Graph Representations of Oxidative Folding Pathways

Vilmos ÁGOSTON[1], Masa CEMAZAR[2,3] and Sándor PONGOR[2]
[1]*Bioinformatics Group, Biological Research Center, Hungarian Academy of Sciences, Temesvári krt. 626726 Szeged, Hungary*
[2]*Protein Structure and Bioinformatics Group, International Centre for Genetic Engineering and Biotechnology, Area Science Park, 34012 Trieste, Italy*
[3]*Current address: Institute for Molecular Bioscience, University of Queensland, St. Lucia 4072, QLD, Australia*

Abstract. Oxidative folding combines the formation of native disulfide bond with the conformational folding resulting in the native three-dimensional fold. Oxidative folding pathways can be described in terms of disulfide intermediate species (DIS) containing a varying number of disulfide bonds and free cysteine residues, which can also be – as opposed to the majority of protein folding states –isolated and experimentally studied. Each DIS corresponds to a family of folding states (conformations) that the given DIS can adopt in three dimensions. The oxidative folding space can be represented as a network of DIS states interconnected by disulfide interchange reactions reactions that can either create/abolish or rearrange disulfide bridges. Such networks can be used to visualize folding pathways in terms of the experimentally observed intermediates. In a number of experimentally studied cases, the observed intermediates appear as part of contiguous oxidative folding pathways.

Introduction

Levinthal's paradox, introduced in 1968 [1], stated that the folding of a protein would last more than the age of the universe, if it went through looking for the native conformation by adapting every single conformation possible. There have been many propositions regarding how the conformational space is restricted so that the folding time is reduced to the experimental range. We know today that most single domain proteins are able to fold effectively *in vitro* to their native folds within seconds. The obvious flaw in stating the paradox itself is actually the fact that the search for the native conformation is unbiased with no stabilisation of particular conformations. Today it has been widely accepted that the native state is the energetically most favourable one on the *potential energy surface*. Actually, each conformational state of the protein assumes a certain position on this surface, which means that not all states are equal in free energy and hence the search for the native fold cannot be unbiased. The way this view has evolved to form theories about folding pathways is the following. Already Levinthal stated that there exist specific pathways for folding. By restricting the molecules to those pathways the polypeptide chain does not need to undergo an extensive search of all the conformational space. In 1973 Anfinsen proposed that the information coded in the amino acid sequence of a protein completely determines its folded structure and that the native state is the global minimum of the free energy [2]. Later, a variety of theories emerged, for example the framework model, the diffusion-collision model, the nucleation model, the hydrophobic-collapse model and the jigsaw model. The hydrophobic-collapse and the framework models were

favoured over the nucleation model, because they imply the existence of folding intermediates, which were discovered soon after. All proposed mechanisms and models were able to explain particular pieces of experimental data, but none provided a clear explanation of the folding principles or a solution to Levinthal's paradox (for a collection of reviews see: [3]).

The current, unified view of protein folding presented in some highly cited reviews by Dobson and co-workers [4,5], underlies the fact that protein folding is a progression in which both native and non-native contacts stabilise native-like structural features. The folding either proceeds through a hydrophobic collapse to a compact globule that has stabilising interactions or through a slow formation of a folding core (nucleus), which then rapidly proceeds towards the native state. Folding is thus seen as a step-wise behaviour, sampling regions of the landscape that are downhill in energy. An important element in the "new view" of protein folding is the folding funnel, which was first introduced by Onuchic and associates [6]. This is one way of representing the folding landscape with the free energy (enthalpy and entropy) as a function of folding progress variable, also known as the fraction of the native contacts. In the light of this simple surface (see Figure 1), it is possible to understand a number of features of the folding process. There are three kinds of states that can be easily distinguished in the folding funnel.

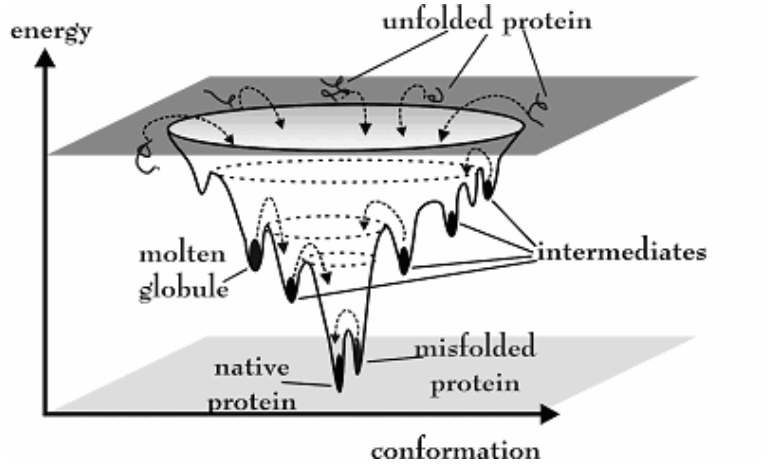


Figure 1. Schematic representation of the energy landscape of protein folding. The energy of a protein is displayed as a function of the topological arrangement of atoms. Adapted from Cemazar [7].

