PCCP

The Ionic Liquid [C_4mpy][Tf_2N] Induces Bound-like Structure in the Intrinsically Disordered Protein FlgM

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-04-2019-001882.R1
Article Type:	Paper
Date Submitted by the Author:	10-Jul-2019
Complete List of Authors:	Carter, Erin; Northern Arizona University, Department of Chemistry and Biochemistry Heyert, Alexanndra; Northern Arizona University, Department of Chemistry and Biochemistry De Souza, Mattheus ; The College of New Jersey, Department of Chemistry Baker, Joseph; The College of New Jersey, Department of Chemistry Lindberg, Gerrick; Northern Arizona University, Department of Chemistry and Biochemistry

The Ionic Liquid [C_4mpy][Tf_2N] Induces Bound-like Structure in the Intrinsically Disordered Protein FlgM

Erin E. Carter,[a] Alexanndra J. Heyert,[a] Mattheus De Souza,[c] Joseph L. Baker,[c] and Gerrick E. Lindberg[a,b,*]

a. Department of Chemistry and Biochemistry, Northern Arizona University, Flagstaff, AZ
b. Center for Materials Interfaces in Research and Applications, Northern Arizona University, Flagstaff, AZ
c. Department of Chemistry, The College of New Jersey, Ewing, NJ
*Email contact: gerrick.lindberg@nau.edu

Abstract

The *A. aeolicus* intrinsically disordered protein FlgM has four well-defined α-helices when bound to σ^{28}, but in water FlgM undergoes a change in tertiary structure. In this work, we investigate the structure of FlgM in aqueous solutions of the ionic liquid [C_4mpy][Tf_2N]. We find that FlgM is induced to fold by the addition of the ionic liquid, achieving average α-helicity values similar to the bound state. Analysis of secondary structure reveals significant similarity with the bound state, but the tertiary structure is found to be more compact. Interestingly, the ionic liquid is not homogeneously dispersed in the water, but instead aggregates near the protein. Separate simulations of aqueous ionic liquid do not show ion clustering, which suggests that FlgM stabilizes ionic liquid aggregation.

Introduction

Room temperature ionic liquids (ILs) are a class of molten salt materials that remain liquid at temperatures relevant for the manipulation of protein structure and function.[1,2] The interplay of aqueous salts and protein behavior is often described with the Hofmeister series[3] and ILs can dramatically expand its resolution and specificity. The ability of ions to induce protein structure and aggregation could be a powerful tool when studying protein function and structure relationships, crystallization, modeling cellular environments, and tackling other outstanding questions in protein science.[4–10] While the use of IL-based materials for such biochemical applications is desirable, it is challenging to select a specific IL because of the immense variety of possible cations, anions, and other components.[1] However, the large number of possible cation/anion combinations also provide great opportunity to design a chemical environment that is tuned to achieve specific physical properties.

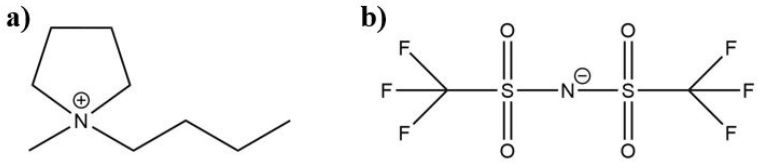


Figure 1: The ionic liquid (IL) 1-butyl-1-methylpyrrolidinium bis(trifluoromethylsulfonyl)imide [C_4mpy][Tf_2N] cationic (a) and anionic (b) components. Together, these ions form one IL pair.

Therefore, it is in the spirit of selecting an IL to have a specific effect on a protein that we have sought to characterize the effect of the IL 1-butyl-1-methylpyrrolidinium bis(trifluoromethylsulfonyl)imide ([C_4mpy][Tf_2N]; shown in Fig. 1) on the structure of the protein FlgM from the thermophilic bacterium *Aquifex aeolicus*.[11] [C_4mpy][Tf_2N] is a well-studied IL that has been shown previously to promote protein helical content in Trp-cage and AKA2,[12] but other work has shown that [C_4mpy] and [Tf_2N] ions individually act to destabilize the structure of ribonuclease A.[13] While promoting helical content in Trp-cage, [C_4mpy][Tf_2N] was found to induce cis/trans isomerizations of the protein backbone that are not observed in water.[14,15] ILs have received attention in the literature as materials to induce or modify protein structure, much of which has been covered in recent reviews.[8,16,17] Nevertheless, we briefly highlight a few salient studies to provide motivation and context for this work. Sajeevan and Roy showed that the IL 1-butyl-3-methylimidazolium chloride can induce a 3_{10} to α-helix conversion, demonstrating that ILs can cause subtle changes in protein structure.[18] Pfaendtner and coworkers have studied five cellulases and shown that ILs can be selected to reduce enzyme secondary structure or even disrupt most of the secondary structure.[19–21] They found that enzymes with increased negative charge on

their surface are more likely to resist secondary structure disruption.[20] Demonstrating that the effects of an IL on a protein are a nontrivial function of the specific IL, protein, and conditions.

