Published in final edited form as:

Biochem Biophys Res Commun. 2015 September 25; 465(3): 523–527. doi:10.1016/j.bbrc.2015.08.052.

A model for non-obligate oligomer formation in protein aggregation

Eamonn F. Healy

Department of Chemistry, St. Edward's University, Austin, TX 78704, USA

Eamonn F. Healy: healy@stedwards.edu

Abstract

Using solvent-exposed intramolecular backbone hydrogen bonds as physico-chemical descriptors for protein packing, a role for transient, non-obligate oligomers in the formation of aberrant protein aggregates is presented. Oligomeric models of the both wild type (wt) and select mutant variants of superoxide dismutase (SOD1) are proposed to provide a structural basis for investigating the etiology of Amyotrophic Lateral Sclerosis (ALS).

Keywords

Dehydron; Obligate non-obligate oligomer; SOD1

1. Introduction

Protein–protein association is essential for a wide variety of cellular processes. Where the complex formed by the interaction of more than one polypeptide chain is the biologically active state then the assembly is termed an obligate oligomer, and represents the protein's quaternary structure. While there exists a wide variety of techniques, including light scattering [1], mass-spectrometry [2] and Nuclear Magnetic Resonance [3], capable of elucidating various aspects of macromolecular structure most of our knowledge of the structural aspects of protein–protein interactions (PPIs) comes from protein crystallography [4]. The Protein data Bank (www.rcsb.org) represents a compilation of 3D protein structures, more than 80% of which were characterized by X-ray crystallography [5]. It has been estimated that monomers represent approximately 25% of the structures deposited in the PDB, with dimers representing a further 15% [6]. While the bias in the data towards soluble proteins is not surprising, the relatively low percentage, less than 7%, of hetero-oligomeric complexes of all sizes is noteworthy.

Given the dominant role of crystallography in protein structure determination it is not surprising that crystal contacts between proteins serve as a common starting point in the analysis of biologically important protein–protein contacts. However it cannot be assumed that the crystallization conditions are an accurate reflection of the biological conditions driving macromolecular assembly and protein function *in vivo*. Additionally the contents of

Transparency document

Transparency document related to this article can be found online at <http://dx.doi.org/10.1016/j.bbrc.2015.08.052>.

the asymmetric symmetry unit (ASU) may represent multiple copies of the biological state of the protein. Conversely the presumed biological molecule might best be represented by applying crystallographic symmetry to the contents of the ASU. While water mediates many protein functions and plays an important, if not determinant, role in protein folding, the fraction of water crystallized with the protein has been shown to correlate with crystal symmetry. A correlation was also observed between the crystal symmetry and the internal symmetry of the protein oligomer [7]. Thus the oligomeric unit presented by the ASU is, at least to some extent, a product of the experimental conditions employed.

However, there is much that we can infer about macromolecular assemblies from the crystalline state. While studies of the various structural determinants of the binding affinity of protein–protein complexes have covered the gamut of physicochemical descriptors, analysis of amino acid preferences and secondary structure packing at the interfaces has yielded some of the most useful insights [8,9]. Buried surface area (BSA) remains a primary descriptor of binding affinity, and this has led to the identification of interface “hot spots” [10], or residues that, when substituted by alanine, correlate with a change in the free energy of dissociation, ΔG_{diss}^0 . Informatics approaches that score putative protein complexes represent an attempt to enumerate possible assemblies by identifying interfaces as either biologically relevant or simply contacts due to crystal packing forces [11]. Since 2007 the PISA (Protein Interface, Surfaces and Assemblies) protocol, a graph-theoretical approach to identify chemically stable and biologically relevant associations that may potentially be formed in a given crystal packing, has been used in the PDB submission process to aid in quaternary structure determination [12]. Alternatively the user can utilize PISA to independently analyze any PDB entry.

