

COMMUNICATION

Binding Modes of Thioflavin T and Congo Red to the Fibril Structure of Amyloid- β (1–42)

Received 00th January 20xx,
Accepted 00th January 20xx

Benedikt Frieg^{a,b}, Lothar Gremer^{b,c}, Henrike, Heise^{b,c}, Dieter Willbold^{b,c}, Holger Gohlke^{a,b,d*}

DOI: 10.1039/x0xx00000x

Binding modes for the amyloid- β (1–42) fibril fluorescent dyes Thioflavin T and Congo red were predicted by molecular dynamics simulations and binding free energy calculations. Both probes bind on the fibril surface to primarily hydrophobic grooves, with their long axis oriented almost parallel to the fibril axis. The computed binding affinities are in agreement with experimental values. The binding modes also explain observables from previous structural studies and, thus, provide a starting point for the systematic search and design of novel molecules, which may improve *in vitro* diagnostics for Alzheimer's disease.

Alzheimer's disease (AD) is a progressive, unremitting, neurodegenerative disorder^{1–4}. In 2018 it was assumed that more than 50 Mio. people worldwide are living with dementia, with two-thirds associated with AD. Recent estimates assume that by 2030 more than 80 Mio. people will be diagnosed with dementia^{5,6}. The considerable number of dementia patients will incur high costs for the healthcare system, which are predicted to rise to worldwide \$2 trillion by 2030⁶. Thus, dementia and AD, in particular, constitute serious challenges for the modern healthcare system.

AD is related to the production-elimination-imbalance of the peptide amyloid-beta ($A\beta$)⁷. At high concentration, $A\beta$ peptides tend to aggregate to oligomers or arrange to symmetric, periodic fibrils, both denoting pathological hallmarks in AD^{4,7–9}. One diagnostic strategy focuses on detecting biomarkers for

AD^{5,6}, such as the accumulation of $A\beta$ fibrils^{5,6}. As changes in the brain begin 20 years or more before AD-related symptoms are expected to occur^{5,6}, a conclusive test will be essential. This is all the more true as a therapeutic intervention in the very early stages of AD is currently considered necessary to stop its progression^{5,6}. Fluorescent dyes that detect $A\beta$ fibrils, for example, Thioflavin T (THT) and Congo red (CR)^{10,11} (Figure S1), are thus essential for an accurate and conclusive diagnosis.

THT and CR form fluorescent complexes with amyloid and amyloid-like fibrils^{12–14}. For both probes, previous studies suggested binding modes^{10,11,15}, but the exact nature of how the probes bind to the $A\beta$ (1–42) fibril remained elusive. Such knowledge is, however, essential for the systematic search for novel molecular probes^{10,11}. The recent success in structure determination of the $A\beta$ (1–42) fibril¹⁶ offers the possibility to predict the binding modes of THT and CR to that fibril.

We applied all-atom molecular dynamics (MD) simulations of free THT and CR diffusion in the presence of the $A\beta$ (1–42) fibril of an aggregate time of 45 μ s to derive such binding modes. The simulations were not biased by any prior information on potential binding epitopes. Over simulation times of 1 μ s, THT repeatedly binds to and unbinds from the fibril (Figure S2, Movie S1). CR, by contrast, does not unbind completely but moves across the fibril surface once bound (Figure S3, Movie S2), suggesting that CR binds with higher affinity.

To identify the most likely binding epitopes, we computed probability densities of the location of either dye. THT predominantly binds in the vicinity of V18, F20, and E22 on either protofilament (Figure 1A). For CR, by contrast, multiple areas are identified across the fibril surface (Figure 2A). Likely, this is because CR can form hydrophobic, ionic, and hydrogen bond interactions, enabling binding to different sites. The vicinity of V18, F20, and E22, however, is weakly occupied by CR, suggesting that both probes bind to distinct epitopes. Stably bound conformations (Figure S4) were clustered to extract the most relevant binding poses. The ten largest clusters were further subjected to molecular mechanics generalized Born surface area (MM-GBSA)¹⁷ and normal mode analysis (NMA)¹⁸

^a John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich GmbH, Jülich, Germany

^b Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich GmbH, Jülich, Germany.