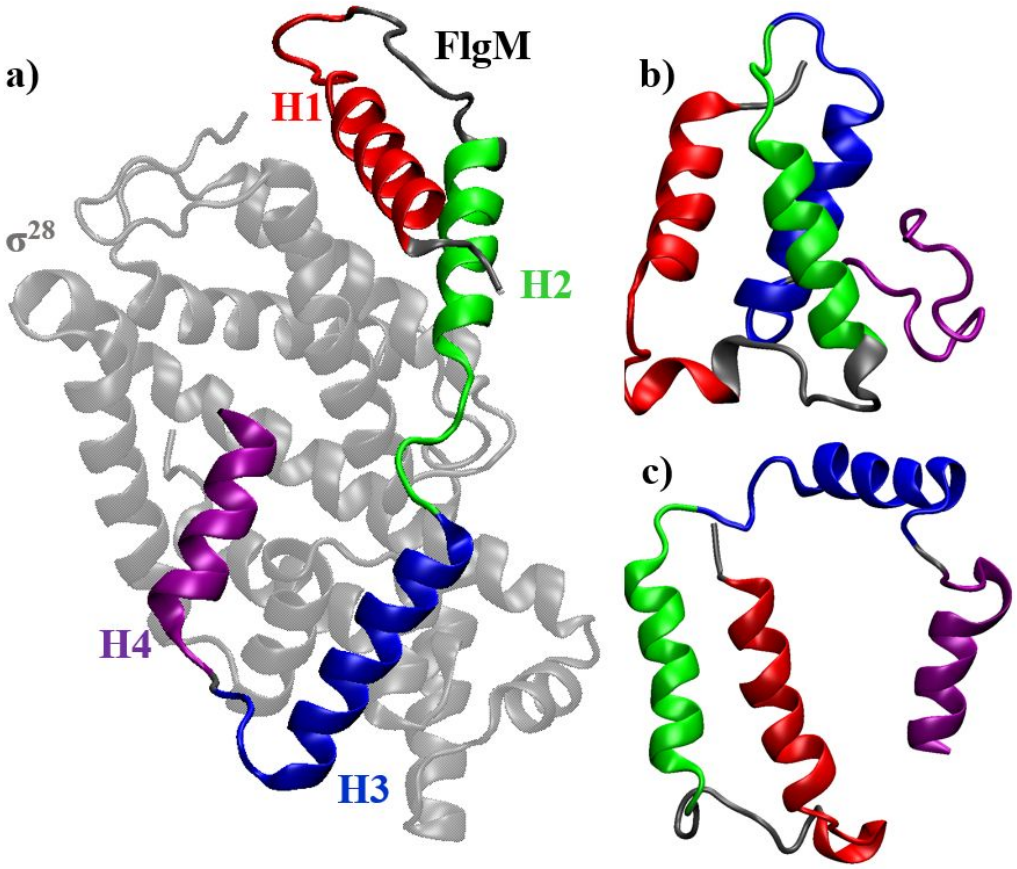


Figure 2: Representative structures of FlgM a) bound to σ^{28} in aqueous solution, b) in aqueous solution, and c) in aqueous IL solution with 50 IL pairs. The colors distinguish the four α-helical regions of FlgM and unstructured, connecting residues are shown in dark grey. Each structure is obtained from a simulation at 358 K. In panel a), σ^{28} is shown in light grey.

A. aeolicus is a thermophilic bacterium that is found underwater at high temperatures, for example in the hot springs of Yellowstone,[22] and has an optimal growing temperature of about 358 K.[23] FlgM proteins have been an important model class for the study of intrinsically disordered proteins,[24] but the disordered nature of FlgM varies between organisms.[24,25] The structure of *A. aeolicus* FlgM contains four α helices connected by disordered regions and no significant overall tertiary structure (Fig. 2).[23–25] The structure shown in Fig. 2a is that of FlgM when complexed with

σ^{28} (when not bound to FlgM, σ^{28} directs flagellar gene transcription[26]). FlgM, however, is less structured in water when not bound to σ^{28} (Fig. 2b). Previous work has shown that the degree of order in *A. aeolicus* FlgM is temperature dependent.[23,25,27] At low temperatures, FlgM is found to be more structured than at the physiological temperature of *A. aeolicus*.[23] Additionally *A. aeolicus* FlgM has α-helical character that decreases as the temperature is increased,[23] and helix 4 (H4) is disordered and fluctuates around an ordered core comprised of helices 1 through 3 (H1, H2, and H3) that retains significant α-helical structure, even at high temperature (358 K).[27] Although *A. aeolicus* FlgM does have a more ordered structure at 293 K than at its physiological temperature of 358 K,[23] it does not contain the structural features often found in thermophilic proteins that are associated with maintaining secondary structure at elevated temperatures (e.g. salt bridges, hydrogen bonding).[28] Intrinsically disordered proteins, however, often possess greater proportions of hydrophilic residues,[29] which suggests that they will be susceptible to manipulation by ions. Additionally, the low vapor pressure, low combustibility, and thermal stability of most ILs, make them an attractive solvent for understanding the high temperature behavior of thermophilic proteins. Therefore, considering the significant interest in designing materials to affect protein structure and the particular challenge of characterizing intrinsically disordered proteins, it is interesting to design solvent mixtures that can induce structure in a protein like FlgM.

In this work, we present results showing that the IL $[C_4mpy][Tf_2N]$ can induce α helicity in FlgM that is similar to the bound state. In section 2, we review the computational molecular dynamics methods used. In section 3, we present our results and provide discussion. Finally, in section 4, we offer conclusions and suggestions for future work.

Methods

The structure of FlgM in complex with σ^{28} was obtained from the Protein Data Bank (PDB ID: 1SC5).[11] FlgM was isolated from σ^{28} and unresolved residues (residues 1-2 and 18-31) were added as random coils using Modeller 9.14.[30] Starting configurations were prepared using Packmol[31] to solvate the protein with [Tf_2N] anions, [C_4mpy] cations, and finally water. A salt concentration of 0.15 M NaCl was achieved using tLEaP.[32] FlgM was modeled with the Amber ff14SB force field,[33] sodium and chloride ions with the parameters developed by Joung and Cheatham,[34] and water with the TIP3P model.[35] [C_4mpy] and [Tf_2N] interactions were described using the parameters developed by Xing et al.[36] with IL partial charges scaled to 0.8 q, which is consistent with previous works.[21,37–40] A total of 27 simulations were carried out with increasing numbers of cation/anion pairs of the IL [C_4mpy][Tf_2N] and other modifications to protein composition. The simulations reported here include FlgM bound to σ^{28}, FlgM in water, FlgM in aqueous IL solutions, and aqueous IL solution (without protein). The simulations of aqueous IL include the addition of 10, 20, 30, 40, and 50 IL pairs. The number of IL pairs are chosen to provide a range of concentrations up to near-saturated aqueous IL conditions, based upon literature solubility information of [C_4mpy][Tf_2N] in water.[41]

We use periodic boundary conditions and bonds involving hydrogen were restrained with the SHAKE algorithm,[42] which permits a timestep of 2 fs. Temperature was controlled with the Langevin thermostat[43] with a collision frequency of 1 ps^{-1} and pressure was controlled with the Monte Carlo barostat.[44] The particle mesh ewald (PME) method was used to handle long-range electrostatics[45] and a cutoff of 8 Å was used for non-bonded interactions.

Minimization and molecular dynamics simulations are performed with the AMBER simulation package version 14.[32] All systems were prepared using a four stage protocol: energy

minimization, heating with restraints, gradual removal of restraints, and unrestrained simulation. The systems were energy minimized in three phases that are comprised of 1000 steps of steepest descent followed by 4000 steps of conjugate gradient. The first phase included 10 kcal mol^{-1} $Å^{-2}$ restraints on the entire protein, the second phase included 10 kcal mol^{-1} $Å^{-2}$ restraints on only the α-carbons, and the third phase included no restraints. Gradual heating was carried out in two phases in the NVT (constant number of atoms, constant volume, and constant temperature) ensemble. In the first phase, 10 kcal mol^{-1} $Å^{-2}$ restraints were imposed on the protein backbone while the temperature was raised from 0 K to either 300 or 358 K over 60 ps. In the second phase, the temperature was held constant for 40 ps. Once the systems were at their final temperature (300 K or 358 K), restraints on the protein were gradually reduced through a series of five simulations in the NPT (constant number of atoms, constant pressure, and constant temperature) ensemble. Each simulation was 1 ns long with decreasing harmonic force constants of 10.0, 5.0, 2.5, 1.0, and 0.5 kcal mol^{-1} $Å^{-2}$. Conventional unrestrained MD simulations were run at the final temperature for 200 ns. Each of the unbound FlgM systems in aqueous IL solution were simulated in duplicate. We report results from eight systems and a total of 27 simulations. In total, 5.4 μs of MD simulation was performed. Table S1 lists all systems studied in this work and the simulation durations.

The visualization of molecular structures is done with VMD.[46] When possible analysis was performed with AmberTools,[47,48] but analysis programs were written in-house otherwise. Much of the analysis focuses on the behavior of the four α-helices that make up the FlgM secondary structure and the residue numbers used to define these regions are provided in Table S2. Many of our simulation workflows have been automated in Parsl and are available upon request.[49] Parsl is a Python-based parallel scripting library intended to facilitate high performance scientific computing workflows.

