structure* **25**, 579-591 e574
14. Lu, Q., Ye, F., Wei, Z., Wen, Z., and Zhang, M. (2012) Antiparallel coiled-coil-mediated dimerization of myosin X. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 17388-17393
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
16. Spink, B. J., Sivaramakrishnan, S., Lipfert, J., Doniach, S., and Spudich, J. A. (2008) Long single alpha-helical tail domains bridge the gap between structure and function of myosin VI. *Nature structural & molecular biology* **15**, 591-597
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. *Quarterly Reviews of Biophysics* **40**, 191-285
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. *Journal of Applied Crystallography* **28**, 768-773

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. *Journal of Biomolecular NMR* **55**, 311-321
20. Lescop, E., Schanda, P., and Brutscher, B. (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *Journal of Magnetic Resonance* **187**, 163-169
21. Favier, A., and Brutscher, B. (2011) Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J Biomol NMR* **49**, 9-15
22. Jie, J., Lohr, 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
23. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of molecular biology* **222**, 311-333
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
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. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4685-4690
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
27. Worrall, J. A., Liu, Y., Crowley, P. B., Nocek, J. M., Hoffman, B. M., and Ubbink, M. (2002) Myoglobin and cytochrome b5: a nuclear magnetic resonance study of a highly dynamic protein complex. *Biochemistry* **41**, 11721-11730
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
29. Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Rance, M., and Skelton, N. J. (2007) *Protein NMR spectroscopy: principles and practice*, 2nd ed., Academic Press, Amsterdam ; Boston
30. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2003) ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315-316
31. Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods in enzymology* **339**, 71-90
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 beta-spectrin. *Journal of molecular biology* **269**, 408-422
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
34. Bermel, W., Bruix, M., Felli, I. C., Kumar, M. V. V., Pierattelli, R., and Serrano, S. (2013) Improving the chemical shift dispersion of multidimensional NMR spectra of intrinsically disordered proteins. *J Biomol NMR* **55**, 231-237
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 pK a measurements. *J Biomol NMR* **60**, 109-129
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
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*

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
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
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
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
42. 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
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
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
45. Wagner, G., Pardi, A., and Wuethrich, K. (1983) Hydrogen bond length and proton NMR chemical shifts in proteins. *Journal of the American Chemical Society* **105**, 5948-5949
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 13C/15N-labeled proteins. *J Biomol NMR* **17**, 305-310
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
48. Cordier, F., and Grzesiek, S. (1999) Direct Observation of Hydrogen Bonds in Proteins by Interresidue 3hJNC‘ Scalar Couplings. *Journal of the American Chemical Society* **121**, 1601-1602
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
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
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 science : a publication of the Protein Society* **19**, 1420-1431
52. Ghosh, T., Garde, S., and Garcia, A. E. (2003) Role of backbone hydration and salt-bridge formation in stability of alpha-helix in solution. *Biophysical journal* **85**, 3187-3193
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
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 (Camb)* **49**, 9824-9826
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
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. *The journal of physical chemistry. B* **118**, 6561-6569
57. Piana, S., Lindorff-Larsen, K., and Shaw, D. E. (2011) How robust are protein folding simulations with respect to force field parameterization? *Biophysical journal* **100**, L47-49

58. Ulrich, A. K., Seeger, M., Schutze, T., Bartlick, N., and Wahl, M. C. (2016) Scaffolding in the Spliceosome via Single alpha Helices. *Structure* **24**, 1972-1983
59. Greenfield, N., and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8**, 4108-4116
60. Gorbet, G., Devlin, T., Hernandez Uribe, Blanca I., Demeler, Aysha K., Lindsey, Zachary L., Ganji, S., Breton, S., Weise-Cross, L., Lafer, Eileen M., Brookes, Emre H., and Demeler, B. (2014) A Parametrically Constrained Optimization Method for Fitting Sedimentation Velocity Experiments. *Biophysical journal* **106**, 1741-1750
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. *Journal of Biomolecular NMR* **6**, 277-293
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
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. *Journal of the American Chemical Society* **111**, 1515-1517
64. Kay, L., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society* **114**, 10663-10665
65. 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. *Journal of Magnetic Resonance, Series B* **101**, 114-119
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. *Journal of the American Chemical Society* **114**, 10974-10975
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
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
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
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
72. Ahlner, A., Carlsson, M., Jonsson, B. H., and Lundstrom, P. (2013) PINT: a software for integration of peak volumes and extraction of relaxation rates. *J Biomol NMR* **56**, 191-202
73. Niklasson, M., Otten, R., Ahlner, A., Andresen, C., Schlagnitweit, J., Petzold, K., and Lundstrom, P. (2017) Comprehensive analysis of NMR data using advanced line shape fitting. *J Biomol NMR* **69**, 93-99
74. Brooks, B. R., Brooks, C. L., III, 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., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry* **30**, 1545-1614
75. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33-38, 27-38

76. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802
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
78. Carter, P., Andersen, C. A., and Rost, B. (2003) DSSPcont: Continuous secondary structure assignments for proteins. *Nucleic acids research* **31**, 3293-3295
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 phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* **8**, 3257-3273
80. Helmus, J. J., and Jaroniec, C. P. (2013) Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *Journal of Biomolecular NMR* **55**, 355-367

Abbreviations

SAH: single alpha helix/helical. MD: molecular dynamics. AUC: analytical ultra-centrifugation. NMR: nuclear magnetic resonance. CD: circular dichroism. MRE: mean residue ellipticity. SAXS: small angle X-ray scattering. TROSY: transverse relaxation optimized spectroscopy.

Table 1. Measured $^{3}J_{NC}$ values, heteronuclear NOE values and relaxation rates for Arg Nε.

Arginine	$^{3}J_{N\varepsilon C\beta}$ (Hz) [b]	hetNOE [c]	R_1 (s^{-1}) [d]	R_2 (s^{-1}) [e]	# Glu [f]
858 [a]	1.1	-1.00 ± 0.09	0.53	2.3	1
860 [a]	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] Overlapping peaks. [b] Errors for $^{3}J_{NC}$ were calculated by finding the range in $^{3}J_{NC}$ values through adding and subtracting the baseline noise level to the peak intensities (<0.1 Hz where not shown). [c] hetNOE errors are standard deviations from the two experiments (TROSY and non-TROSY). [d] R_1 errors are all <0.01 s^{-1}. [e] Other than those shown, R_2 errors are all <0.1 s^{-1}. [f] The last column shows the number of Glu neighbors compatible with forming an ion pair within the helix context for each Arg residue. NR: peaks for Arg873 and Arg920 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 $^{3}J_{NC}$ value.