^c Institut für Physikalische Biologie, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

^d Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

* Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: The Supporting Information document contains detailed instructions for system preparation, simulations, and analysis procedures, and results from RMSD measurements, clustering, and binding affinity computations, additional binding poses, and force field parameters for CR. The Supplemental Movies show representative MD trajectories of THT (Movie S1) and CR (Movie S2) binding.]. See DOI: 10.1039/x0xx00000x

calculation, yielding ΔG_{bind}^0 (eq. S1 - S3). Although the MM-GBSA method is established as a fast and reliable method to estimate effective binding energies, considering the entropic contributions by NMA may introduce some uncertainty^{19, 20}. In the present study, however, the MM-GBSA and NMA calculations were performed on pre-clustered sets of dye poses, such that the total error (eq. S6) for most THT cluster is $< 1.0 \text{ kcal mol}^{-1}$ and for CR cluster is $< 1.5 \text{ kcal mol}^{-1}$ (Table S1, S2) and, thus, comparable to the uncertainty of more rigorous, but computationally less efficient methods²¹. From ΔG_{bind}^0 we derived the dissociation constant $K_{\text{D}}^{\text{comp}}$ (eq. S4) for each cluster. Focusing on the interpretation of $K_{\text{D}}^{\text{comp}}$ and its upper and lower limits, we derive binding mode models that are in agreement with previous observables.

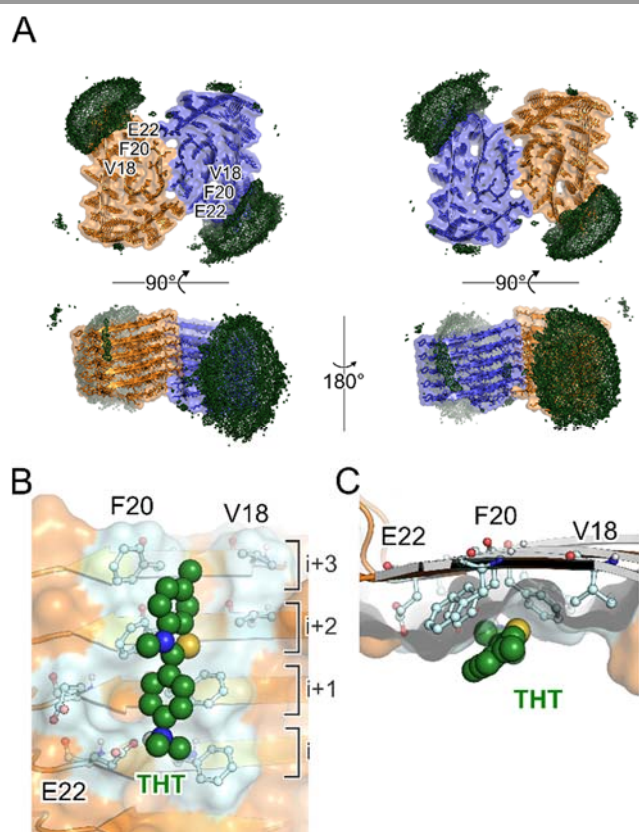


Figure 1: A: 3D density grids representing the probability density of the Thioflavin-T (THT) position around the Aβ(1-42) fibril (blue-orange cartoon-surface representation). B, C: Predominant THT binding pose (sphere representation) from the c0 cluster in side (B)- and top-view (C), with amino acids (cyan ball-stick model) binding THT.