We analyze spatial heterogeneity of the ionic liquid using the following clustering algorithm. Two ions are deemed to be neighbors if the distance between them is less than the neighbor cutoff distance, which is the first minimum in the ion-ion radial distribution function. The ion clusters are determined by the following protocol:

1. A random ion is selected
2. All neighbors within the neighbor cutoff are found
3. If a neighbor is found, that ion is added to the cluster and then all neighbors of that ion within the neighbor cutoff are identified
4. This search continues to progressively populate the cluster until all ions are placed or no more neighbors are found
5. If there are remaining ions not placed in a cluster, then the process continues from Step 1

Following this scheme, all ions are grouped into clusters. Note that a cluster can contain just one ion, if there are no other ions within the neighbor cutoff distance. This is similar to how aggregation has been studied previously; for example, the study by Mustan and coworkers.[50]

Root mean square deviation is calculated with the AmberTools utility cpptraj using the FlgM structure obtained from the protein data bank after missing residues are added. Radial distribution functions were calculated between each amino acid and IL ion for each system. This constitutes a prodigious amount of information, which we digested by looking at the maxima of the radial distribution functions. The radial distribution functions were processed to make local maxima plots by scanning each radial distribution function looking for local maxima (peaks). To avoid spurious effects arising from noise, a point is only selected as a maximum if it is greater than the surrounding six data points in the radial distribution function.

Results and discussion

Since bound FlgM is primarily comprised of α-helical secondary structure and [C_4mpy][Tf_2N] has been shown to increase helical structure in other proteins, it is natural to examine helical content of the protein in solutions of this IL (Fig. 3 and 4). Fig. 3 shows the total percent α-helical secondary structure in FlgM for each concentration. There is a positive trend, which is highlighted by the linear trend lines. Both the 300 and 358 K data show increasing helicity with IL concentration, but the observation is much more subtle for the 300 K simulations. The 300 K behavior of FlgM without IL appears to be an outlier and may be due to kinetic trapping of FlgM in the starting folded state. While this study is the first to look at the effect of an IL on FlgM, other studies have examined the structure of *A. aeolicus* FlgM. For example, Ma et al. showed that at 293.15 K FlgM is 40.5±1.0% α helical,[24] which would indicate over estimation of helicity due to kinetic trapping at 300 K in this work.

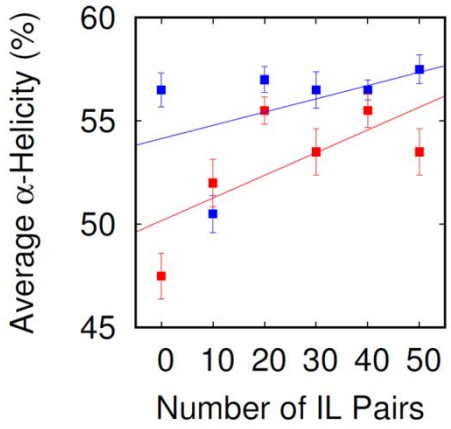


Figure 3: Average percent α-helicity of FlgM plotted against increasing numbers of IL pairs in solution at both 300 K (blue) and 358 K (red).

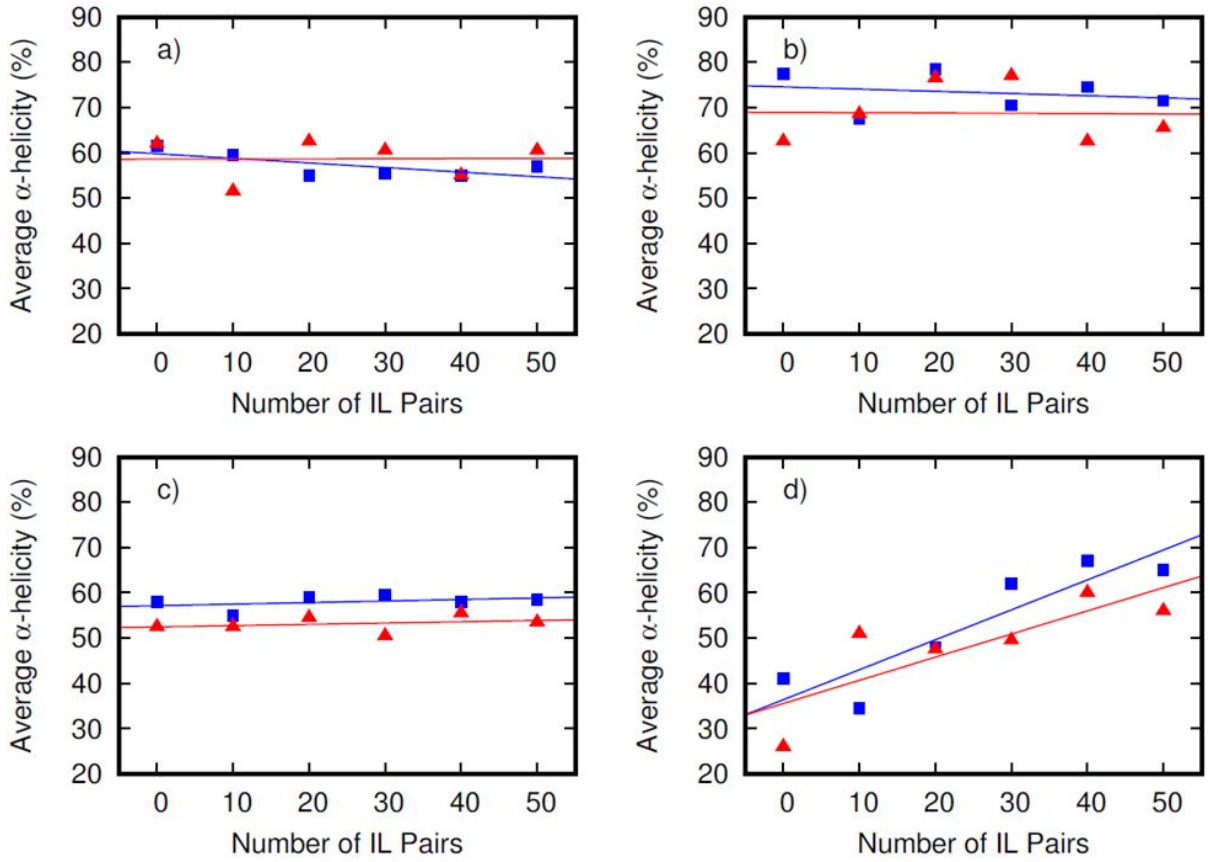


Figure 4: Average percent α-helicity of each helix: a) H1, b) H2, c) H3, and d) H4 with increasing concentration of IL at both 300 K (blue squares) and 358 K (red triangles).