The initial state from which the folding proceeds is extremely heterogeneous and encompasses a large conformational space of rapidly inter-converting states. It seems generally accepted that the unfolded or denatured states are not completely random as one would expect for a theoretical polymer. On the contrary, it has intrinsic propensities for native and non-native like interactions, which funnel the folding process either through global or local conformational preferences. Compact denatured states, commonly known as molten globules, are lower in energy in the folding funnel. These have been in the past defined with a set of well-defined features such as a set of secondary structural elements in the absence of tertiary structure. In contrast, at the bottom of the funnel we find a highly compact state, where the close packing of the side chains is essential for a well-defined conformation. This is the so-called native state [4,5].

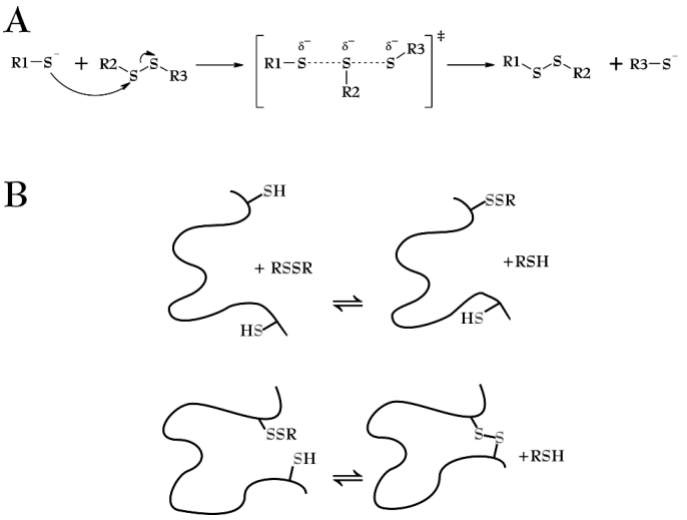


Figure 2. A. Thiol-disulfide exchange mechanism: in the pH range above 8, cysteine thiols are readily converted to thiolate anions (RS^-), which are potent nucleophiles. RS^- anions attack a disulfide bond, displacing one sulfur atom and forming a new bond with the other sulfur atom (nucleophilic substitution). The rate-determining step of this concerted process is the formation of a transition state with a partial transfer of the negative charge (δ^-) over the three sulfur atoms. B. The formation of a disulfide bond on the polypeptide chain (solid curve) with the help of a small molecule reagent (thiol form: RSH, disulfide form: RSSR). The two steps both proceed via a thiol-disulfide exchange reaction. The first step shown is intermolecular and the second intramolecular. The rate of the intramolecular step is relevant to protein folding, since it also involves conformational changes.

The particular kind of folding that this article is concerned with is oxidative folding, which is the fusion of native disulfide bond formation with conformational folding. This complex process is guided by two types of interactions: first, non-covalent interactions giving rise to secondary and tertiary protein structure, and second, covalent interactions between cysteine residues, which transform into native disulfide bridges. The process of disulfide formation is a simple chemical reaction in which two SH groups join to form a disulfide link (Figure 2A). If the SH groups are on a polypeptide chain, the *in vitro* reaction can be promoted by an external redox system such as a mixture of oxidized and reduced glutathione, or cysteine and cystine, respectively. *In vivo,* the oxidative power comes from specific agents such as the molecular chaperones protein disulfide isomerases. The underlying mechanism is disulfide interchange (Figure 2B). There are two kinds of reactions: in a *redox reaction* a protein disulfide bond is created (or abolished), i.e. the oxidative state of the polypeptide is changed. This is the case when one of the participants of the reaction (say RSH) is not part of the protein. In a *shuffling reaction* both participants of the disulfide interchange are protein-bound, so the oxidative state of the polypeptide does not change. In view of these possibilities it becomes obvious that there are a great many ways in which disulfide bridges can form and rearrange during the folding process. Today it is generally accepted that non-covalent interactions guide the process of folding and formation of disulfide bridges will lock the protein into the right conformation. The advantage of oxidative folding as opposed to general protein folding is that disulfide intermediates can be chemically isolated and studied using such techniques as acid trapping of the intermediates and analysis of the disulfide bridges using a combination of enzymatic cleavage and mass spectrometry. There is a body of literature in describing the pathways of oxidative folding in terms of disulfide intermediates [8-10], and our goal is show how graph theory can be used for this purpose.

