Transient Secondary and Tertiary Structure Formation Kinetics in the Intrinsically Disordered State of α-Synuclein from Atomistic Simulations

Timo Graen+,[a] Reinhard Klement+,[a] Asaf Grupi,[b] Elisha Haas,[c] and Helmut Grubmüller*[a]

In the absence of a stable fold, transient secondary structure kinetics define the native state of the prototypical and pharmacologically relevant intrinsically disordered protein (IDP) α-Synuclein (aS). Here, we investigate kinetics preventing ordering and possibly pathogenic β-sheet aggregation. Interestingly, transient β-sheets form frequently at sub µs time scales precisely at the positions observed in aS amyloid fibrils. The formation kinetics competes with rapid secondary structure dissociation rates, thus explaining the low secondary structure content. The fast secondary structure dissociation times are very similar to the dynamics of tertiary structure rearrangements. These findings suggest that the fast dissociation kinetics slows down conformational selection processes for aS aggregation, which may be a general mechanism controlling the aggregation kinetics of IDPs.

1. Introduction

Aggregation of aS in the human brain is linked to the onset of Parkinson's disease.[1] Time scales of transient secondary structure formations and tertiary structure reorganization of free aS monomers prior to aggregation[2] are still unknown. In particular, the interaction of the charged termini in the disordered state has been suggested to prevent permanent secondary structure formation in the fibril core region (residues 61–95),[3] which was suggested[4] to already transiently form β-sheets with unknown rates.

These intrinsic time scales are not only most relevant for the association of aS monomers into low molecular weight oligomers and further for the aggregation kinetics of fibril formation, but also for the interpretation of single molecule FRET experiments.[5,6] Specifically, the relation between structure reorganization (ns–µs), dye dynamics (ps–ns), dye excitation decay (ns) and inter photon correlation times (ms)[7] is largely unknown, but dictates the quantitative interpretation of FRET experiments.

Here, we investigated how transient β-sheet forming regions in the aS monomeric ensemble compare to the known structure forming regions in aggregated aS fibrils and analyzed the β-sheet formation k_{on} and dissociation rates k_{off} to quantify the kinetic differences between the structure forming NAC and

N-terminus and the largely unstructured C-terminus. To this end, we performed unrestraint atomistic MD simulations in explicit solvent and extracted aS ensembles from a set of 20×3 µs wild type aS trajectories using the AMBER03ws/TIP4P2005 s force field.

2. Results and Discussion

Figure 1A depicts the β-sheet population for each individual residue in the simulation ensemble. The majority ($\sim 75\%$) of these occurrences are provided by transiently forming short β-sheets of just 2 to 3 residues in length. The position of β-sheet occurrences are found to be in close agreement with experimental[2,8] β-sheet regions found in aS fibrils, see SI for additional Ramachandran occupations. The observed NAC β-sheet peaks, including the main peak in the NAC region, further agree very well with the previously proposed correlation between transient monomeric kinetics and fibrillated aS.[4] The overall content of α-helices 2% and β-sheets 7.5% in our simulation ensemble compares well to values derived from CD experiments α-helices 2%; β-sheets 11%[9] and FTIR measurements β-sheets 15%[10].

Figure 1B depicts the rates for β-sheet formation k_{on} and Figure 1C the dissociation k_{off} rates. The fastest β-sheet formation rates were observed in the regions between residues 38–44, 47–52, 63–66, 69–74 and 92–96 with five strong peaks at residues 39, 48, 65, 70 and 95. At these positions, β-sheets formed with rates $k_{on} = 1.2–4.5\,\mu s^{-1}$ and dissociated with rates of $k_{off} = 2.3–3.9\,\mu s^{-1}$, see SI for detailed rates. Formation rates were consistently slower than dissociation rates with the exception of the β-sheet peak around residue 70. Here, the association rate $k_{on} = 4.5\,\mu s^{-1}$ was faster than the dissociation rate $k_{off} = 2.9\,\mu s^{-1}$. The lowest occurrences of β-sheets are seen for the highly unstructured C-terminal region between residues 100–140. The changes in k_{on} and k_{off} relative to the average