Experimental binding affinities ($K_{\text{D}}^{\text{exp}}$) of THT to Aβ(1-42) fibrils range from 294 - 4000 nM^{22, 23}. We identified three THT clusters the configurations of which yield $K_{\text{D}}^{\text{comp}}$ in excellent to very good agreement with $K_{\text{D}}^{\text{exp}}$ (Table S1). The c0 cluster results in the most favorable affinity ($\Delta G_{\text{bind}}^0 = -9.06 \pm 0.42 \text{ kcal mol}^{-1}$, $K_{\text{D}}^{\text{comp}} = 113 - 470 \text{ nM}$, Table S1), followed by c3 ($\Delta G_{\text{bind}}^0 = -8.37 \pm 0.52 \text{ kcal mol}^{-1}$, $K_{\text{D}}^{\text{comp}} = 305 - 1759 \text{ nM}$, Tab. S1), and c1 ($\Delta G_{\text{bind}}^0 = -8.14 \pm 0.45 \text{ kcal mol}^{-1}$, $K_{\text{D}}^{\text{comp}} = 506 - 2331 \text{ nM}$, Tab. S1). The intramolecular mobility of THT is reduced in the bound conformation (Figure S5), which is a structural prerequisite for a dye to fluoresce²⁴. THT configurations of these clusters bind to V18, F20, and E22, with THT's protonated amino function

forming a salt bridge with E22, and F20 and V18 accommodating the aromatic moieties (Figure 1B,C, Figure S6). In doing so, the side chains of F20 adopt a V-shaped orientation, forming π - π -stacking interactions with the aromatic rings in THT. THT binds across multiple Aβ(1-42) peptides with its axis oriented almost perpendicularly to the orientation of the stacked β -strands. The observation provides an explanation of why THT recognizes fibrillary structures but no single Aβ(1-42) peptides^{25, 26}.

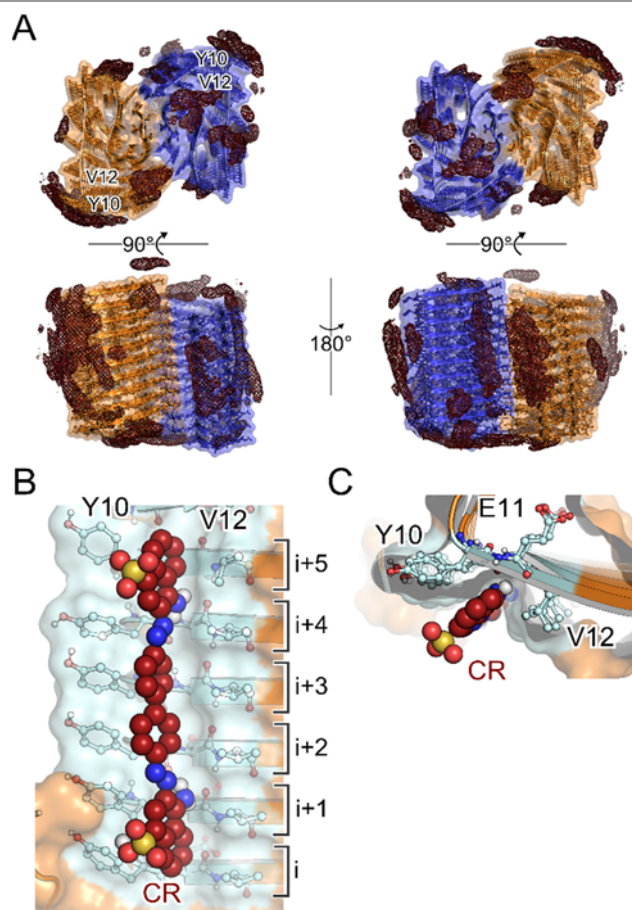


Figure 2: 3D density grids representing the probability density of the Congo red (CR) positions around the Aβ(1-42) fibril (blue-orange cartoon-surface representation). B, C: Predominant CR binding pose (sphere representation) from the c3 cluster in side (B)- and top-view (C), with amino acids (cyan ball-stick model) binding CR.