However, it is difficult to discern local changes in FlgM structure from analysis of the entire protein. Therefore, considering that FlgM can be described in terms of four distinct helical domains, it is logical to analyze each of these domains independently. We find that H1, H2, and H3 do not show significant changes in average percent α-helical secondary structure as a function of the IL concentration (Fig. 4a-c), while H4 is found to increase significantly in average percent α-helicity as the concentration of IL increases (Fig. 4d). The representative structure in Fig. 2b shows that H4 unfolds in water when not bound to σ^{28} and Fig. 2c shows the folded structure of FlgM in aqueous IL with 50 IL pairs. Fig. 4d shows that the IL increases structure of H4, but it is unclear how this structure compares to biologically active structures of FlgM. Therefore, we also simulated FlgM bound to σ^{28} and the average percent α-helicity observed in the bound simulations

is shown in Table 1. In pure water and at each concentration of IL considered, H1, H2, and H3 α-helicity does not significantly differ from when FlgM is bound to σ^{28} in water (Fig. 4a-c and Table 1). At low IL concentrations, however, H4 becomes significantly less structured than in the σ^{28} bound state, but as the IL concentration increases FlgM is found to have secondary structure more similar to FlgM bound to σ^{28}. Comparison of Fig. 3 with the average helicity of FlgM when bound to σ^{28} demonstrates that the secondary structure of FlgM becomes more bound-like when the concentration of IL is increased (Table 1). The similarity of secondary structure in the bound state and at high IL concentrations can be observed quantitatively by comparing the average helicities in the 50 IL solution to the bound state values. Table 1 shows the average helicities of FlgM when bound to σ^{28} and in the 50 IL solution as well as the percent difference between them. At 300 K, all of the observed percent differences are below 12% and the absolute differences are comparable in magnitude to the standard error. At the higher temperature, FlgM is generally observed to be less helical and the standard errors are larger, indicating increased disorder and conformational freedom. Therefore, we conclude that high IL concentration results in secondary structures that are similar to those observed when FlgM is in the bound state. It is interesting to note that despite the similarity of FlgM helicity in each environment, the bound helicities are observed to be greater in all environments except one.

Table 1: Percent α-helicity of FlgM when bound to σ^{28} or in the 50 IL solution and their percent difference

Temperature (K)	Region of FlgM	Helicity when bound to s^{28} (%)	Helicity in 50 IL solution (%)	Percent difference
300	All	59.1±2.3	57.5±1.4	3
	H1	64.9±4.3	57.0±0.6	12
	H2	77.0±6.6	71.5±0.4	7
	H3	55.6±4.3	58.5±0.6	5
	H4	73.1±4.6	65.0±0.9	11
358	All	56.0±3.1	53.5±2.3	4
	H1	69.4±8.0	60.5±0.9	13
	H2	73.6±5.6	65.5±1.1	11
	H3	56.9±4.4	53.5±0.8	6
	H4	73.1±5.4	56.0±1.4	23

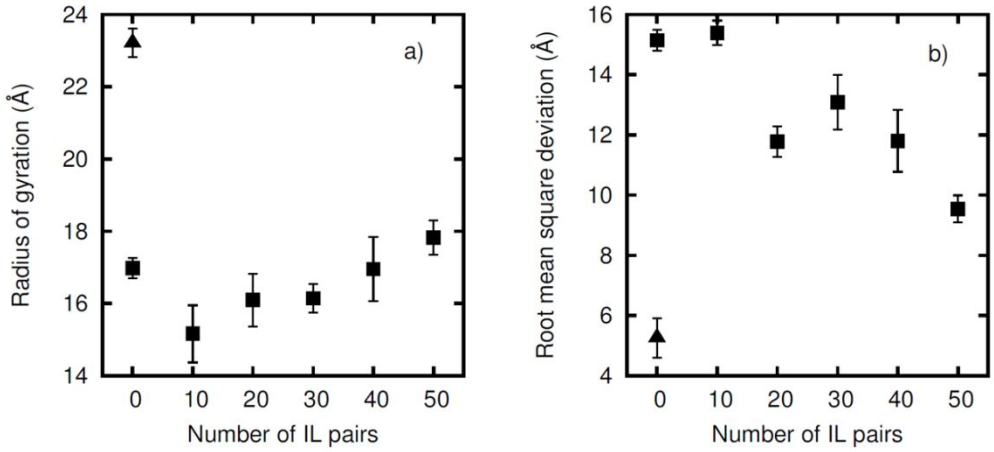


Figure 5: Average a) radius of gyration and b) root mean square deviation of FlgM at 300 K in solution (squares) and when bound to σ^{28} (triangles). The radius of gyration shows that FlgM is more compact when not bound to σ^{28}, and extends a little as the IL concentration is increased. Interestingly, the root mean square deviation of FlgM reveals that FlgM becomes more similar to the bound state in the IL solution.

While the IL is found to induce α-helicity in FlgM that is similar to when it is bound to σ^{28}, the tertiary structure is generally different. Figure 5a shows the average radius of gyration for FlgM in solutions with increasing concentrations of IL and when bound to σ^{28}. The FlgM radius of gyration is found to be significantly smaller in solution than when bound to σ^{28}. The radius of gyration increases as the concentration of IL is increased, but the protein remains significantly more compact at all IL concentrations than when bound. The different tertiary structure between bound Flgm and FlgM in high IL concentration can be observed by comparing Figs. 2a and 2c. The structure in Fig. 2a shows FlgM draped across the surface of σ^{28} in an extended conformation, while in Fig. 2c FlgM takes on a compact conformation. The radius of gyration demonstrates how despite similarities in secondary structure, the tertiary structure of FlgM is significantly different. Such dramatic changes to the tertiary structure in response to environment are common in IDPs.

Figure 5b shows the root mean squared deviation of FlgM from the crystal structure for FlgM in solutions with increasing IL concentrations and when bound to σ^{28}. The root mean squared deviation of FlgM from the crystal structure is found to generally decrease as IL concentration is increased. At the highest concentration considered, the root mean squared deviation remains large compared to that of FlgM bound to σ^{28}, but its similarity to the bound structure is significantly increased. It is natural to wonder if even higher IL concentrations would become more similar, but our highest concentration is at the solubility limit of $[C_4mpy][Tf_2N]$ in water. Characterization of FlgM in neat $[C_4mpy][Tf_2N]$ is also possible, but the significantly higher viscosity of the IL compared to water makes determining the equilibrium ensemble challenging. See Baker et al. for a brief discussion of the viscosity of $[C_4mpy][Tf_2N]$ and a comparison with that of water.[14] The observation that FlgM has a significantly more compact structure that deviates from the bound structure at low concentration is expected because of the unstructured, intrinsically disordered

nature of FlgM when not bound to σ^{28}. Additionally, the observation that high IL concentrations make FlgM take on a slightly more extended conformation with increased similarity to the crystal structure suggests that IL solutions could be a useful environment to mimic many aspects of physiological conditions.