[a] T. Graen,+ R. Klement,+ H. Grubmüller
     Theoretical and Computational Biophysics Department, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany
     E-mail: hgrubmu@gwdg.de
[b] A. Grupi
     Physics Department, Bar-Ilan University, Ramat Gan 52900, Israel
[c] E. Haas
     The Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan 52900, Israel
[+] equal contribution
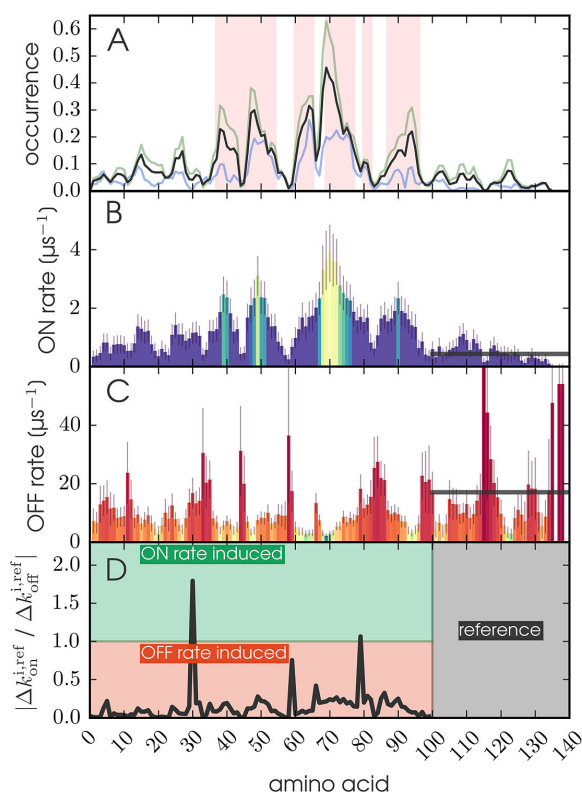   Supporting information for this article is available on the WWW under https://doi.org/10.1002/cphc.201800504

Figure 1. A) Relative occurrence of the secondary structure content for T-REX trajectories (blue), long WT trajectories (green) and weighted with a 40/60 ratio (black). Highlighted background (red) shows aS fibril β-sheet locations[2] (PDB: 2N0A). B) Free monomeric aS β-sheet formation k_{on} and C) β-sheet dissociation rates k_{off} with error bars at 2 σ. Colors equal rates. D) Ratio between $k_{athrmon}$ and k_{off} relative to the unstructured C-terminus (reference). Ratios smaller than 1 indicate structure formation due to changes in the k_{off} rates, ratios larger than 1 indicate formation due to changes in the k_{on}.

Figure 2. A) Residue contact map with N–C contact region (black). B) Mono-exponential amino acid correlation times for short and long range correlations between the transient structure forming regions (top bar) in aS.

rates of this C-terminal region, Figure 1D, suggest that structure formation is caused by changes in the dissociation rates over changes in the formation rates.

We also analyzed the kinetics of α-helices. Whereas the on-rates of $k_{on} = 0.8$–$3.4\ \mu s^{-1}$ were similar to the β-sheet formation, the dissociation rates with $k_{off} = 6.1$–$36\ \mu s^{-1}$ were much faster, see SI.