Uncontrolled oligomerization, in concert with protein misfolding, converts peptides and proteins from their soluble forms into insoluble fibrillar aggregates, that can give rise to a variety of pathological conditions. While molecular chaperones, including the small heat shock proteins (sHsps), play an important role in protecting against protein misfolding and aggregation, for diseases such as Alzheimer’s disease (AD), Parkinson’s (PD) and Amyotrophic Lateral Sclerosis (ALS), the accumulated protein aggregates eventually overtaxes the heat shock response, leading to apoptosis and neuronal cell death. Not surprisingly an important paradigm driving therapeutic development has been that the formation of these highly organized fibrillary aggregates leads to neurodegeneration and death [13], despite the fact that the relevance of plaques, or extracellular fibrils, to AD pathogenesis remains unclear since substantial neuronal dysfunction is observed prior to the appearance of fibrillary deposits [14]. There is also a growing body of evidence to suggest that these fibrils may be just end-stage products, and that the primary toxic element is a pre-fibrillar oligomer [15]. The failure of therapeutic strategies based on fibril elimination [16,17], is leading to a shift in focus away from deposition and towards aggregation and oligomerization [18]. ALS is linked to the misfolding and aggregation of superoxide dismutase (SOD1), with over 90% of cases sporadic and the remaining familial cases associated with a wide array of inherited mutations. While it is known that oligomeric assemblies of both wild type (wt) and mutant SOD1 are precursors to the larger and detergent-insoluble aggregates, the structural events that trigger oligomerization remain

elusive. Furthermore studies have found that mutant, misfolded SOD1 can convert wtSOD1 in a prion-like fashion [19], and that misfolded wtSOD1 can be propagated by release and uptake of protein aggregates [20].

Solvent-exposed intramolecular backbone hydrogen bonds, or dehydrons, have been previously identified as vulnerabilities or structural defects, in the packing of a wide array of proteins [23,24]. Exposure of such dehydrons to an aqueous environment has been shown to weaken protein secondary structures [25,26]. In turn excluding solvent from protein regions containing exposed hydrogen bonds has been implicated as a determinant factor in ligand-protein [27] and protein-protein [28] interactions, as well as protein subunit assembly [29]. This dehydron hypothesis has also been utilized to provide a mechanism of action for sHsp, based on the protection of solvent-exposed backbone hydrogen bonds within the α -crystalline domain [30,31], as well as structural basis for the modulation of SCA3 toxicity by α B-crystalline [32]. This paper seeks to highlight the utility of looking beyond the x-ray crystallographic data when analyzing for protein-protein contacts in putative protein oligomers. Using the protein SOD1 as an example, and focusing on solvent-exposed intramolecular backbone hydrogen bonds as physicochemical descriptors for protein packing, we will present a role for transient, non-obligate oligomers in the formation of aberrant protein aggregates.

2. Methods

The crystal structures for the wild type of human superoxide dismutase are available from the RCSB (www.rcsb.org) as PDB entry 2C9V [33]. A wide array mutant structures are also available, including G37R (1AZV) [34], G85R (2ZKX) and G93A (2WZ6) [35]. After adding hydrogens these proteins were subjected to a short energy minimization using the CHARMM force field [36]. Protein alignments and superimposition were done using the MODELER protocol as implemented in the Discovery Studio program suite.

The extent of hydrogen bond desolvation is quantified as the number of non-bonded, carbonaceous groups, ρ , contained within a domain centered on the residues linked by the interaction [28]. This desolvation domain is defined as two intersecting spheres of fixed radius centered on the C_{α} atoms of the linked residues. Dehydrons are then identified as those backbone hydrogen bonds that are underwrapped by non-polar carbonaceous groups, and defined as those interactions with ρ values at or below the average minus one root mean squared deviation. In this work the default values for domain radius, 6.2 Å, and dehydron cutoff, $\rho - 19$, were used as per reference 28. The dehydrons for SOD1 are shown as green (in web version) connectors in Fig. 1.

The biological assembly for SOD1 shown in Fig. 1 was predicted by PISA (available at <http://www.ebi.ac.uk/pdbe/pisa/>) analysis of the pdb entry 2ZKX [12]. PISA enumerates all macromolecular assemblies that may be potentially formed for a given crystal packing. The Gibb's free energy of interaction as well as an entropy cost of oligomerization is calculated as in Ref. [37]. Those complexes with an overall positive free energy of dissociation are identified as chemically stable. Complexes may be misrepresented due to failures in the energy approximation, or in the inferences made from the crystal packing.