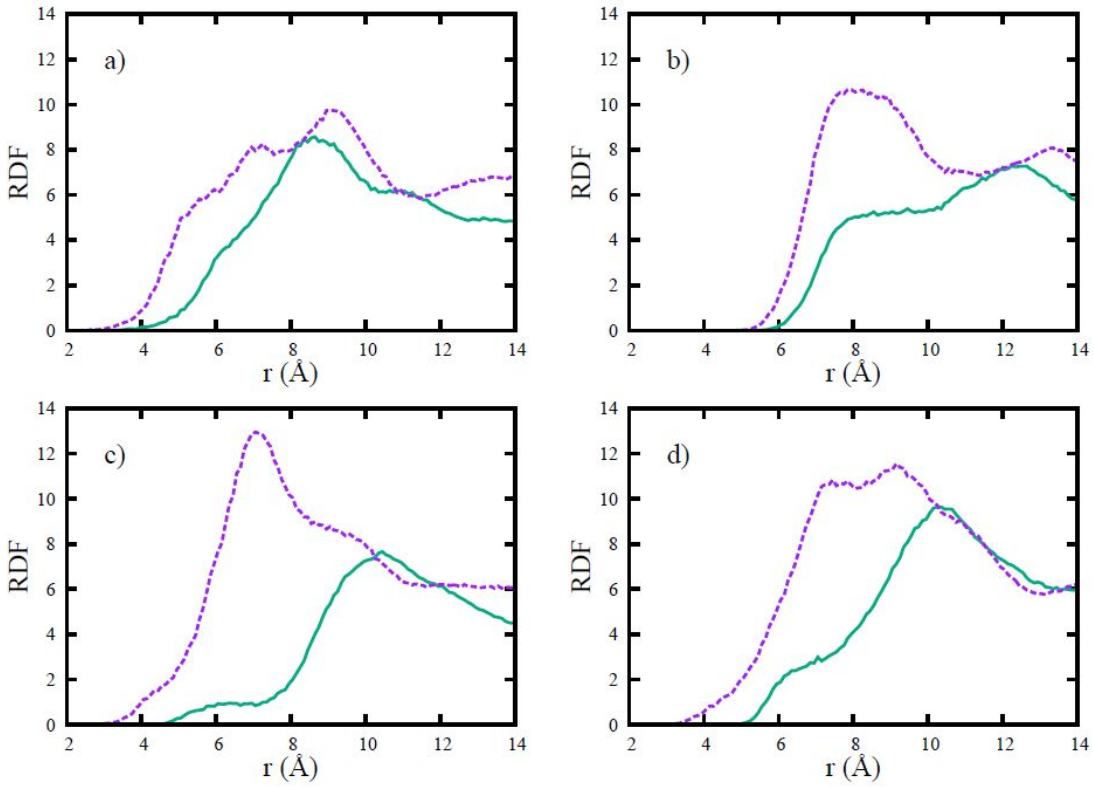


Figure 6: Radial distribution functions (RDFs) of the nitrogen atom in [Tf_2N] (purple) and the nitrogen atom in [C_4mpy] (green) to the center of mass of a) H1, b) H2, c) H3, and d) H4 for the 50 IL system at 300 K.

We next used radial distribution functions between each helix and both ions to investigate the local distribution of the ions around each helix (Fig. 6). For all four helices, the anion [Tf_2N] is found to more closely approach the helix than the cation. For H1 (Fig. 6a), the cation and anion distribution functions are generally similar, but for the other three helices [Tf_2N] has a first peak

near 7 Å, while [C_4mpy] has a shoulder in that region but reaches a maximum near 10 or 12 Å. These suggest that there are differences in the environment around each helix, but these differences are difficult to identify.

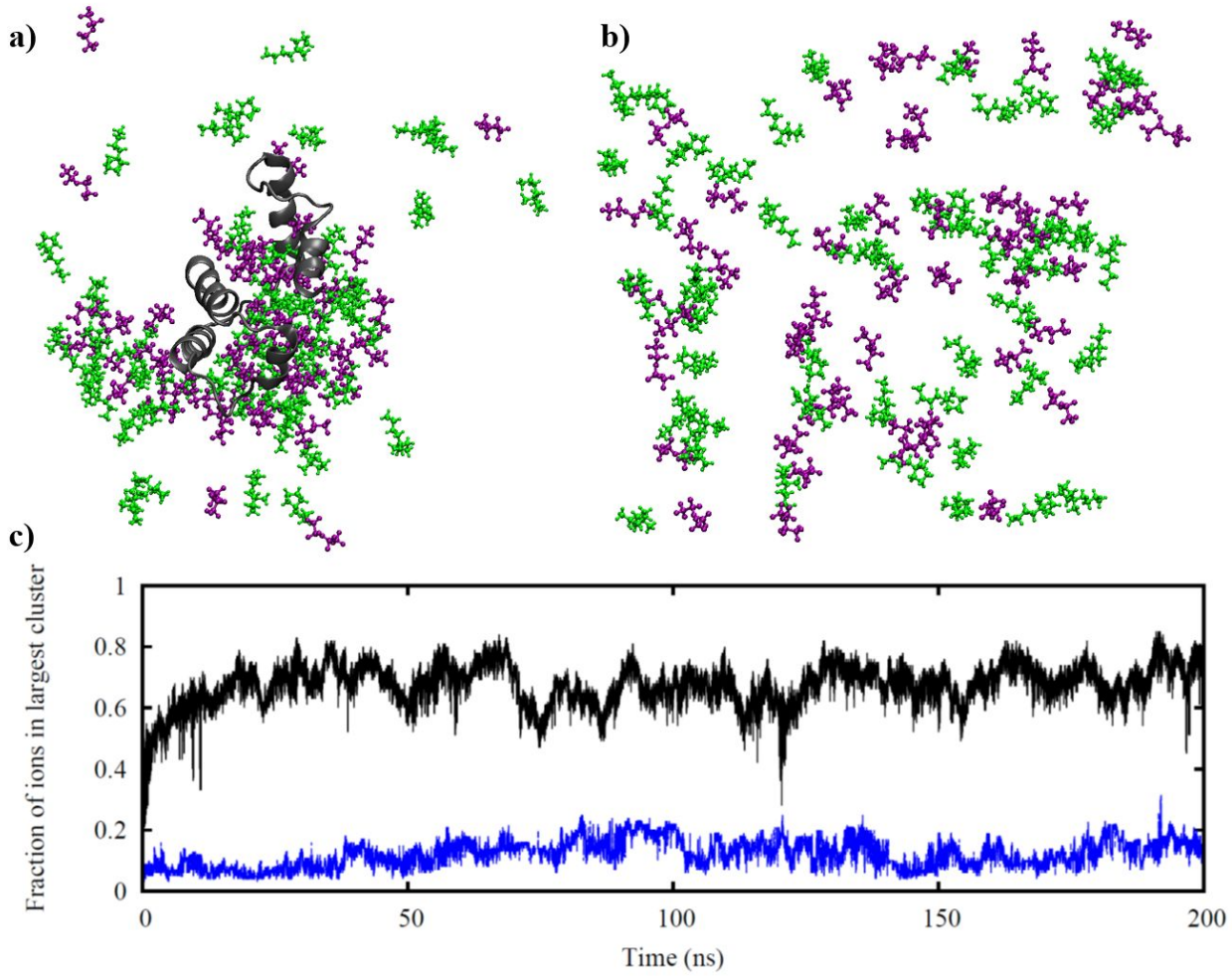


Figure 7: Distribution of IL ions with FlgM (a) and without FlgM (b) for simulations with 50 IL pairs in water at 300 K. [C_4mpy] cations (green) and [Tf_2N] anions (purple) are shown. Water molecules are omitted for clarity. When FlgM is present (shown in grey) the ions aggregate around the protein. Panel (c) shows the fraction of IL ions in the largest IL cluster with (black) and without (blue) FlgM.

To better understand why the α helicity of H4 increases to σ^{28} bound levels in high concentrations of IL, we consider the local environment around the protein in the aqueous IL simulations. The representative snapshot in Fig. 7a shows that the IL forms a cluster in the vicinity of the protein. The configuration in Fig. 7a is from the simulation with 50 IL pairs, but this is a

general feature of FlgM in each of the IL-containing solutions. The structure of the IL cluster is analyzed using radial distribution functions between each amino acid and the ions.