$K_{\text{D}}^{\text{exp}}$ values of CR to Aβ(1-42) fibrils range from 209 - 3070 nM²⁷. We identified two CR clusters the configurations of which yield $K_{\text{D}}^{\text{comp}}$ in excellent agreement with experiments (Table S2). The c3 cluster results in the more favorable binding affinity ($\Delta G_{\text{bind}}^0 = -11.07 \pm 1.31 \text{ kcal mol}^{-1}$, $K_{\text{D}}^{\text{comp}} = 0.84 - 70 \text{ nM}$, Tab. S2) than the c7 cluster ($\Delta G_{\text{bind}}^0 = -9.78 \pm 1.82 \text{ kcal mol}^{-1}$, $K_{\text{D}}^{\text{comp}} = 3 - 1458 \text{ nM}$, Tab. S2). The intramolecular mobility in the bound conformation is reduced markedly in the central biphenyl group and to some extent across the diazen moiety (Figure S7). CR binds to the groove between Y10 and V12, forming edge-to-face aromatic and hydrophobic interactions between its almost planar biaryl core with Y10 and V12, respectively. CR also forms hydrogen bonds with the backbone carbonyl oxygen of E11, while the sulfonic acid and sulfonate groups are solvent-exposed (Figure 2B,C, Figure S8A). CR binds across six layers of

A β (1-42) peptides with its axis oriented almost perpendicularly to the orientation of the stacked β -strands of the A β fibril. In one simulation, CR intercalates into the protofilament (Figure S8B), with $\Delta G_{\text{bind}}^0 = -14.38 \pm 1.95 \text{ kcal mol}^{-1}$ and $K_{\text{D}}^{\text{comp}} = 0.001 - 0.79 \text{ nM}$ (Table S2). However, this pose was only observed in one of 45 simulations, suggesting that it is the consequence of a kinetic trapping rather than a thermodynamically favorable binding.

The reported binding epitopes and modes for THT and CR agree with experimental observables on multiple accounts. First, the differences in the occurrence frequencies of THT in the vicinity of amino acids V18, F29, and E22 (Figure 1) and CR across the entire surface of the A β (1-42) fibril (Figure 2) are in line with *in vitro* data suggesting that THT-type ligands and CR-type ligands bind to distinct, non-overlapping sites on A β aggregates: one type could not be replaced by the other when bound to A β aggregates²³. Second, the binding stoichiometry of CR to insulin fibrils is ~ 20 -fold higher than of THT, suggesting that there are likely more binding sites for CR¹¹, which agrees with our observation that CR tends to bind to the entire fibril surface. By contrast, in early studies on CR binding to amyloid fibrils, a single binding site was identified only²⁸. Recent high-resolution structures of amyloid fibrils^{16, 29-32} revealed highly symmetric and periodic complexes, however, supporting the view of multiple, identical binding sites. Luminescent-conjugated oligothiophene biomarkers were also shown to bind concurrently to the fibril³³.

THT is predicted to bind to V18, F20, and E22 on the fibril surface. The fast kinetics of THT binding to fibrils suggests that THT's binding site is easily accessible³⁴, which holds true for the predicted site. Studies also suggest a minimal THT binding site on the fibril surface that covers four consecutive β -strands^{25, 26}, a feature we also identified (Figure 1B,C). THT binds to A β fibrils but does not bind to monomers or other oligomeric states³⁴, supporting the view that multiple β -strands are involved in THT binding. In the case of amyloid spherulites, THT binds with its long axis almost in parallel to the long axis of the spherulites³⁵. This orientation was also observed in an X-ray structure of THT attached to a PSAM ladder²⁵, during MD simulations of THT binding to a protofilament¹⁵, and herein (Figure 1B,C). Finally, previous studies suggested that THT binds to hydrophobic channels formed by adjacent side chains along the long axis of amyloid fibrils³⁵. By contrast, the X-ray structure of THT attached to a PSAM ladder suggested that THT binds on surfaces formed by tyrosine residues²⁵, which align in a V-shaped orientation²⁵, similar to the orientation of F20 found here (Figure 1B,C). In contrast to a previous study suggesting multiple binding sites for THT to A β (1-42) with different binding affinities³⁶, during our MD simulations, THT binds exclusively to one site with similar affinity. However, no structural information is available that characterizes the previously proposed binding sites³⁶. As the structure of A β (1-42) fibrils strongly depends on the experimental conditions^{16, 37, 38}, the deviation with the previous study may arise from a different fibril structure¹⁶. Note that our binding model still allows the binding of multiple THT molecules, as the number of binding sites increases with fibril size.