ZDOCK is a Fast Fourier Transform rigid docking protocol that searches all possible binding modes in the translational and rotational spaces between two proteins, and evaluates each using an energy scoring function based on shape complementarity [38]. By associating an unfavorable desolvation energy contribution with specified atoms, protein residues can be blocked from being included in binding sites. Results are filtered by clustering poses where the RMSD lies within a specified cutoff of the predicted binding interface. Of the 54,000 docked poses generated for the SOD1 tetramer using the default parameters as implemented in the Discovery Studio program suite from Accelrys Inc. the 2000 top-ranked poses were filtered into 93 clusters. For each cluster an average contact surface area was generated by calculating the change in the solvent accessible surface (SAS) area upon formation of the protein–protein complex for each docked pose. Finally the pose with the largest excluded SAS area, from the cluster with the largest average contact surface area, was selected. This oligomer is shown in Fig. 1.

3. Results and discussion

Analysis of the dehydron pattern for the wtSOD1 monomer is shown in Fig. 1, and reveals that the underwrapped, solvent-exposed backbone hydrogen bonds are located in three relatively well defined regions (circled). One of these regions lies at the dimer interface and is thus dehydrated and stabilized through formation of the asymmetric unit. The other two areas remain exposed in the dimer, raising the possibility that a higher order oligomer that stabilizes the exposed regions through transient protein–protein interactions (PPI) would be favored under biological conditions. The tetramer shown represents the pose from the cluster with the largest average contact surface area, as identified by the ZDOCK protocol, that in turn gives the greatest solvent excluded surface area upon dimer–dimer contact, a value of 1370 Å². For comparison this is approximately 60% greater than the average excluded SAS area calculated for those poses in the highest occupancy cluster. As can be seen this PPI corresponds perfectly with the two erstwhile exposed regions containing hydrogen bonded vulnerabilities. One of these regions is shielded by contact with the adjacent dimer interface, while the other is “patched” by residue Trp₃₂. In turn the D₂ symmetry of the tetramer ensures that a corresponding Trp₃₂ patch is to be found further along the interface. The combination of C₂ point group symmetry and P₂₁ space group symmetry in 2C9V, also allows for a D₂ tetramer to be generated simply by applying crystallographic symmetry to the dimeric ASU. For the G85R mutant, which crystallizes in the I2₁2₁2₁ space group, the ASU itself is represented by a D₂ tetramer, nearly identical structure to that generated by protein docking. PISA analysis of 2ZKX yields the prediction that two of these tetramers will oligomerize to give a stable higher-order biological assembly. As can be seen from the circled region in Fig. 1 this 8 subunit assembly is predicted to occur through contact at a Trp₃₂ “hot spot”. Not surprisingly a cluster was also identified by the ZDOCK protocol yielding just such a Trp–Trp contact point at the oligomer interface. While the ranking [39] for the optimum pose from this cluster was nearly identical to that shown in Fig. 1, the measured SAS excluded area was nearly 15% less.

While the etiology of sporadic ALS is largely unknown the vulnerabilities represented by the solvent-exposed hydrogen bonds in wtSOD1 might well provide possible clues. While the region containing the most underwrapped hydrogen bonded pairs in the monomer, Asn₅₃