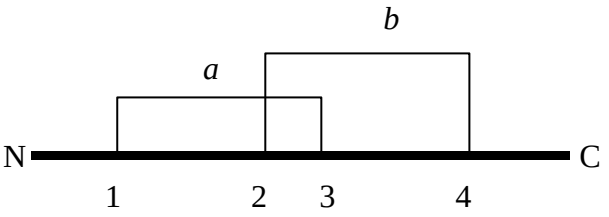
Graph theory has been applied to many aspects of protein research (for a review see [11]). Applications to protein folding followed two broad approaches. First, protein structure itself can be considered as a graph consisting of various interactions (such as covalent bonds, hydrogen bonds, spatial vicinities, contacts etc.) as edges, the nodes being atoms or residues of the protein. It was found, among others, that the so-called contact order, i.e. the average sequence distance between residues in atomic contact, seems to be a key determinant of folding speed [12]. Another line of research concentrates on characteristic networks of interatomic contacts that may form stabilization centres in protein structures and can be the reason of the stability of various proteins [13,14]. It was found that populated conformations seen in molecular dynamics simulations contain characteristic networks of residues [15,16].

Another line of research was triggered by the finding that the robustness and stability of networks may be the result of simple topological properties that are invariant throughout various technical as well as biological systems including social organization, electrical networks, road networks and the Internet [17]. In the following years the network topology of a large number of systems have been described, and it was found that some topology classes, like those characterized by a scale-free distribution of the degree (number of links at each node), or the so called "small world models" that are characterized by densely connected subnetworks loosely linked between each other, are indeed found in various systems within and without biology (for a review see e.g. [18]). The various network types were described in terms of a number of simple measures borrowed from graph theory, such as the clustering coefficient, the diameter of the graph etc. This approach was later extended to descriptions of the entire folding space, using the folding states as nodes, and transitions as links between them. As the folding states of native systems cannot be readily studied by physical methods, the investigations were first directed to model systems. Scala and associates [19] described the folding states of short peptides using Monte Carlo simulation on lattice models. They found that that the geometric properties of this network are similar to those of small-world networks, i.e. the diameter of the conformation space increases for large networks as the logarithm of the number of conformations, while locally the network appears to have low dimensionality. Shahnovitch and co-workers analysed the folding states of proteins during molecular dynamics simulations. It was found that the folding space is reminiscent of scale-free network, characterized by a majority of less populated states as well as some highly populated states reminiscent of "hubs" seen in other systems [20].

Our purpose is to describe the folding space of the oxidative folding process using graph theory. This is an intriguing task since the number of folding states defined in terms of disulfide links is relatively small, as compared to "ordinary" folding. We will approach the problem in two steps: i) using graph theory to describe the disulfide intermediates, and to enumerate the states of the folding space. ii) using a graph-like representation of the folding space to visualize the experimentally studied folding pathways.

1. Graph representation of oxidative folding intermediates

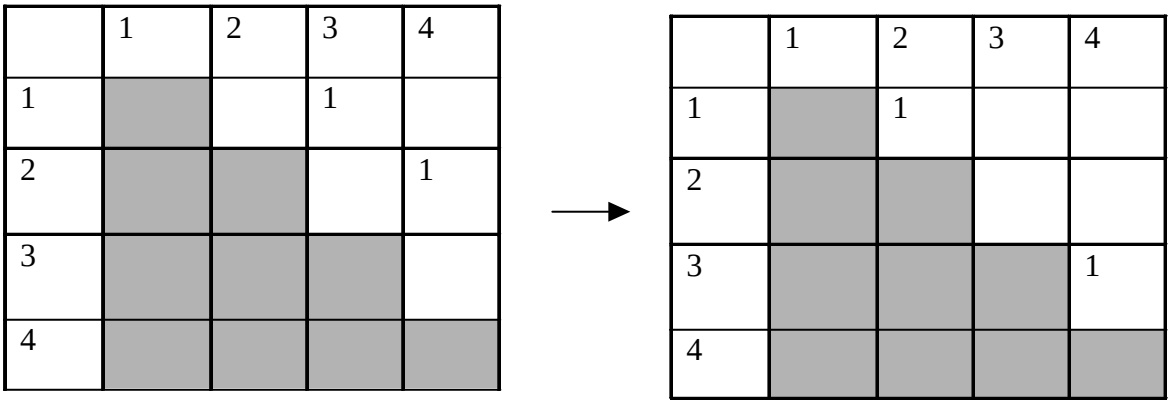
In proteins containing disulfide bonds, usually all cysteines form part of disulfide bridges, and the disulfide topology can be unequivocally described by defining which cysteines are connected. For example, a topology 1-3, 2-4 means that a protein with 4 cysteines has two disulfide bridges that connect cysteines (1,3) and cysteines (2,4) respectively. Cysteines can be labelled by their sequence position, or – as in the previous example – in a serial order from the N-terminus (Figure 3).