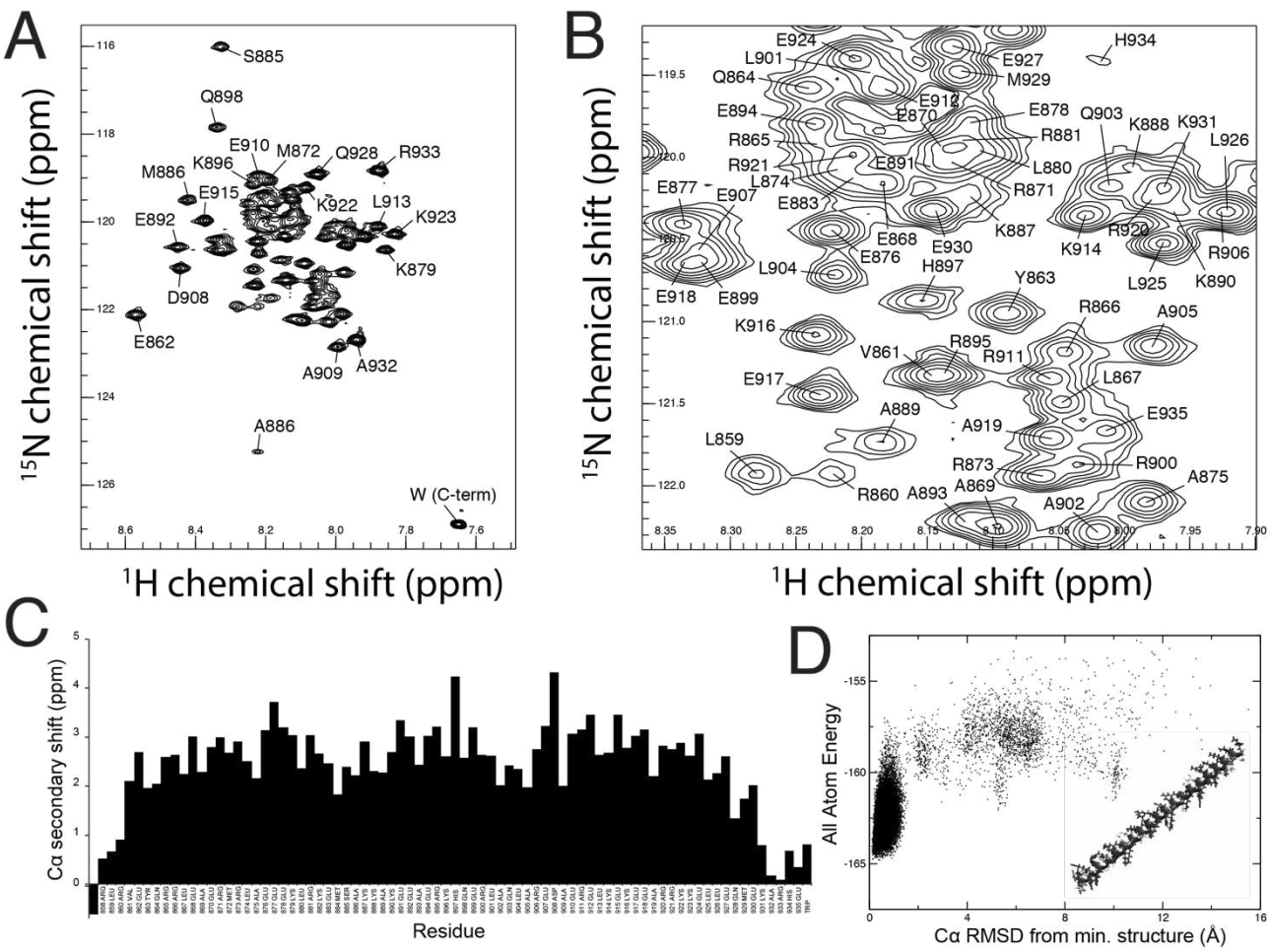

Figure 1. NMR TROSY spectrum for M7A SAH and chemical shift analysis. (A) ^{1}H–^{15}N TROSY spectrum (750 MHz) for M7A SAH. Note the limited dispersion; the vast majority of peaks occur within a 4 ppm (^{15}N) and 0.4 ppm (^{1}H) region, and there is a high degree of peak overlap. **(B)** A zoomed in 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 RMSD for Cα atom positions from those in the lowest energy structure. The lowest energy structure, which after exclusion of flexible tails ranges from Val861 to Glu930, is depicted on the plot.