& Gly₅₆ and Gly₅₆ & Ser₅₉, are desolvated by the subunit interface the other two regions remain exposed in the dimer. A similar pattern is found in the G37R, G85R and G93A mutants, as shown in Fig. 2. While slight variations in wrapping efficiency can be seen, each of the three regions remain extremely vulnerable to solvent exposure across both the wt and mutant variants, indicating that biological assembly to form a transient oligomer is a shared topological requirement. The fact that both wt and mutant SOD1 would be predicted by this analysis to oligomerize to a D₂ tetramer is consistent with the findings that familial and sporadic ALS most likely share a common pathogenic mechanism [40]. Thus, whether caused by metal depletion or mutagenesis, any conformational change that interferes with the formation of a non-obligate tetramer would leave vulnerable protein regions exposed, and subject to structural distortion. Indeed it has previously been shown that the transient structural distortion that accompanies Cu/Zn loss in wtSOD1 leads to destabilized conformations that form non-native oligomers [41]. Clues as to what form these non-native oligomers might take are available from the biological assembly predicted for the G85R SOD1 mutant. As we have previously demonstrated [24,25,32] exposure of a region, such as the short helix patched by Trp₃₂ in the tetramer from Fig. 1, to solvent would lead to a loss of secondary structure. In this case that would be due to the disruption of Asn₁₃₁-Ser₁₃₄, Glu₁₃₂-Thr₁₃₅ and Glu₁₃₃-Lys₁₃₆ intra-helical hydrogen bonds. Such a distortion would no longer require patching by the Trp residue, facilitating oligomeric formation through Trp₃₂-Trp₃₂ contacts, such as that shown in Fig. 1. Such Trp hot spots are not just known to be systemically important in multimeric assembly [42], but Trp₃₂ has also been shown to play a critical role in the transmission and propagation of SOD1 misfolding [19,20]. Thus while misfolded human wtSOD1, with Trp at position 32, has been shown to be capable of inducing the misfolding of natively structured wtSOD1, murine SOD1, with a serine instead of a tryptophan at position 32, has no such capability.

The third of the dehydron regions circled in Fig. 1 encompasses solvent exposed hydrogen bonds spanning the DKDG loop connecting the 5th and 6th β -sheets. This region comprises a critical element of the epitope recognized by the C4F6 monoclonal antibody raised against the G93A ALS variant. In comparing the wt and G93A variants in this region previous researchers have noted the lack of an obvious conformational signature that would explain the specificity of C4F6 binding [43]. Instead it was proposed that the increased flexibility known to accompany most of the familial ALS mutations [44] is propagated through the protein backbone and in turn reduces the rigidity of this “plug”, and allows greater access and greater antibody reactivity. The analysis in Fig. 2 provides for a more direct mechanism, namely the disruption and loss of the already weakened, and highly vulnerable, Asp₉₀(NH)-(O)Val₉₄ hydrogen bond in the G93A mutant caused by exposure to solvent, following in turn from the failure to form a transient oligomer that could serve to “patch” that vulnerability.

In conclusion it is important to note that this analysis seeks to highlight the need to look beyond the crystal structure when investigating biological processes that involve, or may involve, protein association. The utility of a topological analysis based on underwrapped, and thus vulnerable, hydrogen bonds is emphasized. That this perspective may aid in providing a structural basis for analyzing some of the more intriguing biochemical and clinical aspects of conditions such as ALS is hoped. However given the inordinate

complexity of protein recognition and association such an analysis must be considered but one of the many approaches necessary when investigating the functional interactions that underlie the cell's biology.

Acknowledgments

The author wishes to acknowledge the support of the National Institute of General Medical Services (1K12GM102745), as well as Welch Foundation (BH-0018) for its continuing support of the Chemistry Department at St. Edward's University. The author is particularly grateful to Dr. Neil Cashman, University of British Columbia, for both his insight and encouragement in the preparation of this manuscript.

