

# Protein Misfolding, Genetics, and Disease

Anika Nagpal

4<sup>th</sup> December 2014

Genomics and Medicine- Professor Douglas Brutlag

Coined in 1883, the term Protein, comes from the Greek word *proteios*, which means “primary” or “holding the first place.” Indeed, proteins are biological molecules which hold this stature in the context of our living body, having primary responsibility in controlling almost every chemical process which our lives depend on. Their function is closely related to their intricate structure, which is determined by a highly regulate, complex mechanism of folding. Although being composed of sequences of only twenty different amino acids, there are around 100,000 different types of proteins, thus their structures are, perhaps not surprisingly, highly intricate and specific. Despite the highly regulated process of folding, proteins have a tendency to misfold due to the sheer number of proteins and complexity of the system. The term “misfolding” is described as “a process that results in a protein acquiring a sufficient number of non-native interactions to affect its overall architecture and/or its properties in a biologically significant manner”<sup>1</sup>. Many dysfunctions of the body are manifestations of these cases of aberrant folding. The exact mechanism by which this process takes place has not been fully elucidated, but a vast number of clinical trials have given us insights to the biological importance of proteins in the most basic and key bodily functions. Thus, figuring out the process of misfolding may be key to understanding how many bodily processes go awry, and how the protein folding system could inspire viable treatments which target the diseases at their protein core.

---

<sup>1</sup> Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.

Recognizing this importance, since the genome sequencing project completed in 2001, the field of structural genomics has strived to describe the 3-dimensional structure of every protein encoded by a given genome. This genome-based approach allows for a high-throughput method of structure determination by combining experimental and modeling approaches. Although experimental structure determination methods are providing high-resolution structure information about a subset of the proteins, for the large fraction of sequences, whose function is little known, this smaller subset of proteins acts as a template upon which other structures are comparatively modelled<sup>2 3</sup>. Furthermore, computer modelling is not limited by experimental constraints and resistance. Thus a combination of these two approaches is being undertaken in conjunction, giving us an overall broader reach, and better insights than could be achieved by relying on either of these two methods alone<sup>4</sup>.

Understanding the final structure of a protein however, is only one aspect which needs to be tackled. The process of attaining that native structure is equally important for the information to have maximal clinical application against disease. There have been many studies performed and theories proposed on the mechanism of protein folding. One which is backed up by much evidence is that a protein folds such that it finally reaches the most stable, lowest energy structure, with the lowest entropy<sup>5</sup>. Studies have shown that in attaining this final native

---

2 Baker, D. "Protein Structure Prediction and Structural Genomics." *Science* 294.5540 (2001): 93-96. Web.

3 Chance, M. R. "High-Throughput Computational and Experimental Techniques in Structural Genomics." *Genome Research* 14.10b (2004): 2145-154. Web.

4 Luheshi, Leila M., Damian C. Crowther, and Christopher M. Dobson. "Protein Misfolding and Disease: From the Test Tube to the Organism." *Current Opinion in Chemical Biology* 12.1 (2008): 25-31. Web.

<sup>5</sup> Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.

structure, a protein will go through a sequence of transition states<sup>6</sup>. Furthermore, these transition state structures have the key interactions which are integral to the final native structure<sup>7 8</sup>.

The process of folding in an organism, however, is more complex, because the external cellular environment of a protein heavily modulates protein folding in vivo. This is why although we can calculate the propensity of a protein to aggregate based upon a wealth of quantitative data based upon the intrinsic physico-chemical properties taken from experiments conducted in vitro, extrinsic factors are equally important in elucidating the process<sup>9</sup>. One such extrinsic quality is the physiochemical environment such as temperature, which is known to change the strength of hydrophobic interactions<sup>10</sup>. Another example was investigated by a study called “progressive disruption of cellular protein folding in models of polyglutamine diseases”, which showed that the propensity of a protein to aggregate within a cell also depended on the prevailing misfolded proteins in that cell<sup>11</sup>.

These are especially relevant to misfolding, because the surface residues of the protein interact with this environment, and often incompletely-folded transition state proteins have non-native topologies, and so exposed surfaces interact with the environment and molecules in the environment inappropriately<sup>12</sup>. This makes misfolded or partially unfolded proteins highly prone

---

6 Dinner AR, Sali A, Smith LJ, Dobson CM, Karplus M. Understanding protein folding via free energy surfaces from theory and experiment. *Trends Biochem Sci* 2000;25:331–9.

7 Vendruscolo M, Paci E, Dobson CM, Karplus M. Three key residues form a critical contact network in a transition state for protein folding. *Nature* 2001;409:641–6.

