International Journal of Research in Pharmaceutical sciences and Technology

Elucidating protein aggregation in neurodegeneration diseases using computational approaches

Ram Mahinder[1] and Albert Koo[2]*

[1]*Department of Biotechnology, Kuppam college of Science, Andhra Pradesh, India.*
[2]*Department of genetics, Jagiellonian University, Poland.*

ABSTRACT

The formation of harmful non-native protein conformers has arisen as a common thread in diseases like Alzheimer's, Parkinson's, and amyotrophic lateral sclerosis. Atomic-level information on the dynamic alterations which enable protein accumulation, and also the structural properties of large-scale organized aggregation and soluble non-native oligomers, will considerably contribute to the present comprehension of these complicated processes and give prospective ways for limiting the development of cytotoxic species. Moreover, experimental restrictions frequently prevent the accumulation of high-resolution structural and mechanistic data for aggregating systems. Computational techniques, especially those that integrate all-atom and coarse-grained simulations to cover a broad variety of time and length scales, have therefore emerged as critical tools for studying protein aggregation. Here we review the current status of computational methodology for the study of protein self-assembly, with a focus on the application of these methods toward an understanding of protein aggregates in human neurodegenerative disorders.

Keywords: Computational Approaches, Design, Protein, Aggregation, Neurodegenerations.

ISSN: 2581-9143
Review Article

Corresponding Author
Name: M K Aneebuddin

Article Info
Received on: 06-04-2022
Revised on: 26-04-2022
Accepted on: 06-05-2022
DOI: https://doi.org/10.33974/ijrpst.v3i2.290

Copyright© 2022, Ram Mahinder and Albert Koo, Elucidating protein aggregation in neurodegeneration diseases using computational approaches, Production and hosting by *Rubatosis Publications.*

INTRODUCTION

The transition of native proteins to partially unfolded and aggregated species has been involved in several human neurodegenerative disorders. In healthy people, several quality control systems, such as endoplasmic reticulum-associated degradation and the action of chaperones that process unfolded, misfolded, and aggregated proteins, prevent the creation of potentially harmful aggregates. [1-3]. In stressful situations or among elder people, insufficient capacity of quality control mechanisms may allow the aggregate formation to predominate and, eventually, to cross a pathogenic threshold. Furthermore, Aggregation can take place in the absence of significant cellular stress, even when a protein's native form is extremely thermodynamically advantageous [4]. Aggregation preference is also determined by a protein's subcellular location and accessibility to the numerous quality control mechanisms that sustain proteostasis [5]. As a result, variations in the native structures as well as cellular microenvironments of intrinsically disordered proteins and peptides, folded cytosolic proteins, and membrane proteins account for their significantly diverse aggregation inclinations[6]. Protein aggregates are present in both amorphous (disordered) as well as fibrillar (ordered) forms. One of the most common structural features found in ordered protein aggregates is the amyloid fold [7]. Amyloid fibrils contain a characteristic cross-b architecture, which consists of b-strands running parallel to each other, but perpendicular to the direction of fibril extension [8]. Amyloid aggregates can have a variety of gross morphologies, including twisted ribbons, sheets, tubes, and twisted ropes, depending on how the protofibrillar units are arranged [9]. Even though the general tendency of numerous aggregation-prone proteins to establish amyloid is quite well-formed, there may be a growing consensus that neurotoxicity related to protein aggregation mainly tends to happen in lower-order disordered oligomers (soluble assemblies bigger than the native oligomeric state which might be pre-fibrillar intermediates or species not on the fibrillization route) examples of oligomers, protofibrils, and fibrils of amyloid-beta (Ab) peptide [10]. These are some of the mechanisms proved to assist initial phases of aggregation in connection with membranes, which can be resulting in membrane breakdown and consequent cell death [11]. Prefibrillar oligomers produce non-specific ion channels at the

surface of a membrane during the process of protein misfolding [12]; the resulting change in membrane permeability affects ion homeostasis, resulting in mitochondrial dysfunction and impaired synaptic transmission, between several other neurotoxic consequences [13]. Conformational features of disordered and pre-fibrillar oligomers of various neurodegeneration-linked proteins are discussed in more detail in the latter half of the review.