We finally analyzed the tertiary kinetics and asked whether there are transient tertiary mechanisms beyond the secondary structure forming regions described above. To this aim, Fig-
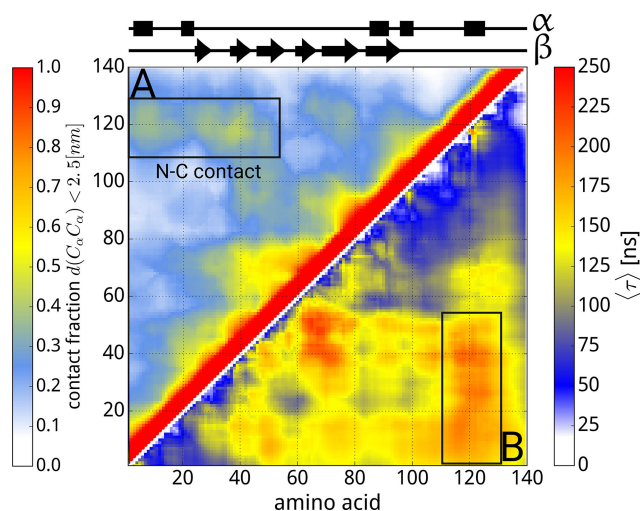
ure 2A (upper left half) depicts the residue contact map, which shows a prominent interaction between the N- and C-terminus, in agreement with previously identified NMR-PRE contact regions.[11] To quantify the time scales at which conformational tertiary rearrangements take place, Figure 2B (lower right half) shows pairwise residue distance correlation times. Interestingly, the correlation times $\tau_{NC} \approx 200$–250 ns of the N–C contact region are very similar to the secondary structure formation rate peak around residue 70 in the fibril core region, Figure 2B, with a formation time of 220 ns but also with its dissociation time of 260 ns. In contrast, the correlation times within the N- and C-termini are very fast 20–100 ns, except for a small region around residue 121. Taken together, the contact map and correlation times are consistent with an aS model that has a structure forming fibril core region between amino acids 35–90 and two unstructured but dynamically interacting,[3] termini between amino acids 1–35 and 90–140, suggesting a possible involvement of the termini in structure formation in the NAC region.

Figure 3 summarizes the above times scales of tertiary and secondary structure formation and dissociation in aS. Secondary structures form and dissociate frequently but no permanent secondary structure persists.
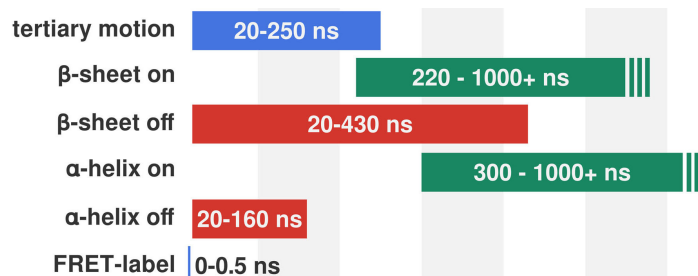

Figure 3. Overview of the observed time scales for tertiary motion and FRET labels (blue) and the secondary structure formation (green) and dissociation (red) time scales from 20×3 μs AMBER03ws/TIP4P2005 s simulation trajectories.

We tentatively attribute this very rich conformational dynamics mainly to markedly faster sub μs dissociation kinetics compared to the unfolded state of folding proteins. Although not too much data is available for the secondary structure kinetics of the unfolded state of folding proteins, e.g. secondary structure motifs of the unfolded ACBP protein have been found to persist tens of μs.[12,13] These differences may in part be due to the use of a different force field, but it seems unlikely that this is the sole cause of difference of almost two orders of magnitude. Indeed, very recent single molecule spectroscopy measurements of aS (*B. Schuler, private communication*) and other unfolded and intrinsically disordered proteins[14–16] show similarly fast reorganization dynamics at 10–50 ns time scale. Further, a similar effect was previously described for urea induced unfolding of a protein, where urea was found to destabilize the folded state not by accelerating unfolding, but rather by slowing down refolding.[17]

The initial step of aS aggregation is the formation of low molecular weight oligomers. Recent studies have shown[18,19] that the assembly of aS monomers produces different types of oligomers. Some of them are more neurotoxic than others which has been attributed to the internal structure of these oligomers being more β-rich. The polymorphic aggregation of aS hints towards a conformational selection mechanism to form toxic oligomers. Therefore, β-sheets are not only a prominent structural feature in aS fibrils but are also desease relevant in low molecular wheigt oligomers. With this the overall secondary and tertiary structure kinetics of aS becomes important for the aggregation process. The average collision time of aS monomers in solution ranges from $\tau_c \sim 1.5 \mu s$ for 0.1 mM protein concentration and decreases to $\tau_c \sim 150$ ns for aS concentrations of 1 mM used in aS aggregation assays. For medium and high aS concentrations secondary structure formation, tertiary rearrangement and average collisions happen in a comparable time regime between 100 ns to 1μs favorable for the formation of β-rich oligomers.