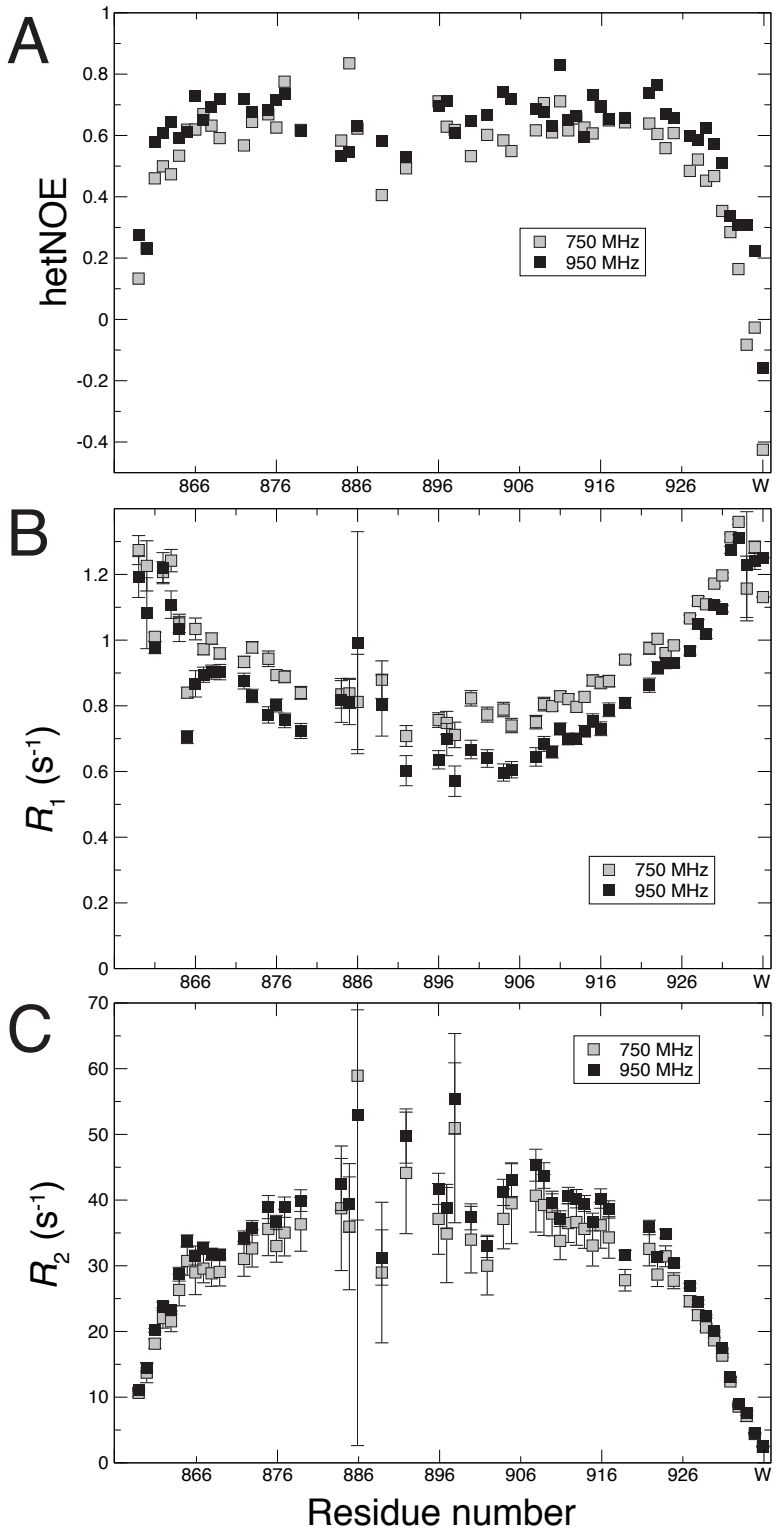

Figure 2. Dynamic properties from ^{1}H–^{15}N NOE and ^{15}N relaxation studies of M7A SAH. Plots of **(A)** ^{1}H–^{15}N NOE, **(B)** longitudinal (R_1), and **(C)** transverse (R_2) ^{15}N relaxation rates, as a function of residue number. Relaxation data were not recorded for a number of residues, most particularly between residue 876 and 896; many of the associated peaks were heavily overlapped with higher intensity neighbors. Error bars in (B) and (C) represent the error estimation from Monte Carlo-based fitting.

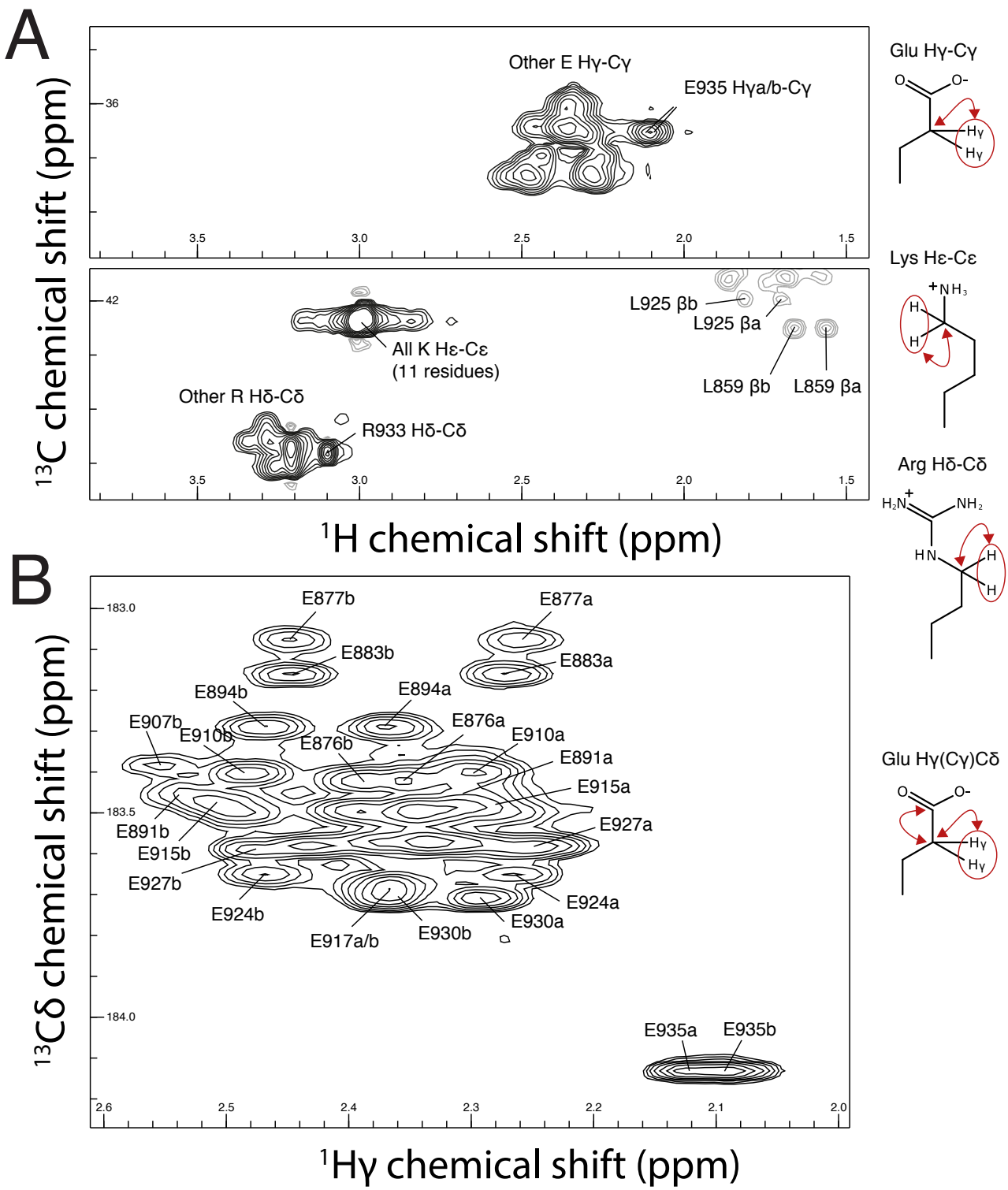

Figure 3. NMR spectra to interrogate side chain nuclei in Glu, Lys and Arg. (A) Sections of the ^{13}C-HSQC spectrum for M7A highlighting the most distal H–C correlation from the side chain of Glu (Hɣ–Cɣ), Lys (Hε–Cε) and Arg (Hδ–Cδ). **(B)** The Hɣ(Cɣ)Cδ Glu-specific spectrum exhibits improved peak dispersion over the Hɣ–Cɣ correlations in the ^{13}C-HSQC.