1-3, 2-4 or *abab* topology

Figure 3. Nomenclature for disulfide topologies. Disulfides can be labeled by the sequence positions, or simply by the sequential number of the cysteine residues they connect (*1-3, 2-4* topology). Alternatively, it is customary to alphabetically label the disulfide bridges, and describe the topology by assigning the bridge label to the cysteines, starting from the N terminus (*abab* topology).

The number of fully connected (disulfide bonded) isomers in a protein chain with n disulfide bonds ($2n$ cysteines) can be deduced from simple combinatorial considerations as $(2n)!/(n!*2^n)$. According to this formula proteins with two disulfide bridges have 3 fully oxidized isomers, 3-disulfide proteins have 15 and 4-disulfide proteins have 105. In other words, the number of intermediates increases very fast as a function of the number of constituent cysteines, and it has been hypothesized that the reason why the number of cysteines in autonomously folding protein domains is not very large is because the too high number of possible intermediates would slow down the folding process.



1-3, 2-4 1-2, 3-4

Figure 4 Adjacency matrices of two disulfide topologies of a peptide with two disulfide bridges

For a complete description of the folding process we have to consider both fully oxidized intermediates and the ones with free cysteine residues. For this purpose we will use a formal description of the intermediates as (undirected) graphs, with cysteines as nodes and disulfide bridges as edges (the main chain will not be represented). For the majority of naturally occurring protein structures the resulting graphs will be extremely simple especially if described as an adjacency matrix. Such an adjacency matrix is symmetrical, and contains 1 if two cysteines form a disulfide bonds and zero otherwise. As one cysteine can form only one disufide bridge, each column and each row of the resulting matrix will have atmost one value of 1. The adjacency matrix of two disulfide topologies of a 2-disulfide proteins are shown in Figure 4.

2. Description of the oxidative folding space as graphs

The graph descriptions introduced above can be applied both to fully and to partially oxidized intermediates, and the transitions between them can be conveniently described by comparing the adjacency matrices of the two states. The sum of the elements in the i-th column plus the i-th row ($S_i = \sum_j A_{ji} \sum_j A_{ij}$) shows if the i-th cysteine forms a bridge. The sum of the differences calculated between these measures of two adjacency matrices describing two intermediates, ($SD = \sum_i S_i$) shows how many cyesteins gained or lost a pair. If two states are connected by a disulfide interchange reaction, the number of disulfide bridges NB remains the same by definition, and it is easy to show that SD will differ exactly by 2. For redox steps in which one disulfide bridge is established or lost, NB and SD will increase or decrease by one and two, respectively. On the above basis one can easily enumerate, for a protein with any number of cysteine residues, a) the oxidative folding states and b) the possible transition steps between them. In other words one can draw a network of all possible oxidative folding pathways. The characteristics of a few systems are summarized in Table 1.

Table 1. Number of possible intermediates in and graph parameters of oxidative folding networks.

N of cysteines	N of intermediates (nodes)	Redox transitions	Shuffling transitions	Total no of transitions (edges)	Clustering coefficient C	Average path length
1	1	0	0	0	1.000	0.000
2	2	1	0	1	1.000	1.000
3	4	3	3	6	1.000	1.000
4	10	12	12	24	0.400	1.467
5	26	40	60	100	0.410	1.810
6	76	150	240	390	0.247	2.293
7	232	546	1050	1596	0.253	2.640
8	764	2128	8736	10864	0.181	3.149
9	2620	8352	19152	27504	0.182	3.550
10	9496	34380	83520	117900	0.142	3.977

The results show that on one hand, the clustering coefficient of the system decreases while on the other, the average path length increases with the number of cysteines. Both findings are consistent with the view that the folding space of peptides with many cysteines may be too complex and thus the systems may be unable to fold fast enough.