The accuracy of aS simulations and IDPs in general strongly depends on the choice of force field,[20,21] an observation which was previously described also for aS by Piana et al.[22] Therefore, we assessed the simulation accuracy from μs temperature replica exchange simulations using four different force fields by comparison to experimental data from NMR,[23–25] SAXS[24] and FRET[26,27] measurements.

Figure 4 compares radius of gyration (R_G) histograms for the four tested force fields. The AMBER03ws/TIP4P2005 s[21,28,29] (green) ensemble agrees best with experiment[24,30] with a $R_G = 3.0$ nm. The more compact CHARMM22*/TIP4P−D[22,31–33] (magenta) ensemble resulted in a smaller radius of gyration $R_G = 2.2$ nm, whereas AMBER99sb*ildn/TIP3P[34–37] (yellow) and CHARMM22*/TIP3P−C[31–33] (blue) are clearly well outside the range (black lines) set by the experiments. The same trend was observed for the direct comparison with SAXS and NMR-PRE profiles, see SI.

The differences between the aS ensembles are indeed dramatic and underscore the importance of careful force field choice and assessment. Even though there does not seem to be one single 'optimal' force field for all IDPs,[20] the
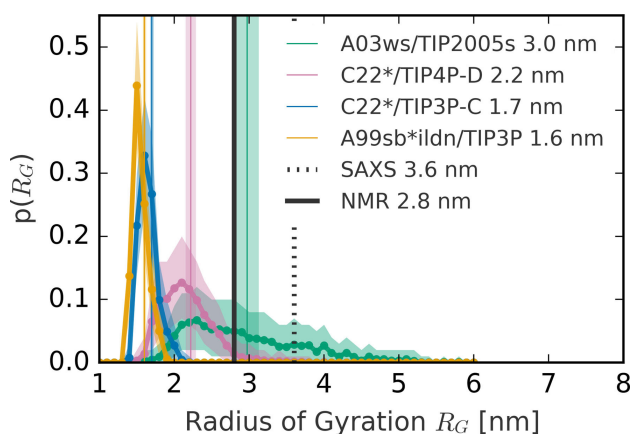

Figure 4. Radius of gyration histograms from replica exchange simulations after WHAM temperature reweighting to 300 K. Averages are shown (colored vertical lines) with their statistical uncertainty at 2 σ (shaded regions).

AMBER03ws/TIP4P2005 s based ensemble closely reproduces the expanded shape of the previous fitted aS ensembles with a $R_G = 3.1$ nm[23,24].

To also assess the long range conformational dynamics, we doubly labeled the AMBER03ws/ TIP4P2005 s ensemble with FRET dyes at eight positions. We compared the distances derived from MD simulations to those predicted from trFRET experiments[26,27] (Figure 5). This comparison depends on an *a priori* unknown orientational dynamics of the dyes.[38]
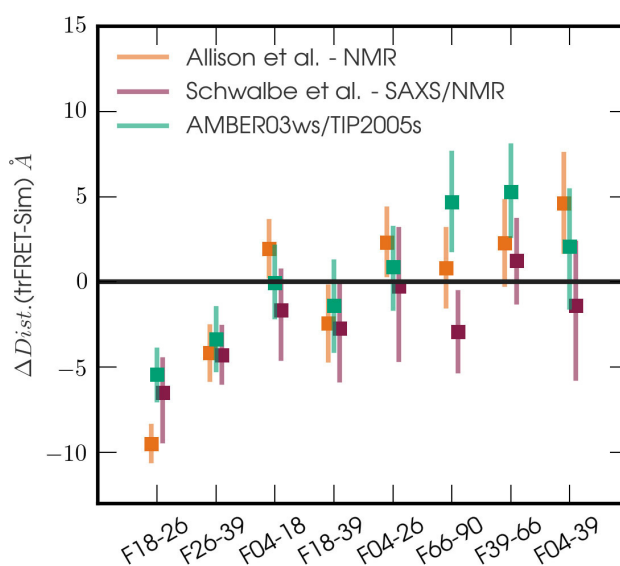