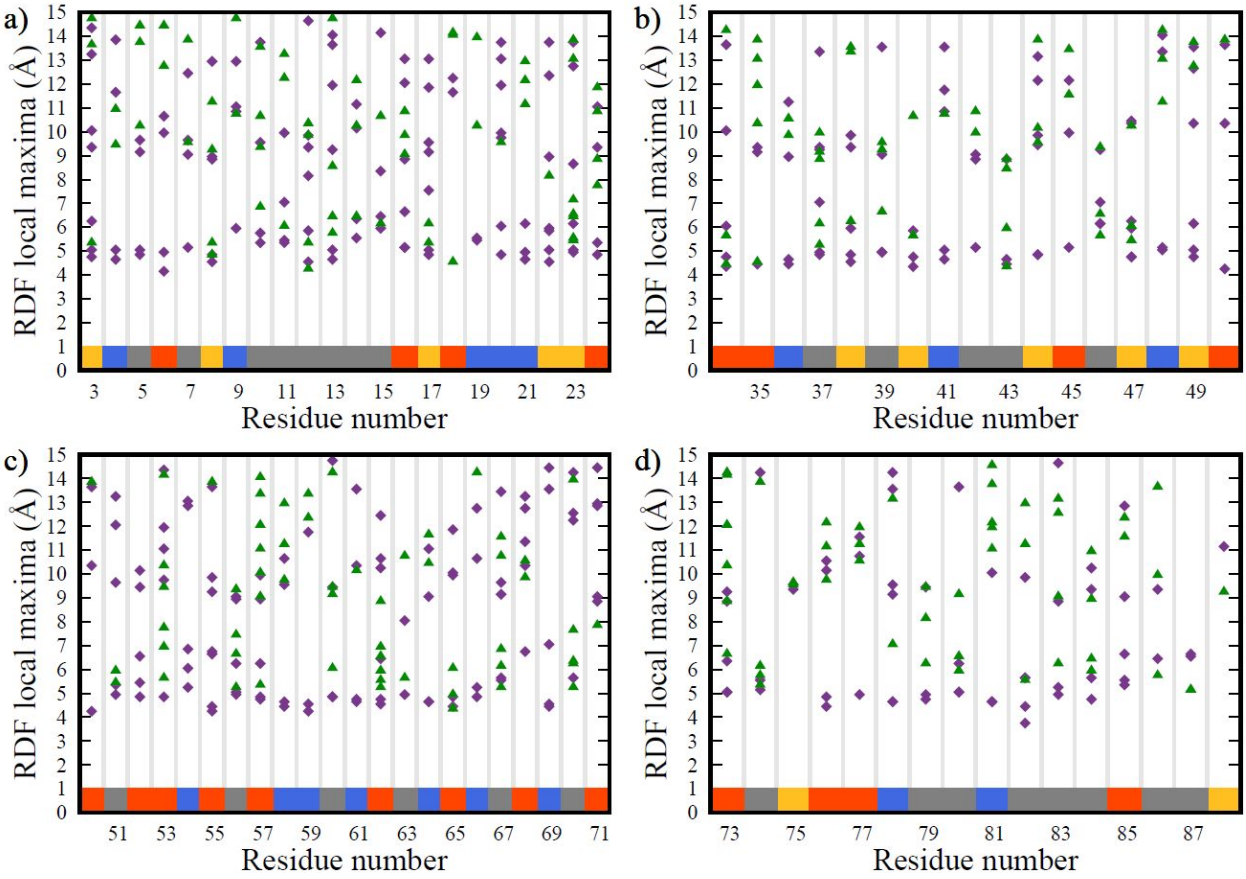


Figure 8: Local maxima present in the radial distribution function (RDF) between the center of mass of each residue and the nearest atom in a [C_4mpy] ion (green) or a [Tf_2N] ion (purple) for (a) H1, (b) H2, (c) H3, and (d) H4. The bar on the bottom of each graph classifies the residue as positively charged (blue), negatively charged (red), polar (yellow), or nonpolar (grey). This is a representative plot obtained from the 10 IL pair system at 300 K, but the trends are similar for other simulations.

The radial distribution functions between the IL ions and each residue reveal that there is order within the cluster. In Fig. 8 we plot each local maximum in the radial distribution function against the residue number for each IL ion. Similar to Fig. 6, Fig. 8 shows that [Tf_2N] more closely coordinates all helices. A few amino acids are observed to be closer to [C_4mpy] ions on average than [Tf_2N], but the cation is never found to be a lot closer. Interestingly, the closest ion to every amino acid in H1, H2, and H3 is about 4 to 6 Å away, but in H4 there are multiple

amino acids with ion first solvation shell maxima beyond 9 Å, suggesting that the absence of IL could be an important aspect of the observed behavior. The color bar at the bottom of each panel in Fig. 8 shows the type of each residue (see Fig. 8 caption), which reveals that even negatively charged amino acids are generally more closely coordinated by the anion [Tf_2N].

In an effort to understand if the cluster of IL ions is actually aggregation induced by the presence of the protein, we performed an identical simulation at the highest IL concentration without FlgM. If the IL is found to aggregate without the protein, then the aggregation phenomena would likely be independent of the protein (possibly due to saturation of the aqueous solution), however if the aggregation does not occur then this lends support to the hypothesis that the protein either induces IL aggregation or expedites the process. As shown in Fig. 7b, in the absence of FlgM the IL is not found to form large clusters in aqueous solution, instead remaining essentially evenly dispersed with only much smaller clusters spontaneously forming and disintegrating. This was quantified by calculating the number of ions in the largest cluster (Fig 7c). When the protein is present, the IL aggregates to form a large cluster containing about 70% of the ions in the system within 20 ns. However, without the protein, the IL is not observed to form a cluster with more than 25% of the ions in the system. This demonstrates that not only does the IL affect the protein, but likewise the protein affects the IL, facilitating aggregation of the ions.

The paradigm of protein function is that biological activity is directly tied to a defined three-dimensional structure. Globular proteins typically have a structure that is stable enough that they maintain it even in dilute aqueous solution, while intrinsically disordered proteins confoundingly often lack their functionally active three-dimensional structure outside the environment in which they are active.[51] The biological environment in which a protein functions

317 is often crowded with significant, non-negligible concentrations of many species. Proteins in
318 such heterogeneous environments are difficult to study directly because of their complexity and
319 the physiologically relevant protein structures cannot easily be deduced from simplified aqueous
320 environments. Therefore, relatively simple chemical analogues have been developed to mimic
321 cellular environments.[52,53] Similarly, solutions like the IL described here could serve as a
322 powerful intermediate between 'simple' aqueous solutions and complex *in vivo* environments,
323 permitting the study of protein structure in low hydration, crowded environments.