References

1. Liu T, Chu B. Light scattering by proteins. *Encycl Surf Colloid Sci*. 2002; 3:3023–3043.
2. Dass, C. Principles and Practice of Biological Mass Spectrometry. John Wiley & Sons, Inc; New York: 2001.
3. Cavanagh, J.; Fairbrother, WJ.; Palmer, AG., III; Skelton, NJ. Protein NMR Spectroscopy, Principles and Practice. Academic Press; San Diego: 1995.
4. Blundell, TL.; Johnson, LN. Protein Crystallography. Academic Press Inc; London: 1976.
5. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007; 372:774–797. [PubMed: 17681537]
6. Matthews, JM.; Sunde, M. Advances in Experimental Medicine and Biology. Vol. 747. Springer; New York: 2012. Protein Dimerization and Oligomerization in Biology; p. 1-18.
7. Chruszcz M, Potrzebowski W, Zimmerman MD, Grabowski M, Zheng H, Lasota P, Minor W. Analysis of solvent content and oligomeric states in protein crystals—does symmetry matter? *Protein Sci*. 2008; 17:623–632. [PubMed: 18359856]
8. Griffin, MDW.; Gerrard, JA. Advances in Experimental Medicine and Biology. Vol. 747. Springer; New York: 2012. Protein Dimerization and Oligomerization in Biology; p. 74-87.
9. Jones, S. Advances in Experimental Medicine and Biology. Vol. 747. Springer; New York: 2012. Protein Dimerization and Oligomerization in Biology; p. 42-54.
10. Moal IH, Fernández-Recio J. SKEMPI: a Structural Kinetic and energetic database of mutant protein interactions and its use in empirical models. *Bioinformatics*. 2012; 28:2600–2607. [PubMed: 22859501]
11. Ponstingl H, Henrick K, Thornton JM. Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins Struct Funct Bioinform*. 2000; 41:47–57.
12. Krissinel E. Macromolecular complexes in crystals and solutions. *Acta Crystallogr Sect D Biol Crystallogr*. 2011; 67:376–385. [PubMed: 21460456]
13. Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science*. 1992; 256:184. [PubMed: 1566067]
14. Klein WL, Krafft GA, Finch CE. Targeting small A β oligomers: the solution to an Alzheimer's disease conundrum? *Trends Neurosci*. 2001; 24:219–224. [PubMed: 11250006]
15. Caughey B, Lansbury PT Jr. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu Rev Neurosci*. 2003; 26:267–298. [PubMed: 12704221]
16. Kirkitadze MD, Bitan G, Teplow DB. Paradigm shifts in Alzheimer's disease and other neurodegenerative disorders: the emerging role of oligomeric assemblies. *J Neurosci Res*. 2002; 69:567–577. [PubMed: 12210822]
17. Haass C, Selkoe DJ. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide. *Nat Rev Mol Cell Biol*. 2007; 8:101–112. [PubMed: 17245412]
18. Redler RL, Fee L, Fay JM, Caplow M, Dokholyan NV. Non-native soluble oligomers of Cu/Zn superoxide dismutase (SOD1) contain a conformational epitope linked to cytotoxicity in amyotrophic lateral sclerosis (ALS). *Biochemistry*. 2014; 53:2423–2432. [PubMed: 24660965]