The pathways can also be graphically represented, and in order to simplify the resulting picture, we chose a 3D representation wherein the states having the same number of disulfide bridges are placed on separate planes. In this representation, the shuffling transitions are within the planes, and the redox edges connect adjacent planes.

It is noted that the experimental methods do not reveal all possible intermediates; some of them may be too short-lived or not abundant enough so as to be noticed an isolated. In spite of these limitations, the folding pathways appear as connected subgraphs within the network of all possible intermediates, showing that the experimental techniques actually identified states that can interconvert into one another. Only in EGF do we see an "isolated" intermediate which suggest that some intermediates of the pathway were not observed experimentally.

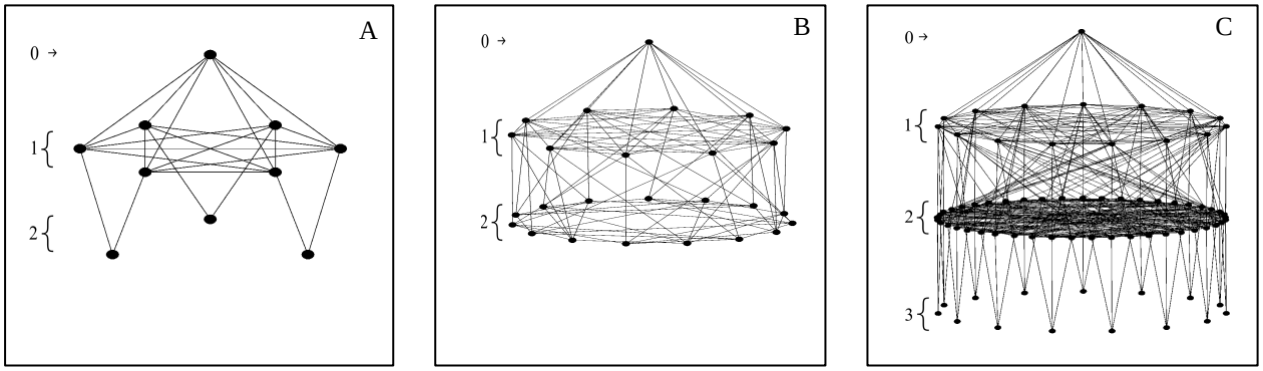


Figure 5. Three dimensional representation of the oxidative folding space of polypeptides with 4,5 and 6 cysteine residues (A, B and C, respectively). The nodes represent intermediates, the number of disulfide bridges is indicated with numbers on the left of each panel. The edges indicate disulfide exchange transitions. Zero indicates the fully reduced state, nodes in the lowest plane are the fully oxidized intermediates, one of which is the native state. Edges within the same plane indicate shuffling reactions (interchange between two protein-bound disulfides), edges between planes are redox transitions in which a disulfide bridge is created or abolished.

The network representations shown in Figure 3 are three-dimensional representation of the entire oxidative folding space described in terms of chemically well-defined disulfide intermediates. Species with the same number of disulfide bridges are placed on the same plane, so shuffling reactions, which do not change the number of disulfide bridges are represented as edges within the same plane. On the contrary, reactions in which a disulfide bridge is gained or lost, are represented as edges between two neighbouring planes. The fully reduced state (zero disulfide bridges) is on top, the fully oxidized species, on of which is the native state, is on the bottom. Panel B shows a peptide with 5 cysteines, such as granulocyte-colony stimulating-factor [21, 22] in which the native state contains one free cysteine residue that is not part of a disulfide bridge. In this case the native state can in principle rearrange into other species, so there are shuffling edges also in the lowest plane in the figure. In most of the known cases, the number of cysteines is an even number, so the fully oxidized DISs cannot readily interconvert into each other. In some cases this might be an obstacle: the propeptide of BPTI contains an additional free cysteine that seems to facilitate the folding of the molecule. The propeptide is subsequently cleaved and in this way the structure is locked into the native disulfide configuration [23]. The oxidative folding pathways can be pictured as routes within the full network, starting at the fully reduced species and ending at the native state. In the literature there are a few well-studied examples in which folding intermediates have been determined. Three examples, bovine pancreatic trypsin inhibitor, insulin-like growth factor and epidermal growth factor are shown in Table 2 and Figure 6.