Figure 5. Comparison of the eight trFRET[26,27] mean dye-dye distances to the AMBER03ws/TIP4P2005s MD ensemble, to the NMR-PRE restraint ensemble[40] and to the NMR-SAXS fitted ensemble.[24]

To decide whether the orientational dynamics of the dye is faster (dynamic case) or slower than the FRET transfer process (static case)[7] of about 5 ns, we calculated the correlation time of $\kappa^2(t)$ for the tryptophan/coumarin dye pair. The mean correlation time for the eight label positions was 0.4 ± 0.1 ns, i.e.

much faster than the average decay time. The average $\langle \kappa^2 \rangle = 0.65 \pm 0.02$ was derived following earlier works[39] and was found to be very close to the isotropic $\kappa_{rmiso}^2 = 2/3$ for all tested force fields, despite the very crowded protein matrix.

These two findings strongly suggest that the FRET labels sample isotropic orientations within the lifetime of the FRET process, such that the derived distance distributions should be rather accurate. We attribute this encouraging result to the choice of the particularly small dyes tryptophan and coumarin, which also seem to perturb the system only slightly.

The distances for the two label pair positions F18-26 and F26-39 were found in good agreement between the AMBER03ws/TIP4P2005 s ensemble and the two fitted ensembles,[23,24] whereas the FRET experiments suggest these distances to be shorter. Compared to the other tested force fields, the AMBER03ws/TIP4P2005 s ensemble further also agreed closest with the FRET experiments when directly comparing to the raw fluorescence intensity decay curves from experiment, see SI. Overall, the AMBER03ws/TIP4P2005 s ensemble predicted the trFRET distances as well as the two fitted ensembles.

We next assessed our simulations against the more local properties probed by NMR experiments, in particular against $^3J_{HNH\alpha}$ (Figure 6), $^1J_{C\alpha H\alpha}$, $^1J_{NC\alpha}$, $^2J_{NC\alpha}$ couplings and chemical shifts (SI) for the AMBER03ws ensemble and compared them to NMR experiments.[25] The J-couplings correlate with the experimental reference and in most cases agree within error bars. Despite the large error bars, this agreement supports our simulations, as evidenced by a P-value below 0.4% for such correlation by chance. The analysis of the Ramachandran populations[25] (SI) suggests a slight overestimation of valine β-sheet populations.

Further, comparison to NMR spectral densities[25] suggests that the mobility in this region is somewhat slower in our simulations than in the NMR experiments (SI), although the main result of an off-rate dominated β-sheet population is unchanged.

3. Conclusions

Overall, we have demonstrated that without any aS specific experimental input (except using 2 bits for force field selection), our unrestraint MD simulations predicted both its expanded shape from SAXS experiments and trFRET distances as well as previous static fitted ensembles. The AMBER03ws ensemble shows a good correlation with NMR J-couplings. Not unexpectedly, the deviation for the secondary chemical shifts is somewhat larger than for the previous ensemble which was fitted against NMR data.[24]