Computational methods for studying protein misfolding and aggregation:

Computational methods, particularly when integrated with experimental methodologies, are becoming more effective for characterizing changes in protein dynamics and identifying uncommon molecular processes that cause aggregation [14]. In this part, we'll go through some of the most prevalent computational methodologies for studying protein aggregation, beginning with algorithms that predict aggregation propensity based only on amino acid series or in the circumstance of protein tertiary structure. Then, using molecular dynamics (MD) simulations, we address methods for simulating protein aggregation and modeling aggregate structures, including ways for reducing the computational load of simulations including large polypeptides and long-time scales. Evaluation of protein aggregation sequence determinants, the question is how amino acid composition influences aggregation propensity is highly relevant to protein design, as well as to our understanding of protein evolution and the pathogenicity of certain amino acid substitutions. Considering that hydrophobic interactions are a significant driving factor in protein self-assembly, it is predicted that a peptide with a higher hydrophobic content will lead to a higher aggregation tendency, whereas a peptide with a net charge will prevent the aggregation. Furthermore, given the evidence for similar cross-b architecture in between amyloid fibrils of many proteins [15-20], a stretch of amino acids with a higher proclivity to adopt b-strand secondary structure would be predicted to increase fibril formation. Such physicochemical characteristics of amino acids constitute the foundation of the majority of the algorithms that predict either the rate or propensity of aggregation of different regions of a protein. Predictors are based on overall protein composition. The prediction of the effect of a point mutation on the aggregation kinetics of a protein was one of the earliest attempts to rationally anticipate the influence of protein sequence on its aggregation propensity [21]. Fitting the coefficients of an empirical formula to a collection of available data for the aggregation kinetics of 50-point mutations of human muscle acylphosphatase (ACP) in comparison to the wild-type protein provided the basis for this technique **(Chiti et al., 2003)**. The empirical formula had been a linear equation that assimilate the shift in hydrophobicity, net charge, and an inclination for helix-b-sheet transition following mutation. This is a simple version because it treats protein characteristics as a simple sum without taking the position of the mutation in the structure into consideration. Furthermore, two of the three components of the equation, hydrophobicity and net charge, are strongly reliant on one another. Despite these limitations, projections of aggregation rate alterations associated with 27 mutations in a range of disease-related and model proteins produced using this model have a statistically important association with experimental results. This work emphasizes the importance of amino acid physicochemical qualities in determining, to a significant extent, changes in aggregation rate. The two constraints of this model, interdependent terms and free parameter fitting, were solved in a more complex model developed by the Caflisch group [22]. The components of their equation included changes in aromaticity, dipole moment, the ratio of accessible surface area, and b-sheet propensity depending on the kind of mutation (e.g., polar to apolar or apolar to polar). This equation is more sophisticated than the basic linear relationship of chiti et al, but it lacks redundant terms and, more crucially, it lacks free parameters that must be determined by fitting the equation to a training set. Chiti et al model was expanded to estimate the absolute aggregation rate by including external parameters such as ionic strength and pH at which aggregation occurs in vitro [23].

Predictors of peptides in amyloidogenic proteins:

These predictors are found on the conjecture of a well-folded beginning state and the amyloidogenic potential of a peptide inside a protein rather than simply across the complete protein. TANGO [24] is among the models that incorporate these requirements. Rather than a basic linear equation for the whole protein, the TANGO algorithm assumes that a specific amino acid location in a protein can presume four probable states: unfolded, turn, helix, or b-aggregated. The probability of a specific residue existing within every state is determined by its secondary structure propensity, charge-charge interactions, and solvation penalty. TANGO deliberates the division function of these various states to every residue, presuming that the likelihood of a polypeptide chain containing more than two amyloidogenic areas is insignificant. This presumption restricts TANGO's capacity to predict the aggregation propensities of proteins with more than 50 amino acids. To point out the reality that perhaps the folded state represents the beginning point of aggregation, the stability of the folded state is occupied into deliberation within the TANGO algorithm through Fold-X [25-28] computations. In practice, TANGO is extra effective for determining if a protein is going to aggregate than it is for estimating exact aggregation rates. Despite these limitations, TANGO proved successful in estimating the expanse of these proteins that were experimentally known to be liable to aggregate in benchmark testing of disease-associated proteins and their variants.

Recent trends in computational approaches

The latest extension of this approach is WALTZ [29], which, among other developments, significantly expanded the peptide training set in order to reduce biases in the original algorithm that were attributed to the smaller training set. There are numerous other algorithms in the same group of predictors of aggregation-prone stretches of amino acids in a protein, along with an extension of the indigenous technique. There is even a server that incorporates the results of multiple algorithms and provides a consent prediction [30]. An MD analysis of aggregation kinetics [31-34] reveals a two-phase procedure of fibril development, with largely disordered chains forming the contact interface, accompanied by structural transition and aggregation of b-sheet content. It also suggests a critical role of protein–solvent interactions in a-synuclein aggregation. Structures of fibrils have also been a focus of computer-assisted studies. Fibril rupture simulations accept AFM experimentations and suggest the existence of extremely mechanically stable structures in a-synuclein fibrils with higher b-sheet content [34-39].