8 Makarov DE, Plaxco KW. The topomer search model: a simple, quantitative theory of two-state protein folding kinetics. *Prot Sci* 2003;12:17–26.

9 Luheshi, Leila M., Damian C. Crowther, and Christopher M. Dobson. "Protein Misfolding and Disease: From the Test Tube to the Organism." *Current Opinion in Chemical Biology* 12.1 (2008): 25-31. Web.

10 Bross, Peter, and Brage Andresen. "Protein Misfolding and Degredation in Genetic Disease." (2011): n. pag. Web.

11 T, Gidalevitz, and Ben-Zvi A. "Progressive Disruption of Cellular Protein Folding in Models of Polyglutamine Diseases." *Science* (2006): n. pag. Web.

12 Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.

to aggregation, and closely linked to disease, and so there are inherent cellular mechanisms which prevent misfolding in this complex process, including molecular chaperones, catalysts and degradation pathways. During incomplete folding states, regions rich in hydrophobic residues are usually buried within the structure in the native structure, but become exposed in the process of folding. Molecular chaperones have been shown to bind to nascent chains and protect these regions from such interactions during incomplete folding states<sup>13</sup>. Other types of molecular chaperones are able to solubilize incorrectly folded proteins to give them a second chance to fold correctly. Catalysts play a similar role in preventing misfolding by speeding up folding reactions which would otherwise pass through transition states slowly, making it more likely that they interact with the cellular environment incorrectly<sup>14</sup>.

In vivo, despite these protective measures, the tendency of a protein to misfold is still so high and there are a large number of proteins which circumvent these protective systems to form incorrect structures. Thus there is a secondary system of a highly regulated degradation system to selectively target such proteins, carried out either by ubiquitin-proteasome systems, or lysosomes. Lysosomes are responsible for degrading extracellular proteins which are taken into the cell, through a process of largely unspecific degradation. However, some lysosome autophagy is mediated by chaperones which selectively degrade proteins with a decreased folding capacity<sup>15</sup>. Proteasomes, on the other hand, deal with proteins synthesized inside the cell. The central element of the proteasome system is the covalent linkage of a protein called ubiquitin to proteins targeted for destruction, which are then recognized by the regulatory particle of the

---

13 Hartl FU, Hayer-Hartl M. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 2002;295:1852–8.

14 Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.

15 Bross, Peter, and Brage Andresen. "Protein Misfolding and Degredation in Genetic Disease." (2011): n. pag. Web.

proteasome. The protein is unfolded using energy from ATP, after which the unfolded polypeptide chain is translocated to the core particle where the enzyme breaks key peptide bonds integral to the structure, and thus the functioning of the protein<sup>16</sup>. Degradation pathways must be highly regulated because they involves unfolding and digesting the protein into smaller fragments which may be more prone to misfolding at certain stages as compared to the native protein<sup>17</sup>.

Because proteins regulate a number of key processes in the functioning of the body, common diseases such as Alzheimer's, Parkinson's, and Diabetes to name a few, are increasingly being traced back to underlying cases of misfolded proteins. The two mechanisms by which misfolding can be linked to disease are defined as a gain-of-function and loss-of-function. In the first case, the misfolded protein aggregates to become toxic and thus gains a pathological function<sup>18</sup>, whereas in loss-of-function, the native formation is not acquired, and thus there is less of the functional protein to carry out key cellular processes in the body.

This mechanism of toxic gain-of-function further reveals the process by which misfolded proteins associate with many diseases, especially neurodegenerative diseases. A key characteristic of diseases associated with protein misfolding is the deposition of proteins in the form of amyloid fibrils and plaques. The aggregate forms of these insoluble deposits vary between different diseases, but they have also been shown to share key structural properties under optical observation, such as the "cross-beta" x-ray fibre diffraction pattern<sup>19</sup>, which is

---

16 J, Wang, and Maldonado MA. "The Ubiquitin-proteasome System and Its Role in Inflammatory and Autoimmune Diseases." (2006): n. pag. Web.

17 Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.

18 *ibid*

19 Sunde M, Blake CCF. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv Prot Chem* 1997;50:123-59.

composed of beta sheets running perpendicular to the fibril axis. The formation of the amyloid fibril has been shown to first involve partially unfolded intermediates or largely unfolded denatured forms of the native protein structure, and then conformational reconfiguration into the amyloid structure into the cross beta structure. The process of this reconfiguration is unclear, yet intriguing because amyloid formation defies the theory discussed earlier which states that the final protein structure is one with the lowest energy, and thus the highest stability, because amyloids are not the lowest energy structures<sup>20</sup>. It has been proposed that factors such as “protein degradation upon turnover, interaction with membranes or posttranscriptional modifications, including phosphorylation or glycation<sup>21</sup>” may be involved in reconfiguration. A partially unfolded intermediate of the reconfiguration likely plays an important role in the aggregation of amyloid fibers which is seen as a characteristic of many diseases associated with protein misfolding.