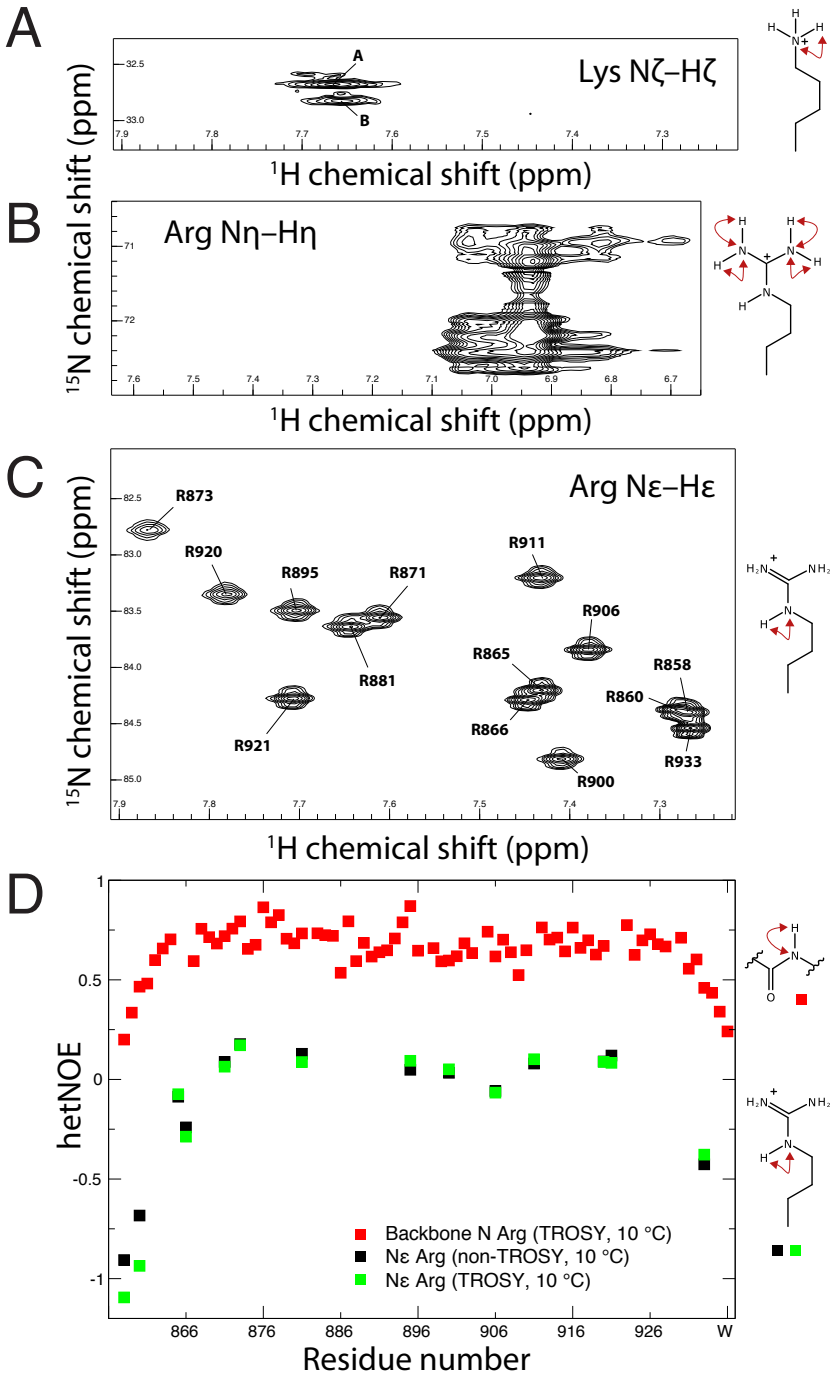

Figure 4. ^{15}N-HiSQC NMR spectra for side chain Lys or Arg N–H groups, and heteronuclear NOE data for Arg Nε–Hε at pH 5.5, 10 °C. Sections of ^{15}N-HiSQC spectra for M7A SAH highlighting the lysine NH_3 region ~33 ppm **(A)**, Arg Nη–Hη region ~72 ppm **(B)**, and Arg Hε–Nε region ~80 ppm **(C)**. **(D)** Heteronuclear NOEs for backbone ^{15}N nuclei (red) and for Nε nuclei in all fourteen Arg residues (black/green). A TROSY version of the heteronuclear NOE experiment was required in order to resolve the backbone amide correlations at this temperature.