We observed that short lived β-sheets frequently formed in regions of slower dissociation rates precisely at the fibril forming NAC region, supporting a conformational selection mechanism to form toxic oligomers. The tertiary 250 ns N–C contact kinetics occur on a similar time scale as the secondary structure kinetics of the central NAC region around residue 70. This result suggests that either tertiary structure changes govern secondary formation, e.g., via changed hydrophobicities, or that the presence of secondary structure elements is required for and triggers the formation of tertiary structure, e.g., in a diffusion-collision scenario, or both. It is conceivable that such coupling of structural dynamics, too, will be observed
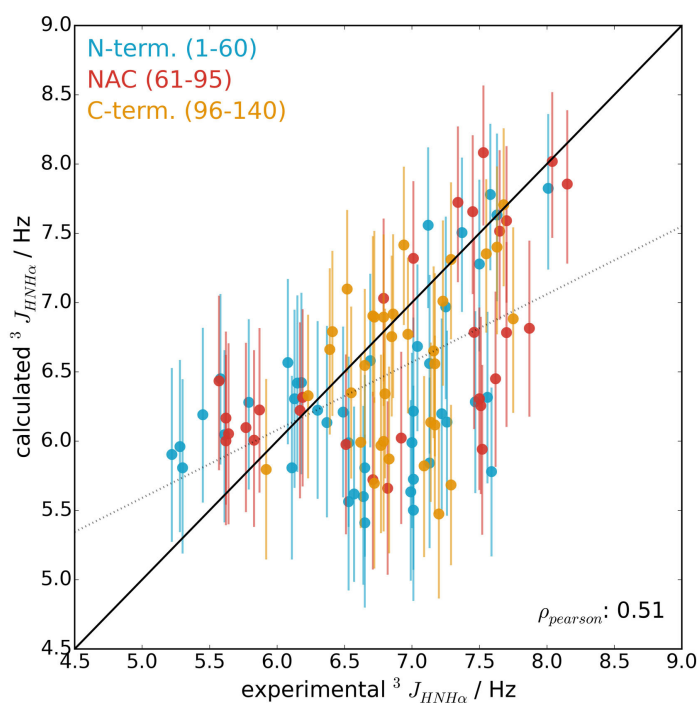

Figure 6. Comparison of $^3J_{HNH\alpha}$ couplings between calculated values from 20×3 μs AMBER03ws MD simulations and experimental values of Mantsyzov et al.[25]. The optimal linear fit (dashed line), N-terminus (blue dots), NAC region (red dots) and C-terminus (yellow dots) are highlighted. Error bars (colored lines).