Interestingly, recent studies have shown that the ability of polypeptide chains to form amyloid structures is a common feature of all polypeptide chains, instead of being encoded by specific sequence motifs present in only some proteins associated with disease<sup>22</sup>. A series of experiments by Dobson and co-workers, have shown that proteins that are not relevant to amyloid disease such as myoglobins, too are capable of forming amyloid fibrils under destabilized conditions<sup>23</sup>. When the structure of both disease-causing fibrils and non-disease causing fibrils were examined, it was shown that the core structure of the fibril was determined by hydrogen bond interactions involving the polypeptide main chain- a chain common to all polypeptides. In

---

20 Ohnishi, S., and K. Takano. "Amyloid Fibrils from the Viewpoint of Protein Folding." Cellular and Molecular Life Sciences (CMLS) 61.5 (2004): 511-24. Web.

21 ibid

22 Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." Seminars in Cell & Developmental Biology 15.1 (2004): 3-16. Web.

23 Ohnishi, S., and K. Takano. "Amyloid Fibrils from the Viewpoint of Protein Folding." Cellular and Molecular Life Sciences (CMLS) 61.5 (2004): 511-24. Web.

contrast, the core structure of non-amyloid fibril proteins have primary interactions which many at times involve the side chains over the main chain. However, although each polypeptide chain has the ability to form amyloid structures, the propensity to do so is determined by a variety of factors, which are particular to each chain, including, but not limited to charge, secondary structure propensities and hydrophobicity. Furthermore, even though the ability to form amyloid structures is an intrinsic physico-chemical quality common to all proteins, as we discussed earlier, extrinsic factors are equally important, and aggregation is also highly modulated by cellular environmental conditions. Cellular conditions which favor formation of fibrils are those in which also cause proteins to be partially unfolded, such as low pH<sup>24</sup>.

The first traces of prion diseases, dates back to the Fore tribe of Papua New Guinea in the 1950's. Through a cultural practice called funerary cannibalism, members of the tribe ingested the dead bodies of other members. Especially by consuming the brain, where prions, misfolded proteins responsible for transmissible spongiform encephalopathies, were most prevalent, through which they passed on to other members. Although transmission of prion diseases is no longer most commonly passed on through the ingestion of the deceased, familial prion diseases inherited through genes cause protein misfolding diseases such as Familial Creutzfeldt-Jakob disease, Gerstmann-Sträussler-Scheinker syndrome, and fatal familial insomnia. Disease causing mutations do not cause misfolding by directly affecting the residues which make up the active sites, or other functional sites of the protein. Rather, most of the mutations involved in protein folding occur in the majority of residues which are important in creating a stable structure for the polypeptide chain. Single mutations affecting these structural residues alter properties such as charge, secondary structure propensities and hydrophobicity which determine the propensity to

---

24 Kelly J. Alternative conformation of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol* 1998;8:101–6.

aggregate<sup>25</sup>. Mutations affect folding in two main ways. The first is by impairing the folding process, resulting in decreased amounts of stable, functional protein. The other is by the mutations causing the protein to fold into stable, non-functional conformations which would aggregate and disturb normal processes<sup>26</sup>.

Studies have hypothesized that Familial Creutzfeldt-Jakob Disease could be affected by protein-misfolding related mutations in both ways. Fatal familial insomnia is a prion disorder of which 15% of cases are inherited in an autosomal dominant pattern. It is clinically characterized by insomnia with or without a diurnal dreaming state, hallucinations, delirium, and dysautonomia preceding motor and cognitive deterioration<sup>27</sup>. Interestingly, the prion protein is a predominantly  $\alpha$ -helical protein constituent of mammalian cells under normal conditions, but converts its structure into a  $\beta$ -sheet form into the stable amyloid fibril formation<sup>28</sup>. This can be explained by studies which show that many of the mutations associated with the familial deposition diseases increase the population of partially unfolded states, which as discussed are a key pre-cursor to the formation of amyloid fibrils<sup>29</sup>. Mutations in the PRNP gene which cause misfolding of the dominantly alpha helical regions into beta pleated sheets are hypothesized to be one possible cause for Familial Creutzfeldt-Jakob disease. Additionally, studies done in PrP-null mice, Collinge et al. (1994) also provided evidence that prion protein is necessary for normal

---

25 Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.