CONCLUSION

Protein aggregation was already associated with a number of neurodegenerative conditions, the causes of which are undefined, and there are no effective therapies for these disorders. In specific, the initial stage soluble oligomers are assumed to be more harmful, although their heterogeneity and transience frequently prevent experimental characterization, which might lead to the development of techniques to stop their gathering or toxicity. As a result, the high-resolution structural, dynamic, and mechanistic insights provided by computational studies of protein aggregation hold the special potential to allow the rational modulation of oligomer formation. This capacity will allow for more direct verification of the 'cytotoxic oligomer hypothesis,' as well as the identification of possible techniques for limiting the development of toxic oligomers in neurodegenerative diseases.

REFERENCES

1. Redler, R. L., Shirvanyants, D., Dagliyan, O., Ding, F., Kim, D. N., Kota, P., ... & Dokholyan, N. V. (2014). Computational approaches to understanding protein aggregation in neurodegeneration. Journal of molecular cell biology, 6(2), 104-115.
2. Wei, G., Mousseau, N., & Derreumaux, P. (2007). Computational simulations of the early steps of protein aggregation. Prion, 1(1), 3-8.
3. Ross, C. A., & Poirier, M. A. (2004). Protein aggregation and neurodegenerative disease. Nature medicine, 10(7), S10-S17.
4. Sarkar, A., Kumar, S., Grover, A., & Sundar, D. (2012). Protein aggregation in neurodegenerative diseases: Insights from computational analyses. Current Bioinformatics, 7(1), 87-95.
5. Tripathi, V. P., Ahmed, H., Renulos, N., & Aneebuddin, M. K. (2022). Elucidating the Role of Chemokines in Infectious Diseases and Gastric Cancer. Arch Gastroenterol Res, 3(1), 23-26.
6. Haenig, C., Atias, N., Taylor, A. K., Mazza, A., Schaefer, M. H., Russ, J., ... & Wanker, E. E. (2020). Interactome mapping provides a network of neurodegenerative disease proteins and uncovers widespread protein aggregation in affected brains. Cell reports, 32(7), 108050.
7. Thangakani, A. M., Nagarajan, R., Kumar, S., Sakthivel, R., Velmurugan, D., & Gromiha, M. M. (2016). CPAD, curated protein aggregation database: a repository of manually curated experimental data on protein and peptide aggregation. PLoS One, 11(4), e0152949.
8. Tripathi, V. P., & Jetsu, R. (2022). Traditional Herbs Role in Enhancement of Antitumor Responses of Agonistic CD40-Antibody by Reducing Myeloid-derived Suppressor Cells. Chem. Res. J, 7, 1-5.
9. Vaquer-Alicea, J., & Diamond, M. I. (2019). Propagation of protein aggregation in neurodegenerative diseases. Annual review of biochemistry, 88, 785-810.
10. Pedersen, J. T., & Heegaard, N. H. (2013). Analysis of protein aggregation in neurodegenerative disease. Analytical chemistry, 85(9), 4215-4227.
11. Pappu, R. V., Wang, X., Vitalis, A., & Crick, S. L. (2008). A polymer physics perspective on driving forces and mechanisms for protein aggregation. Archives of biochemistry and biophysics, 469(1), 132-141.
12. Santos, J., Pujols, J., Pallarès, I., Iglesias, V., & Ventura, S. (2020). Computational prediction of protein aggregation: advances in proteomics, conformation-specific algorithms and biotechnological applications. Computational and structural biotechnology journal, 18, 1403-1413.
13. Tripathi, V. P., Goo, D., Maidya, B. N., & Aneebuddin, M. K. (2022). New Trends in Interrelation of Infectious Colorectal Cancer with Intestinal Microbiota. Arch Gastroenterol Res, 3(1), 18-22.
14. Khabirova, E. (2016). Models of neurodegeneration using computational approaches (Doctoral dissertation, University of Cambridge).
15. Straub, J. E., & Thirumalai, D. (2011). Toward a molecular theory of early and late events in monomer to amyloid fibril formation. Annual review of physical chemistry, 62, 437-463.
16. Davis, M. Y., Trinh, K., Thomas, R. E., Yu, S., Germanos, A. A., Whitley, B. N., ... & Pallanck, L. J. (2016). Glucocerebrosidase deficiency in Drosophila results in α-synuclein-independent protein aggregation and neurodegeneration. PLoS genetics, 12(3), e1005944.
17. Morris, A. M., Watzky, M. A., & Finke, R. G. (2009). Protein aggregation kinetics, mechanism, and

curve-fitting: a review of the literature. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1794(3), 375-397.