and is functionally relevant in many other intrinsically disorder-edproteins.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] M. Goedert, *Nat. Rev. Neurosci.* **2001**, *2*, 492–501.
[2] M. D. Tuttle, G. Comellas, A. J. Nieuwkoop, D. J. Covell, D. A. Berthold, K. D. Kloepper, J. M. Courtney, J. K. Kim, A. M. Barclay, A. Kendall, *Nat. Struct. Mol. Biol.* **2016**.
[3] M.-K. Cho, G. Nodet, H.-Y. Kim, M. R. Jensen, P. Bernado, C. O. Fernandez, S. Becker, M. Blackledge, M. Zweckstetter, *Prot. Science* **2009**, *18*, 1840–1846.
[4] H.-Y. Kim, H. Heise, C. O. Fernandez, M. Baldus, M. Zweckstetter, *ChemBioChem* **2007**, *8*, 1671–1674.
[5] A. J. Trexler, E. Rhoades, *Biochemistry* **2009**, *48*, 2304–2306.
[6] A. J. Trexler, E. Rhoades, *Biophys. J.* **2010**, *99*, 3048–3055.
[7] I. V. Gopich, A. Szabo, *PNAS* **2012**, *109*, 7747–7752.
[8] H. Shaykhalishahi, A. Gauhar, M. M. Wördehoff, C. S. Grüning, A. N. Klein, O. Bannach, M. Stoldt, D. Willbold, T. Härd, W. Hoyer, *Angewandte Int. Ed.* **2015**, *54*, 8837–8840.
[9] A. Rekas, R. B. Knott, A. Sokolova, K. J. Barnham, K. A. Perez, C. L. Masters, S. C. Drew, R. Cappai, C. C. Curtain, C. L. L. Pham, *Eur. Biophys. J.* **2010**, *39*, 1407–1419.
[10] V. N. Uversky, J. Li, L. Fink, *J. Biol. Chem.* **2001**, *276*, 10737–10744.
[11] S. Esteban-Martʹın, J. Silvestre-Ryan, C. W. Bertoncini, X. Salvatella, *Biophys. J.* **2013**, *105*, 1192–1198.
[12] K. Lindorff-Larsen, N. Trbovic, P. Maragakis, S. Piana, D. E. Shaw, *JACS* **2012**, *134*, 3787–3791.
[13] V. A. Voelz, M. Jäger, S. Yao, Y. Chen, L. Zhu, S. A. Waldauer, G. R. Bowman, M. Friedrichs, O. Bakajin, L. J. Lapidus, *JACS* **2012**, *134*, 12565–12577.
[14] A. Soranno, B. Buchli, D. Nettels, R. R. Cheng, S. Muller-Spath, S. H. Pfeil, A. Hoffmann, E. A. Lipman, D. E. Makarov, B. Schuler, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17800–17806.
[15] B. Schuler, A. Soranno, H. Hofmann, D. Nettels, *Annu. Rev. Biophys.* **2016**, *45*, 207–231.
[16] A. Soranno, A. Holla, F. Dingfelder, D. Nettels, D. E. Makarov, B. Schuler, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1833–E1839.
[17] M. C. Stumpe, H. Grubmu"ller, *Biophys. J.* **2009**, *96*, 3744–3752.
[18] N. Cremades, S. I. Cohen, E. Deas, A. Y. Abramov, A. Y. Chen, A. Orte, M. Sandal, R. W. Clarke, P. Dunne, F. A. Aprile, C. W. Bertoncini, N. W. Wood, T. P. Knowles, C. M. Dobson, D. Klenerman, *Cell* **2012**, *149*, 1048–1059.
[19] J. A. Fauerbach, T. M. Jovin, *Eur. Biophys. J.* **2017**.
[20] S. Rauscher, V. Gapsys, M. J. Gajda, M. Zweckstetter, B. L. de Groot, H. Grubmu"ller, *JCTC* **2015**, *11*, 5513–5524.
[21] R. B. Best, W. Zheng, J. Mittal, *JCTC* **2014**, *10*, 5113–5124.
[22] S. Piana, A. G. Donchev, P. Robustelli, D. E. Shaw, *J. Phys. Chem. B* **2015**, *119*, 5113–5123.
[23] J. R. Allison, P. Varnai, C. M. Dobson, M. Vendruscolo, *JACS* **2009**, *131*, 18314–18326.
[24] M. Schwalbe, V. Ozenne, S. Bibow, M. Jaremko, L. Jaremko, M. Gajda, M. R. Jensen, J. Biernat, S. Becker, E. Mandelkow, M. Zweckstetter, M. Blackledge, *Structure* **2014**, *22*, 238–249.
[25] A. B. Mantsyzov, A. S. Maltsev, J. Ying, Y. Shen, G. Hummer, A. Bax, *Protein Sci.* **2014**, *23*, 1275–1290.
[26] A. Grupi, E. Haas, *J. Mol. Biol.* **2011**, *405*, 1267–1283.
[27] A. Grupi, E. Haas, *J. Mol. Biol.* **2011**, *411*, 234–247.
[28] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, *J. Comput. Chem.* **2003**, *24*, 1999–2012.
[29] J. L. Abascal, C. Vega, *J. Chem. Phys.* **2005**, *123*, 234505.
[30] A. S. Morar, A. Olteanu, G. B. Young, G. J. Pielak, *Prot. Sci.* **2001**, *10*, 2195–2199.
[31] S. Piana, K. Lindorff-Larsen, D. E. Shaw, *Biophys. J.* **2011**, *100*, L47–L49.
[32] A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
[33] A. D. MacKerell, M. Feig, C. L. Brooks, *J. Comput. Chem.* **2004**, *25*, 1400–1415.
[34] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Bioinf.* **2006**, *65*, 712–725.
[35] R. B. Best, G. Hummer, *J. Phys. Chem. B* **2009**, *113*, 9004–9015.
[36] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins Struct. Funct. Bioinf.* **2010**, *78*, 1950–1958.
[37] M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **2000**, *112*, 8910–8922.
[38] E. Haas, E. Katchalski-Katzir, I. Z. Steinberg, *Biopolymers* **1978**, *17*, 11–31.
[39] M. Hoefling, N. Lima, D. Haenni, C. A. Seidel, B. Schuler, H. Grubmüller, *PLoS One* **2011**, *6*, e19791.
[40] C. W. Bertoncini, Y.-S. Jung, C. O. Fernandez, W. Hoyer, C. Griesinger, T. M. Jovin, M. Zweckstetter, *PNAS* **2005**, *102*, 1430–1435.