19. Grad LI, Guest WC, Yanai A, Pokrishevsky E, O'Neill MA, Gibbs MAE, Cashman NR. Intermolecular transmission of superoxide dismutase 1 mis-folding in living cells. *Proc Natl Acad Sci*. 2011; 108:16398–16403. [PubMed: 21930926]
20. Grad LI, Yerbury JJ, Turner BJ, Guest WC, Pokrishevsky E, O'Neill MA, Cashman NA. Intercellular propagated misfolding of wild-type Cu/Zn superoxide dismutase occurs via exosome-dependent and-independent mechanisms. *Proc Natl Acad Sci*. 2014; 111:3620–3625. [PubMed: 24550511]
23. Fernández A, Scheraga H. Insufficiently dehydrated hydrogen bonds as determinants of protein interactions. *Proc Natl Acad Sci U S A*. 2003; 100:113–118. [PubMed: 12518060]
24. Healy EF, Johnson S, Hauser C, King P. Tyrosine kinase inhibition: ligand binding and conformational change in c-Kit and c-Abl. *FEBS Lett*. 2009; 583:2899–2906. [PubMed: 19660459]
25. Healy EF. The effect of desolvation on nucleophilic halogenase activity. *Comput Theor Chem*. 2011; 964:91–93.
26. Healy EF, Romano P, Mejia M, Lindfors G III. Acetylenic Inhibitors of ADAM10 and ADAM17: in silico analysis of potency and selectivity. *J Mol Graph Model*. 2010; 29:436–442. [PubMed: 20863729]
27. Fernández A, Ridgway S. Dehydron: a structure-encoded signal for protein interactions. *Biophysical J*. 2003; 85:1914–1928.
28. Maddipati S, Fernández A. Feature-similarity protein classifier as a ligand engineering tool. *Biomol Eng*. 2006; 23:307–315. [PubMed: 17110166]
29. Fernández A, Lynch M. Non-adaptive origins of interactome complexity. *Nature*. 2011; 474:502–505. [PubMed: 21593762]
30. Healy EF, King PJ. A mechanism of action for small heat shock proteins. *Biochem Biophys Res Commun*. 2012; 417:268–273. [PubMed: 22155251]
31. Healy EF. A model for heterooligomer formation in the heat shock response of *Escherichia coli*. *Biochem Biophys Res Commun*. 2012; 420:639–643. [PubMed: 22450329]
32. Healy EF, Little C, King PJ. A model for small heat shock protein inhibition of polyglutamine aggregation. *Cell Biochem Biophys*. 2014; 69:275–281. [PubMed: 24242192]
33. Strange RW, Antonyuk SV, Hough MA, Doucette PA, Valentine JS, Hasnain SS. Variable metallation of human superoxide dismutase: atomic resolution crystal structures of Cu–Zn, Zn–Zn and as-isolated wild-type enzymes. *J Mol Biol*. 2006; 356:1152–1162. [PubMed: 16406071]
34. Eisenberg D, Hart PJ, Liu H, Pellegrini M, Nersissian AM, Gralla EB, Valentine EJS. Subunit asymmetry in the three-dimensional structure of a human CuZnSOD mutant found in familial amyotrophic lateral sclerosis. *Protein Sci*. 1988; 7:545–555. [PubMed: 9541385]
35. Galaldeen A, Strange RW, Whitson LJ, Antonyuk SV, Narayana N, Taylor AB, Hart PJ. Structural and biophysical properties of metal-free pathogenic SOD1 mutants A4V and G93A. *Arch Biochem Biophys*. 2009; 492:40–47. [PubMed: 19800308]
36. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem*. 1983; 4:187–217.
37. Krissinel E. Crystal contacts as nature's docking solutions. *J Comput Chem*. 2010; 31:133–143. [PubMed: 19421996]
38. Chen R, Li L, Weng Z. ZDOCK: an initial-stage protein-docking algorithm. *Proteins*. 2003; 52:80–87. [PubMed: 12784371]
39. Pierce B, Weng Z. ZRANK: reranking protein docking predictions with an optimized energy function. *Proteins*. 2007; 67:1078–1086. [PubMed: 17373710]
40. Bosco DA, Morfini G, Karabacak NM, Song Y, Gros-Louis F, Pasinelli P, Brown RH Jr. Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. *Nat Neurosci*. 2010; 13:1396–1403. [PubMed: 20953194]
41. Teilum K, Smith MH, Schulz E, Christensen LC, Solomentsev G, Oliveberg M, Akke M. Transient structural distortion of metal-free Cu/Zn superoxide dismutase triggers aberrant oligomerization. *Proc Natl Acad Sci*. 2009; 106:18273–18278. [PubMed: 19828437]

42. Moreira IS, Fernandes PA, Ramos MJ. Hot spots—a review of the protein–protein interface determinant amino-acid residues. *Proteins Struct Funct Bioinform.* 2007; 68:803–812.
43. Ayers JI, Xu G, Pletnikova O, Troncoso JC, Hart PJ, Borchelt DR. Conformational specificity of the C4F6 SOD1 antibody; low frequency of reactivity in sporadic ALS cases. *Acta Neuropathol Commun.* 2014; 2:55. [PubMed: 24887207]
44. Prudencio M, Borchelt DR. Superoxide dismutase 1 encoding mutations linked to ALS adopts a spectrum of misfolded states. *Mol Neurodegener.* 2011; 6:1–19. [PubMed: 21211002]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1. Dehydron analysis (upper left) and the asymmetric unit (upper right) for wtSOD1, with the three regions of vulnerability circled; the optimum pose for oligomer formation generated by ZDOCK (lower right), with the three regions of vulnerability circled and Trp₃₂ shown as spacefill, and the biological assembly predicted by PISA for the G85R mutant (lower left), with the Trp₃₂ hot spot circled.

Fig. 2.
Key donor–acceptor hydrogen bond distances (carbonaceous group wrapping value) in the regions of vulnerability identified for wt and mutant SOD1.