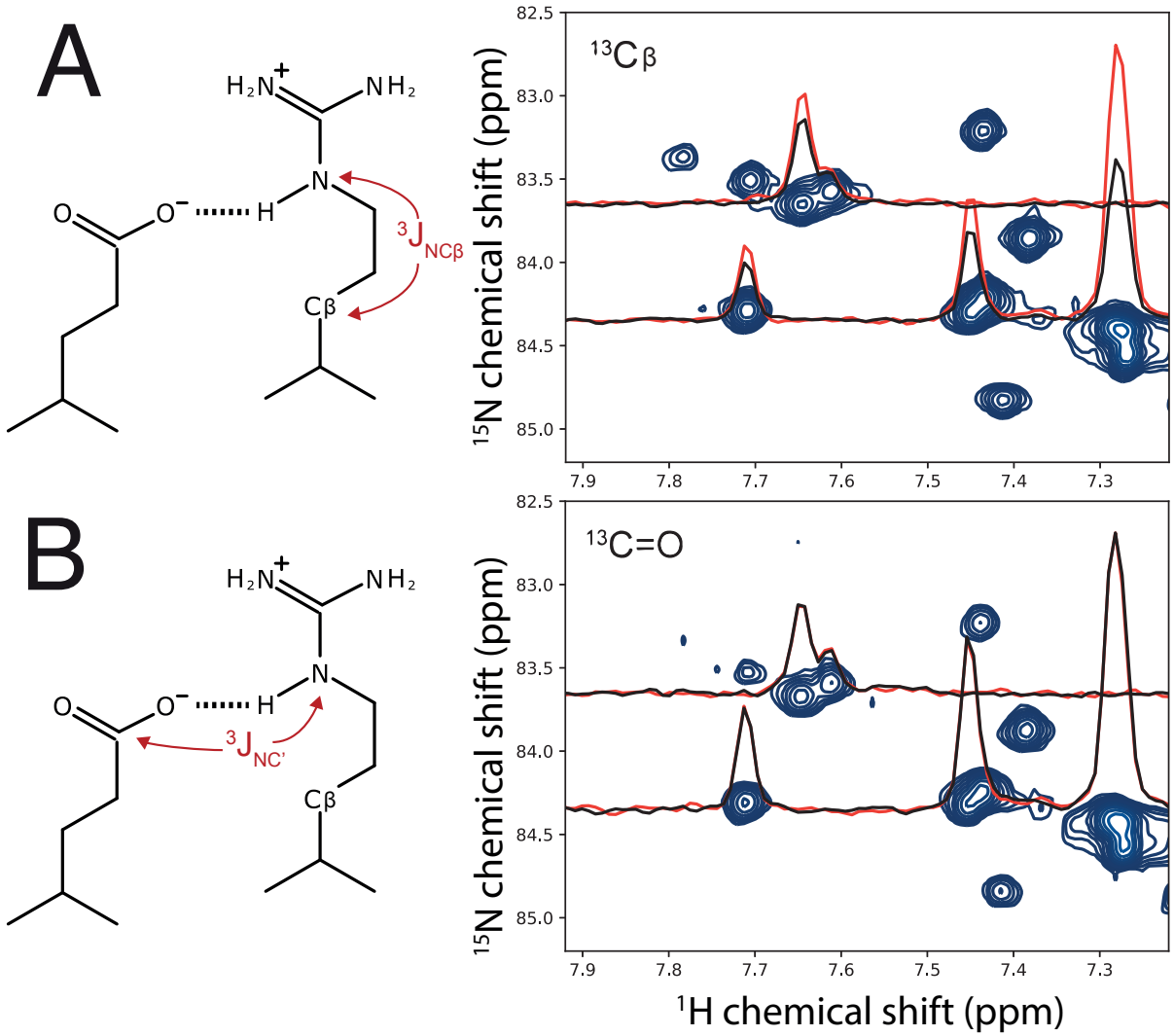

Figure 5. Spin-echo difference ^{1}H–^{15}N correlation experiments to measure $^{3}J_{NC}$. (A) Reference sub-spectrum for the Arg NεCβ experiment. Overlain are two example line scans from the reference (red) and the attenuated sub-spectrum (black). A reduction in peak intensity between the two sub-spectra indicates a measurable $^{3}J_{NC}$ coupling for each peak. Peaks for Arg873 (and Arg920) 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 $^{3}J_{NC}$ value. **(B)** Reference sub-spectrum for the through-hydrogen bond $^{3}J_{NC}$ 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 $^{3}J_{NC}$ for each peak. Peaks for Arg873 and Arg920 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).