324 **Conclusion**

325 ILs are a diverse class of materials with the potential to manipulate biological systems,
326 including affecting protein secondary and tertiary structure. In this work, we have shown that the
327 IL $[C_4mpy][Tf_2N]$ can induce secondary structure in the intrinsically disordered protein FlgM to
328 a state with secondary structure that is similar to a physiological bound state. FlgM is comprised
329 of four α-helices, the first three of which are stable and have similar α-helicity when bound to σ^{28}
330 or unbound in water, but H4 is found to increase α-helicity significantly from 41% in water to
331 65% in saturated IL solution at 300 K. Future exploration of the effects of the IL on FlgM could
332 be done by using an enhanced sampling method to explore the precise interdependence of α-
333 helicity and coordination are explored. Metrics for quantitatively understanding IL effects on
334 protein structure are sought after and the RDF local maxima analysis appears to be a useful tool
335 to reduce the immense amount of solvation information in protein systems to reveal local trends.

336 **Conflicts of Interest**

337 There are no conflicts of interest to declare

338 **Author Contributions**

Conceived and designed the analysis: E.E.C.; J.L.B.; and G.E.L. Collected the data: E.E.C. and G.E.L. Performed the analysis: E.E.C.; A.J.H.; M.D.; J.L.B.; and G.E.L. Wrote the paper: E.E.C. and G.E.L.

Acknowledgements

We are grateful to Profs. Cynthia Hartzell and Andrew Koppisch for numerous helpful discussions. The simulations reported were performed with the Monsoon high performance computing cluster at Northern Arizona University, which is funded by Arizona's Technology and Research Initiative Fund. J.L.B. acknowledges the use of the ELSA high performance computing cluster at The College of New Jersey to perform some of the analysis for the research reported in this work. The ELSA cluster is funded by the NSF under award number OAC-1828163. J.L.B. acknowledges support from NSF award ACI-1550528 and G.E.L. from NSF award ACI-1550562.

References

1 K. N. Marsh, J. A. Boxall and R. Lichtenthaler, Room temperature ionic liquids and their mixtures—a review, *Fluid Phase Equilib.*, 2004, **219**, 93–98.

2 K. Ghandi and G. Khashayar, A Review of Ionic Liquids, Their Limits and Applications, *Green and Sustainable Chemistry*, 2014, **04**, 44–53.

3 H. I. Okur, J. Hladílková, K. B. Rembert, Y. Cho, J. Heyda, J. Dzubiella, P. S. Cremer and P. Jungwirth, Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions, *J. Phys. Chem. B*, 2017, **121**, 1997–2014.

4 M. L. Pusey, M. S. Paley, M. B. Turner and R. D. Rogers, Protein Crystallization Using Room Temperature Ionic Liquids, *Cryst. Growth Des.*, 2007, **7**, 787–793.

5 D. Hekmat, D. Hebel, S. Joswig, M. Schmidt and D. Weuster-Botz, Advanced protein crystallization using water-soluble ionic liquids as crystallization additives, *Biotechnol. Lett.*, 2007, **29**, 1703–1711.

6 R. A. Judge, S. Takahashi, K. L. Longenecker, E. H. Fry, C. Abad-Zapatero and M. L. Chiu, The Effect of Ionic Liquids on Protein Crystallization and X-ray Diffraction Resolution, *Cryst. Growth Des.*, 2009, **9**, 3463–3469.

7 N. Debeljuh, C. J. Barrow, L. Henderson and N. Byrne, Structure inducing ionic liquids—enhancement of alpha helicity in the Abeta (1--40) peptide from Alzheimer's disease, *Chem. Commun.* , 2011, **47**, 6371–6373.

8 K. S. Egorova, E. G. Gordeev and V. P. Ananikov, Biological Activity of Ionic Liquids and Their Application in Pharmaceutics and Medicine, *Chem. Rev.*, 2017, **117**, 7132–7189.

9 S. P. M. Ventura, F. A. E Silva, M. V. Quental, D. Mondal, M. G. Freire and J. A. P. Coutinho, Ionic-Liquid-Mediated Extraction and Separation Processes for Bioactive Compounds: Past, Present, and Future Trends, *Chem. Rev.*, 2017, **117**, 6984–7052.

10 E. M. Kohn, J. Y. Lee, A. Calabro, T. D. Vaden and G. A. Caputo, Heme Dissociation from Myoglobin in the Presence of the Zwitterionic Detergent N,N-Dimethyl-N-Dodecylglycine Betaine: Effects of Ionic Liquids, *Biomolecules*, , DOI:10.3390/biom8040126.

11 M. K. Sorenson, S. S. Ray and S. A. Darst, Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation, *Mol. Cell*, 2004, **14**, 127–138.

12 J. L. Huang, M. E. Noss, K. M. Schmidt, L. Murray and M. R. Bunagan, The effect of neat ionic liquid on the folding of short peptides, *Chem. Commun.* , 2011, **47**, 8007–8009.

13 D. Constantinescu, H. Weingärtner and C. Herrmann, Protein denaturation by ionic liquids and the Hofmeister series: a case study of aqueous solutions of ribonuclease A, *Angew. Chem. Int. Ed Engl.*, 2007, **46**, 8887–8889.

14 J. L. Baker, J. Furbish and G. E. Lindberg, Influence of the ionic liquid [C4mpy][Tf2N] on the structure of the miniprotein Trp-cage, *J. Mol. Graph. Model.*, 2015, **62**, 202–212.

15 A. J. Heyert, S. L. Knox, G. E. Lindberg and J. L. Baker, Influence of an ionic liquid on the conformational sampling of Xaa-Pro dipeptides, *J. Mol. Liq.*, 2017, **227**, 66–75.

16 J. Smiatek, Aqueous ionic liquids and their effects on protein structures: an overview on recent theoretical and experimental results, *J. Phys. Condens. Matter*, 2017, **29**, 233001.

17 E. A. Oprzeska-Zingrebe and J. Smiatek, Aqueous ionic liquids in comparison with standard co-solutes : Differences and common principles in their interaction with protein and DNA structures, *Biophys. Rev.*, 2018, **10**, 809–824.

18 K. A. Sajeevan and D. Roy, Temperature-dependent molecular dynamics study reveals an ionic liquid induced 310 - to α-helical switch in a neurotoxin, *Biopolymers*, 2017, **108**, e23009.

19 V. W. Jaeger and J. Pfaendtner, Structure, dynamics, and activity of xylanase solvated in binary mixtures of ionic liquid and water, *ACS Chem. Biol.*, 2013, **8**, 1179–1186.

20 V. Jaeger, P. Burney and J. Pfaendtner, Comparison of three ionic liquid-tolerant cellulases by molecular dynamics, *Biophys. J.*, 2015, **108**, 880–892.