26 Bross, Peter, and Brage Andresen. "Protein Misfolding and Degredation in Genetic Disease." (2011): n. pag. Web.

27 <http://www.omim.org/entry/600072>

28 Ohnishi, S., and K. Takano. "Amyloid Fibrils from the Viewpoint of Protein Folding." *Cellular and Molecular Life Sciences (CMLS)* 61.5 (2004): 511-24. Web.

29 Dobson, Christopher M. "Principles of Protein Folding, Misfolding and Aggregation." *Seminars in Cell & Developmental Biology* 15.1 (2004): 3-16. Web.



synaptic function<sup>30</sup>, and the inherited mutations may be causing a post-translationally modified form of cellular PrPc, which ultimately leads to progressive loss of functional PrPc.

Cystic Fibrosis is another genetic disease which is well known to be associated with a mutation which causes overall loss of functional protein through misfolding. The disease is associated with symptoms involving the disruption of exocrine function of the pancreas, intestinal glands, biliary tree, bronchial glands, and sweat glands. A deletion of phenylalanine 508 ( $\Delta F508$ ) in the CFTR protein has been shown to cause the first domain of the protein to fold more slowly than the wild type domain does, leading to impaired folding of the CF transmembrane conductance regulator, and hence the degradation of the protein and overall loss-of-function<sup>31</sup>.

However, despite there being a number of diseases linked to protein misfolding, as discussed earlier, the reason most of our bodies function proficiently, and in general every possibility for mishap doesn't manifest itself as a disease is because our body has the inherent, built in mechanisms for dealing with problems if they go wrong, or to prevent them from going wrong in the first place. The term Proteosis is defined as the sum of all the genetic and environmental factors that safeguard the native conformation of the proteome and remove damaging, misfolded products<sup>32</sup>. Viable therapy for diseases of misfolded proteins could involve strengthening these existing proteosis mechanisms which our body has naturally developed to cope with problems, instead of building our own mechanisms of intervention. For example, it has been shown that the

---

30 Collinge, John, Miles A. Whittington, Katie C. L. Sidle, Corinne J. Smith, Mark S. Palmer, Anthony R. Clarke, and John G. R. Jefferys. "Prion Protein Is Necessary for Normal Synaptic Function." *Nature* 370.6487 (1994): 295-97. Web.

31 Lukacs, Gergely L., and A.s. Verkman. "CFTR: Folding, Misfolding and Correcting the  $\Delta F508$  Conformational Defect." *Trends in Molecular Medicine* 18.2 (2012): 81-91. Web.

32 Powers, Evan T., Richard I. Morimoto, Andrew Dillin, Jeffery W. Kelly, and William E. Balch. "Biological and Chemical Approaches to Diseases of Proteostasis Deficiency." *Annual Review of Biochemistry* 78.1 (2009): 959-91. Web.

introduction of co-factors involved in guiding and stabilizing the native structure could be administered by increasing expression levels to overcome the symptoms of many disorders simply by using inherent systems to our advantage<sup>33</sup>. One consideration to be taken into account is that overexpressing more stringent quality control systems may affect normal proteins in a negative way, however, the benefit to the defective protein may be enough to cause an overall gain in function.

Another key consideration in such therapies is the point of intervention in the course of the progression of the disease. With the existing healthcare system, such a therapy would be undertaken at times of stress such as at old age in which case there is a loss of such inherent protective systems in the body. Here, proteosis would be used as a way to overcome the loss to re-stabilize the body's original functions. But what if the concept of Proteosis was involved and undertaken as more of a preventative measure than a reactive measure. Following the idea of vaccinations, if we began to strengthen the natural defense system against a particular diseases before it is contracted, our body would not have to undergo the symptoms at all. Preventative healthcare currently incorporates genetic tests to analyze predisposition to disease, as well as lifestyle changes, but our preventative healthcare system should be just as aggressive as our reactive one. Patients with increased genetic predispositions to protein-misfolding diseases could begin more than just lifestyle changes, but instead begin therapy which would involve overexpression of the body's inherent protective molecular chaperones before the disease is contracted to decrease the chances of the patient become sick in the first place. Understanding the process of protein folding and misfolding further could give us more information to form

---

33Auluck, Pavan, and H. Y. Edwin Chan. "Chaperone Suppression of  $\alpha$ -Synuclein Toxicity in a Drosophila Model for Parkinson's Disease." *Science* 1 295 (2002): n. pag. Web

specific therapies that effectively combat the increasing number of cases of diseases associated with protein misfolding and aggregation.