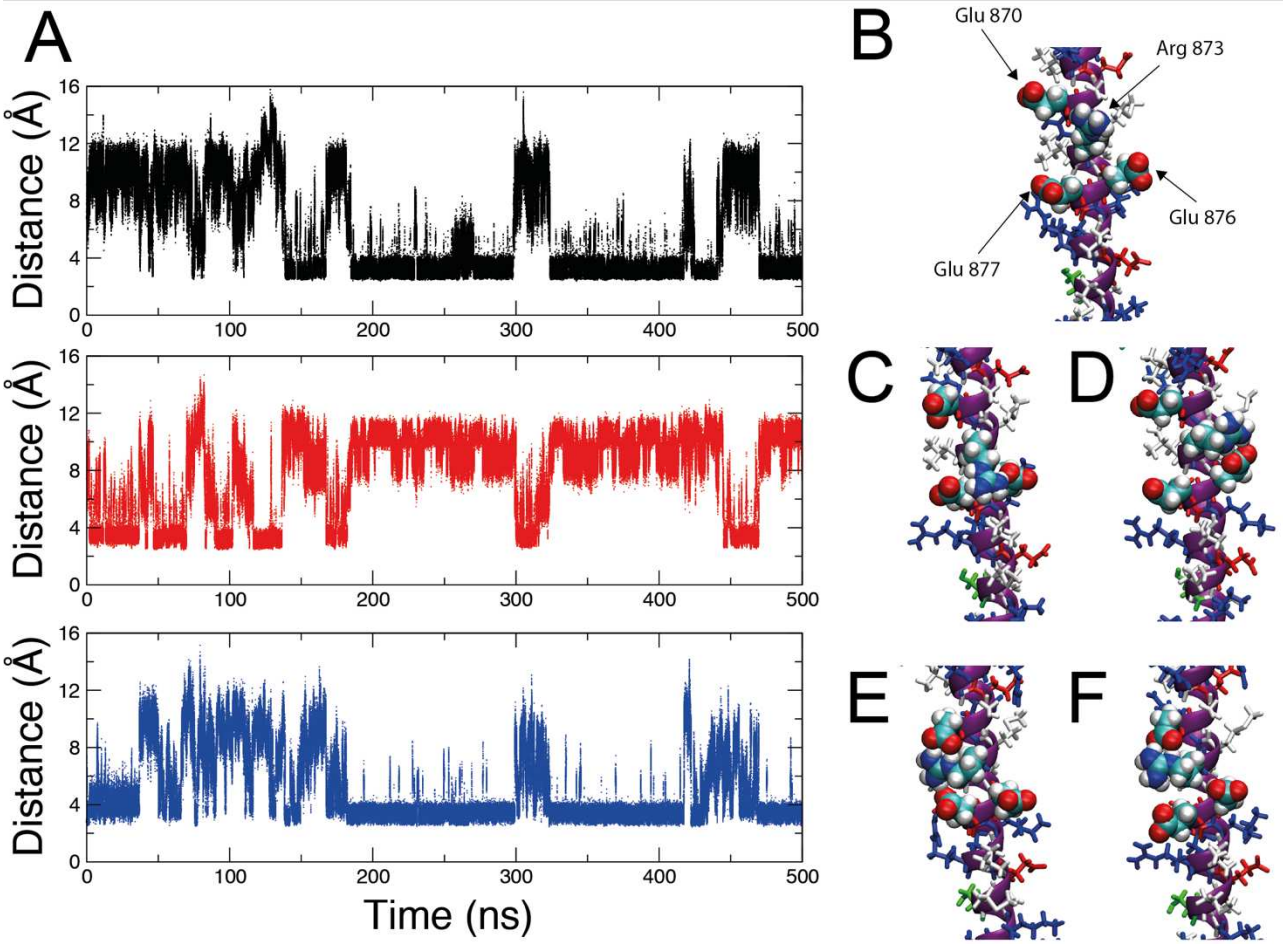

Figure 6. Simulation results highlighting the variable nature of ion pairs for an example residue Arg873. (A) The distances between charged residue pairs throughout the simulation. Arg873–Glu870 (black), Arg873–Glu876 (red) and Arg873–Glu877 (blue). The distance criterion chosen to mark the formation of an ion pair was 4 Å (5). Parts **(B–F)** show snapshots from the simulation indicating many interaction modes. A movie of the full simulation trajectory centered on Arg873 is provided (Movie S1). 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. **(B)** No ion pair formation. **(C)** Simultaneous Arg873–Glu876 and Arg873–Glu877 ion pair formation. **(D)** Arg873–Glu876 only. **(E)** Simultaneous Arg873–Glu870 and Arg873–Glu877 ion pair formation. **(F)** Arg873–Glu870 only.