21 P. R. Burney, E. M. Nordwald, K. Hickman, J. L. Kaar and J. Pfaendtner, Molecular dynamics

investigation of the ionic liquid/enzyme interface: application to engineering enzyme surface charge, *Proteins*, 2015, **83**, 670–680.
22 G. Deckert, P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen and R. V. Swanson, The complete genome of the hyperthermophilic bacterium Aquifex aeolicus, *Nature*, 1998, **392**, 353–358.
23 R. G. Molloy, W. K. Ma, A. C. Allen, K. Greenwood, L. Bryan, R. Sacora, L. Williams and M. J. Gage, Aquifex aeolicus FlgM protein exhibits a temperature-dependent disordered nature, *Biochim. Biophys. Acta*, 2010, **1804**, 1457–1466.
24 W. K. Ma, R. Hendrix, C. Stewart, E. V. Campbell, M. Lavarias, K. Morris, S. Nichol and M. J. Gage, FlgM proteins from different bacteria exhibit different structural characteristics, *Biochim. Biophys. Acta*, 2013, **1834**, 808–816.
25 J. Wang, Y. Yang, Z. Cao, Z. Li, H. Zhao and Y. Zhou, The role of semidisorder in temperature adaptation of bacterial FlgM proteins, *Biophys. J.*, 2013, **105**, 2598–2605.
26 K. Fredrick and J. D. Helmann, FlgM is a primary regulator of sigmaD activity, and its absence restores motility to a sinR mutant, *J. Bacteriol.*, 1996, **178**, 7010–7013.
27 Y.-F. Zhu, Z.-X. Cao, L.-L. Zhao and J.-H. Wang, Conformational Change Characteristics in the Intrinsically Disordered FlgM Protein from a Hyperthermophile at Two Different Temperatures, *Acta Phys. Chim. Sin.*, 2015, **31**, 384–392.
28 A. Razvi and J. M. Scholtz, Lessons in stability from thermophilic proteins, *Protein Sci.*, 2006, **15**, 1569–1578.
29 V. N. Uversky, What does it mean to be natively unfolded?, *Eur. J. Biochem.*, 2002, **269**, 2–12.
30 A. Sali and T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.*, 1993, **234**, 779–815.
31 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, PACKMOL: a package for building initial configurations for molecular dynamics simulations, *J. Comput. Chem.*, 2009, **30**, 2157–2164.
32 D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York and P.A. Kollman, *Amber*, University of California, San Francisco, 2015.
33 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. Hauser and C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theory Comput.*, **0**, null.
34 I. S. Joung and T. E. Cheatham 3rd, Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations, *J. Phys. Chem. B*, 2008, **112**, 9020–9041.
35 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**, 926–935.
36 H. Xing, X. Zhao, Q. Yang, B. Su, Z. Bao, Y. Yang and Q. Ren, Molecular Dynamics Simulation Study on the Absorption of Ethylene and Acetylene in Ionic Liquids, *Ind. Eng. Chem. Res.*, 2013, **52**, 9308–9316.
37 L. I. N. Tomé, M. Jorge, J. R. B. Gomes and J. A. P. Coutinho, Molecular dynamics simulation studies of the interactions between ionic liquids and amino acids in aqueous solution, *J. Phys. Chem. B*, 2012, **116**, 1831–1842.
38 K. G. Sprenger, V. W. Jaeger and J. Pfaendtner, The general AMBER force field (GAFF) can accurately predict thermodynamic and transport properties of many ionic liquids, *J. Phys. Chem. B*, 2015, **119**, 5882–5895.
39 K. Palunas, K. G. Sprenger, T. Weidner and J. Pfaendtner, Effect of an ionic liquid/air Interface on the structure and dynamics of amphiphilic peptides, *J. Mol. Liq.*, 2017, **236**, 404–413.
40 K. G. Sprenger, J. G. Plaks, J. L. Kaar and J. Pfaendtner, Elucidating sequence and solvent specific

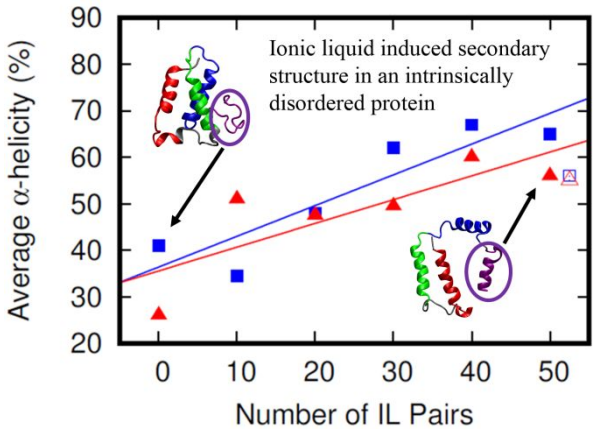
452 design targets to protect and stabilize enzymes for biocatalysis in ionic liquids, *Phys. Chem. Chem.*
453 *Phys.*, 2017, **19**, 17426–17433.
454 41 M. G. Freire, C. M. S. S. Neves, S. P. M. Ventura, M. J. Pratas, I. M. Marrucho, O. João, J. A. P.
455 Coutinho and A. M. Fernandes, Solubility of non-aromatic ionic liquids in water and correlation
456 using a QSPR approach, *Fluid Phase Equilib.*, 2010, **294**, 234–240.
457 42 J.-P. Ryckaert, R. Jean-Paul, C. Giovanni and H. J. C. Berendsen, Numerical integration of the
458 cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J.*
459 *Comput. Phys.*, 1977, **23**, 327–341.
460 43 R. J. Loncharich, B. R. Brooks and R. W. Pastor, Langevin dynamics of peptides: The frictional
461 dependence of isomerization rates of N-acetylalanyl-N'-methylamide, *Biopolymers*, 1992, **32**, 523–
462 535.
463 44 D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke,
464 T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S.
465 Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H.
466 Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L.
467 Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York and P.A.
468 Kollman, *Amber*, University of California, San Francisco, 2015.
469 45 T. Darden, D. Tom, Y. Darrin and P. Lee, Particle mesh Ewald: An N·log(N) method for Ewald
470 sums in large systems, *J. Chem. Phys.*, 1993, **98**, 10089.
471 46 W. Humphrey, A. Dalke and K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.*, 1996,
472 **14**, 33–8, 27–8.
473 47 D. R. Roe and T. E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of
474 Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
475 48 D.A. Case, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W.
476 Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu,
477 T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev,
478 F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails,
479 R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman, *AMBER 2017*,
480 University of California, San Francisco, 2017.
481 49 Y. Babuji, A. Brizius, K. Chard, I. Foster, D. S. Katz, M. Wilde and J. Wozniak, *Introducing Parsl:*
482 *A Python Parallel Scripting Library*, 2017.
483 50 F. Mustan, A. Ivanova, G. Madjarova, S. Tcholakova and N. Denkov, Molecular Dynamics
484 Simulation of the Aggregation Patterns in Aqueous Solutions of Bile Salts at Physiological
485 Conditions, *J. Phys. Chem. B*, 2015, **119**, 15631–15643.
486 51 J. S. Yoneda, A. J. Miles, A. P. U. Araujo and B. A. Wallace, Differential dehydration effects on
487 globular proteins and intrinsically disordered proteins during film formation, *Protein Sci.*, 2017, **26**,
488 718–726.
489 52 M. M. Dedmon, C. N. Patel, G. B. Young and G. J. Pielak, FlgM gains structure in living cells, *Proc.*
490 *Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12681–12684.
491 53 J. Tyrrell, K. M. Weeks and G. J. Pielak, Challenge of mimicking the influences of the cellular
492 environment on RNA structure by PEG-induced macromolecular crowding, *Biochemistry*, 2015, **54**,
493 6447–6453.

494

495

496

497 **Graphical abstract:**



498

499 Text highlighting the novelty of the work:

500 The ionic liquid 1-butyl-1-methylpyrrolidinium bis(trifluoromethylsulfonyl)imide is shown to induce
501 secondary structure similar to a bioactive state in the protein FlgM.