18. Tripathi, V. P., Aneebuddin, M. K., Alex, C., & Moharir, K. S. (2021). Evaluating Clinical Applications of Liquid Biopsy by Combining Circulating Tumor DNA and Tumor cells. Current Research in Pharmaceutical Sciences, 107-111.
19. Papy-Garcia, D., Christophe, M., Minh Bao, H., Fernando, S., Ludmilla, S., Diaz Julia Elisa, S., & Rita, R. V. (2011). Glycosaminoglycans, protein aggregation and neurodegeneration. Current Protein and Peptide Science, 12(3), 258-268.
20. Candelise, N., Scaricamazza, S., Salvatori, I., Ferri, A., Valle, C., Manganelli, V., ... & Misasi, R. (2021). Protein aggregation landscape in neurodegenerative diseases: Clinical relevance and future applications. International Journal of Molecular Sciences, 22(11), 6016.
21. Louka, A., Zacco, E., Temussi, P. A., Tartaglia, G. G., & Pastore, A. (2020). RNA as the stone guest of protein aggregation. Nucleic Acids Research, 48(21), 11880-11889.
22. Gil-Garcia, M., Iglesias, V., Pallarès, I., & Ventura, S. (2021). Prion-like proteins: from computational approaches to proteome-wide analysis. FEBS open bio, 11(9), 2400.
23. Uversky, V. N. (2010). Mysterious oligomerization of the amyloidogenic proteins. The FEBS journal, 277(14), 2940-2953.
24. Tripathi, V. P., & Dubey, D. D. (2017). A replication-time-controlling sequence element in Schizosaccharomyces pombe. Chromosoma, 126(4), 465-471.
25. Viet, M.H., Ngo, S.T., Lam, N.S., et al. (2011). Inhibition of aggregation of amyloid peptides by b-sheet breaker peptides and their binding affinity. J. Phys. Chem. B 115, 7433-7446.
26. Vitalis, A., Lyle, N., and Pappu, R.V. (2009). Thermodynamics of b-sheet formation in polyglutamine. Biophys. J. 97, 303-311.
27. Voegler Smith, A., and Hall, C.K. (2001). a-helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model. Proteins 44, 344-360.
28. Wang, Q., Johnson, J.L., Agar, N.Y., et al. (2008). Protein aggregation and protein instability govern familial amyotrophic lateral sclerosis patient survival. PLoS Biol. 6, e170.
29. Yadav, M. P., Padmanabhan, S., Tripathi, V. P., Mishra, R. K., & Dubey, D. D. (2012). Analysis of stress-induced duplex destabilization (SIDD) properties of replication origins, genes and intergenes in the fission yeast, Schizosaccharomyces pombe. BMC research notes, 5(1), 1-10.
30. Wilcox, K.C., Zhou, L., Jordon, J.K., et al. (2009). Modifications of superoxide dismutase (SOD1) in human erythrocytes: a possible role in amyotrophic lateral sclerosis. J. Biol. Chem. 284, 13940-13947.
31. Pratihar, A. S., Tripathi, V. P., Yadav, M. P., & Dubey, D. D. (2015). Chromosomal context and replication properties of ARS plasmids in Schizosaccharomyces pombe. Journal of biosciences, 40(5), 845-853.
32. Wise-Scira, O., Aloglu, A. K., Dunn, A., Sakallioglu, I. T., & Coskuner, O. (2013). Structures and free energy landscapes of the wild-type and A30P mutant-type α-synuclein proteins with dynamics. ACS chemical neuroscience, 4(3), 486-497.
33. Mishra, A., Nizammuddin, S., Mallick, C. B., Singh, S., Prakash, S., Siddiqui, N. A., ... & Thangaraj, K. (2017). Genotype-phenotype study of the middle Gangetic plain in India shows association of rs2470102 with skin pigmentation. Journal of Investigative Dermatology, 137(3), 670-677.
34. Wu, C., & Shea, J. E. (2011). Coarse-grained models for protein aggregation. Current opinion in structural biology, 21(2), 209-220.
35. Wu, K.-P., Weinstock, D.S., Narayanan, C., et al. (2009). Structural reorganization of a-synuclein at low pH observed byNMR and REMD simulations. J. Mol. Biol. 391, 784-796.
36. Yun, S., Urbanc, B., Cruz, L., et al. (2007). Role of electrostatic interactions in amyloid b-protein (Ab) oligomer formation: a discrete molecular dynamics study. Biophys. J. 92, 4064-4077.
37. Poche, D. M., Garlapati, R. B., Mukherjee, S., Torres-Poché, Z., Hasker, E., Rahman, T., ... & Poché, R. M. (2018). Bionomics of Phlebotomus argentipes in villages in Bihar, India with insights into efficacy of IRS-based control measures. PLoS neglected tropical diseases, 12(1), e0006168.
38. Zhang, Z., Chen, H., and Lai, L. (2007). Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential.Bioinformatics 23, 2218-2225.
39. Zhou, B., Hao, Y., Wang, C., et al. (2012). Conversion of natively unstructured a-synuclein to its a-helical conformation significantly attenuates production of reactive oxygen species. J. Inorg. Biochem. 118C, 68-73.