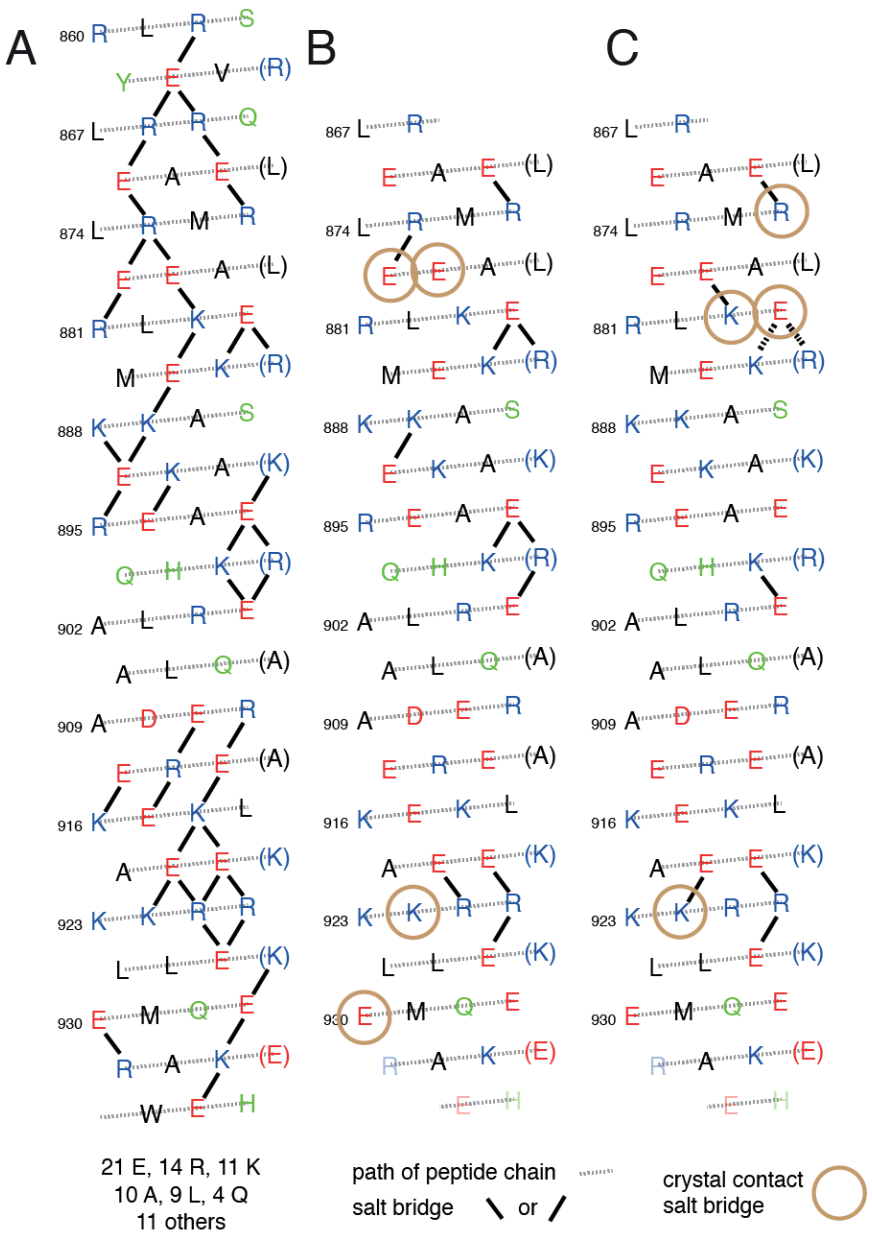

Figure 7. Helical net plot representation highlighting potential pairs in M7A$_{858–935}$ (M7A SAH) and observed ion pairs between side chains in the crystal structure of M7A$_{866–935}$. (A) The sequence of M7A SAH including the additional N-terminal Ser and C-terminal Trp shown as a helical net plot. All potential K/R$_i$–E$_{i+4}$, E$_i$–K/R$_{i+4}$, E$_i$–K/R$_{i+3}$, and K/R$_i$–E$_{i+3}$ ion pairs and the number of E, R, K and other residues are shown. **(B)** In chain A from PDB: 5WST, 11 ion pairs are observed. **(C)** In chain B from PDB: 5WST, 8 ion pairs are observed. The side chain of Glu878 displays two conformers (a and b) in the crystal structure that form an ion pair to either Arg881 or Lys882 (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 Glu876(A)–Lys879(B) (N–O 3.7 Å), Glu877(A)–Lys922(B) (N–O 2.7 Å), Lys922(A)–Glu878b(B) (N–O 2.2 Å), and Glu930(A)–Arg871(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.).

Dynamic ion pair behavior stabilizes single alpha helices in proteins

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

J. Biol. Chem. published online December 28, 2018