CR is predicted to bind to a hydrophobic channel formed by Y10, E11, and V12. Two structural features are essential for CR binding to A β aggregates: first, two negative charges separated by $\sim 19 \text{ \AA}$ ¹⁰ and, second, a biaryl core framework¹⁰. Modifying the gap between the negative charges reduces the binding affinity¹⁰, but modifying the substituents of the biaryl core barely influences the binding affinity^{10, 39}. The spacing of $\sim 19 \text{ \AA}$ corresponds to the spacing of five pleated and stacked β -strands^{11, 16}, suggesting that CR binds across multiple layers, thereby forming ionic interactions. CR was shown to bind to the surface of stacked β -strands, that way creating ionic interactions between the sulfonic groups of CR and lysine residues of the HET fibril⁴⁰. As the CR molecule considered in our simulations, however, does not carry two negatively charged groups (Figure S1B), it is not surprising that we did not observe the ionic interaction model. Still, we find that CR binds to the surface of stacked β -strands (Figure 2B,C). In our case, however, CR forms mainly hydrophobic contacts, which may be explained by a suggested second binding site for CR-type ligands¹⁰, for which the importance of hydrophobic interactions for CR binding has been stressed^{11, 15, 41}. By contrast, some experimental results cannot be explained by the ionic model¹⁰. Interestingly, for CR derivatives where no ionic state is expected at physiological pH, the binding affinity increased relative to CR¹⁰, supporting the view that hydrophobic interactions are crucial for CR binding. The affinity of CR-like probes also decreases significantly after replacing the amino by hydroxyl groups¹⁰, which is surprising at first sight, as amino and hydroxyl groups both are hydrogen bond acceptor and donor groups. As observed for other azo dyes^{42, 43}, a hydroxyl group in the *ortho* position relative to the azo group implements azo-hydrazone tautomerism, a feature that must be considered during evaluations of the experiments¹⁰. Interestingly, in our model, the amino groups form hydrogen bonds with the carbonyl oxygen of E11, and the same stabilizing interactions were also observed for CR binding to HET fibrils⁴⁰. Finally, most studies suggest that CR binds with its long axis perpendicular to the direction of the β -strands^{11, 44}, as shown herein (Figure 2B,C).

We corroborated the binding modes by comparing $K_{\text{D}}^{\text{comp}}$ with available $K_{\text{D}}^{\text{exp}}$, which are in good agreement; note that, to our knowledge, binding affinities of the probes to the recently resolved A β (1-42) fibril structure at pH 2 have not yet been determined¹⁶. Besides the A β (1-42) fibril structure obtained at pH 2 used here, other A β structures provide further evidence for two symmetrically arranged stacks of A β peptides^{37, 45-47}.

Throughout the simulations, the fibril preserved its structural organization, and dye binding has no significant influence on the fibril structure (Figure S9), although minor structural changes were observed at the fibril ends (see Movie S2 after 37 s). Changing the pH may be of relevance for THT, as the affinity to insulin and HET fibrils decreases at acidic pH conditions, suggesting a pH dependency of THT binding^{11, 48}. As to CR, there is no evidence suggesting a pH dependency, although most of the affinity studies are done at neutral pH⁴⁸.

In conclusion, we present binding sites and modes for THT and CR to the A β (1-42) fibril. Binding sites, modes, and affinities agree with previous experimental observations. The insights

provide a starting point for the systematic search and design of novel probes that bind to A β (1-42) fibrils, which is essential for conclusively diagnosing A β (1-42) fibrils-related diseases.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. J. Wang, B. J. Gu, C. L. Masters and Y. J. Wang, *Nat. Rev. Neurol.*, 2017, **13**, 612-623.
2. P. Scheltens, K. Blennow, M. M. B. Breteler, B. de Strooper, G. B. Frisoni, S. Salloway and W. M. Van der Flier, *Lancet*, 2016, **388**, 505-517.
3. C. L. Masters, R. Bateman, K. Blennow, C. C. Rowe, R. A. Sperling and J. L. Cummings, *Nat. Rev. Dis. Primers.*, 2015, **1**, 1-18.
4. H. V. Vinters, *Annu. Rev. Pathol.: Mech. Dis.*, 2015, **10**, 291-319.
5. A. s. Association®, *2019 Alzheimer's disease facts and figures*, 2019.
6. A. s. D. International, *World Alzheimer Report 2018: The state of the art of dementia research: New frontiers*, 2018.
7. D. J. Selkoe and J. Hardy, *EMBO Mol. Med.*, 2016, **8**, 595-608.
8. M. Goedert, F. Clavaguera and M. Tolnay, *Trends Neurosci.*, 2010, **33**, 317-325.
9. M. Jucker and L. C. Walker, *Nature*, 2013, **501**, 45-51.
10. L. S. Cai, R. B. Innis and V. W. Pike, *Curr. Med. Chem.*, 2007, **14**, 19-52.
11. M. Groenning, *J. Chem. Biol.*, 2010, **3**, 1-18.
12. A. I. Sulatskaya, I. M. Kuznetsova and K. K. Turoverov, *J. Phys. Chem. B*, 2011, **115**, 11519-11524.
13. C. Xue, T. Y. W. Lin, D. Chang and Z. F. Guo, *Roy. Soc. Open Sci.*, 2017, **4**.
14. E. I. Yakupova, L. G. Bobyleva, I. M. Vikhlyantsev and A. G. Bobylev, *Biosci. Rep.*, 2019, **39**.
15. C. Wu, Z. X. Wang, H. X. Lei, W. Zhang and Y. Duan, *J. Am. Chem. Soc.*, 2007, **129**, 1225-1232.
16. L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R. B. G. Ravelli, M. Tusche, C. Lopez-Iglesias, W. Hoyer, H. Heise, D. Willbold and G. F. Schröder, *Science*, 2017, **358**, 116-119.
17. B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *J. Chem. Theory Comput.*, 2012, **8**, 3314-3321.
18. D. A. McQuarrie, *Statistical mechanics*, University Science Books, 2000.
19. H. Y. Sun, L. L. Duan, F. Chen, H. Liu, Z. Wang, P. C. Pan, F. Zhu, J. Z. H. Zhang and T. J. Hou, *Phys. Chem. Chem. Phys.*, 2018, **20**, 14450-14460.
20. H. Gohlke and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 238-250.
21. Z. Cournia, B. Allen and W. Sherman, *J. Chem. Inf. Model.*, 2017, **57**, 2911-2937.
22. J. Zhang, A. Sandberg, A. Konsmo, X. Y. Wu, S. Nystrom, K. P. R. Nilsson, P. Konradsson, H. LeVine, M. Lindgren and P. Hammarstrom, *Chem. - Eur. J.*, 2018, **24**, 7210-7216.
23. Z. P. Zhuang, M. P. Kung, C. Hou, D. M. Skovronsky, T. L. Gur, K. Plossl, J. Q. Trojanowski, V. M. Y. Lee and H. F. Kung, *J. Med. Chem.*, 2001, **44**, 1905-1914.
24. A. I. Sulatskaya, A. A. Maskevich, I. M. Kuznetsova, V. N. Uversky and K. K. Turoverov, *PLoS One*, 2010, **5**.
25. M. Biancalana, K. Makabe, A. Koide and S. Koide, *J. Mol. Biol.*, 2009, **385**, 1052-1063.
26. C. Wu, M. Biancalana, S. Koide and J. E. Shea, *J. Mol. Biol.*, 2009, **394**, 627-633.
27. M. Stravalaci, M. Beeg, M. Salmona and M. Gobbi, *Biosens Bioelectron*, 2011, **26**, 2772-2775.
28. W. E. Klunk, J. W. Pettegrew and D. J. Abraham, *J. Histochem. Cytochem.*, 1989, **37**, 1273-1281.
29. M. G. Iadanza, R. Silvers, J. Boardman, H. I. Smith, T. K. Karamanos, G. T. Debelouchina, Y. C. Su, R. G. Griffin, N. A. Ranson and S. E. Radford, *Nat. Commun.*, 2018, **9**.
30. B. S. Li, P. Ge, K. A. Murray, P. Sheth, M. Zhang, G. Nair, M. R. Sawaya, W. S. Shin, D. R. Boyer, S. L. Ye, D. S. Eisenberg, Z. H. Zhou and L. Jiang, *Nat. Commun.*, 2018, **9**.
31. R. Guerrero-Ferreira, N. M. I. Taylor, D. Mona, P. Ringler, M. E. Lauer, R. Riek, M. Britschgi and H. Stahlberg, *Elife*, 2018, **7**.
32. C. Röder, N. Vettore, L. N. Mangels, L. Gremer, R. B. Rayelli, D. Willbold, W. Hoyer, A. K. Buell and G. F. Schröder, *Nat. Commun.*, 2019, **10**.
33. C. König, R. Skanberg, I. Hotz, A. Ynnerman, P. Norman and M. Linares, *Chem. Commun.*, 2018, **54**, 3030-3033.
34. H. LeVine, *Arch. Biochem. Biophys.*, 1997, **342**, 306-316.
35. M. R. H. Krebs, E. H. C. Bromley and A. M. Donald, *J. Struct. Biol.*, 2005, **149**, 30-37.
36. I. M. Kuznetsova, A. I. Sulatskaya, V. N. Uversky and K. K. Turoverov, *Mol Neurobiol*, 2012, **45**, 488-498.
37. M. T. Colvin, R. Silvers, Q. Z. Ni, T. V. Can, I. Sergeyev, M. Rosay, K. J. Donovan, B. Michael, J. Wall, S. Linse and R. G. Griffin, *J. Am. Chem. Soc.*, 2016, **138**, 9663-9674.
38. Y. L. Xiao, B. Y. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov and Y. Ishii, *Nat Struct Mol Biol*, 2015, **22**, 499-U497.
39. W. Zhen, H. Han, M. Anguiano, C. A. Lemere, C. G. Cho and P. T. Lansbury, *J. Med. Chem.*, 1999, **42**, 2805-2815.
40. A. K. Schütz, A. Soragni, S. Hornemann, A. Aguzzi, M. Ernst, A. Bockmann and B. H. Meier, *Angew. Chem. Int. Edit.*, 2011, **50**, 5956-5960.
41. J. H. Cooper, *Lab. Invest.*, 1974, **31**, 232-238.
42. A. S. Ozen, P. Doruker and V. Aviyente, *J. Phys. Chem. A*, 2007, **111**, 13506-13514.
43. H. Zollinger, *Color chemistry: syntheses, properties, and applications of organic dyes and pigments*, John Wiley & Sons, 2003.
44. L. W. Jin, K. A. Claborn, M. Kurimoto, M. A. Geday, I. Maezawa, F. Sohraby, M. Estrada, W. Kaminsky and B. Kahr, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 15294-15298.
45. M. Kollmer, W. Close, L. Funk, J. Rasmussen, A. Bsoul, A. Schierhorn, M. Schmidt, C. J. Sigurdson, M. Jucker and M. Fändrich, *Nat. Commun.*, 2019, **10**.
46. M. A. Wälti, F. Ravotti, H. Arai, C. G. Glabe, J. S. Wall, A. Böckmann, P. Güntert, B. H. Meier and R. Riek, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E4976-E4984.
47. A. K. Schütz, T. Vagt, M. Huber, O. Y. Ovchinnikova, R. Cadalbert, J. Wall, P. Güntert, A. Böckmann, R. Glockshuber and B. H. Meier, *Angew. Chem. Int. Edit.*, 2015, **54**, 331-335.
48. M. Groenning, M. Norrman, J. M. Flink, M. van de Weert, J. T. Bukrinsky, G. Schluckebier and S. Frokjaer, *J. Struct. Biol.*, 2007, **159**, 483-497.