BPTI's folding pathway was the subject of an intense dispute in the early 1990's, but later resulted in one of the most extensively studied oxidative folding pathways and a major protein folding model. With some differences, BPTI's pathway was characterised with the predominance of only a limited number of folding intermediates that adopt mainly native disulfide bridges and native-like structures. It is important to remember that 1- and 2-disulfide intermediates were present, but no 3-disulfide species apart from the native protein was detected on this pathway. One of the most abundant intermediates is a two disulfide species with two native disulfide bonds and a native-like structure. Formation of the third disulfide (Cys14-Cys38) is the last step of the folding process. A prevalence of the native-like structures and native disulfide bridges points to the conclusion that non-covalent

Table 2. Disulfied intermediates experimentally observed in the oxidative folding of various proteins

Protein	Disulfide intermediates[1]	Ref.
Bovine pancreatic trypsin inhibitor (BPTI)	3-5; 1-6; 3-5, 1-2; 3-5, 1-4; 3-5, 2-4; 1-6, 2-4; 3-5, 1-6; **1-6, 3-5, 2-4**;	[24,25]
Insulin-like growth factor (IGF)	2-6; 2-6, 3-5; 2-6, 1-4; 2-6, 4-5; 2-6, 1-3; 2-6, 1-3, 4-5; **1-4, 2-6, 3-5;**	[26-28]8]
Epidermal growth factor (EGF)	2-3; 1-2; 4-6; 5-6; 3-4; 2-4, 5-6; 2-5, 3-4; 1-6, 2-5, 3-4; 1-2, 3-4, 5-6; **1-3, 2-4, 5-6;**	[29]

[1]The intermediates are described with the notation given in Figure 3. The native disulfide connectivity is given in bold, the fully reduced species is not explicitly included.

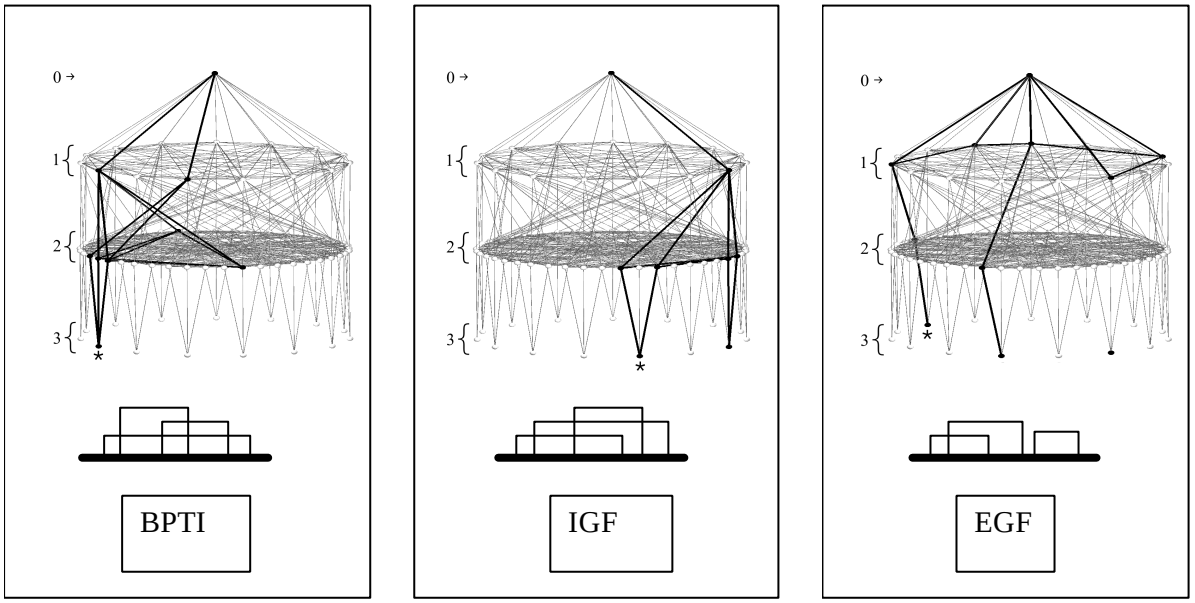


Figure 6. The oxidative folding pathways of bovine pancreatic trypsin inhibitor (BPTI), insulin-like growth factor (IGF) and epidermal growth factor (EGF). The native state is marked by asterisk.

interactions that are specific to the amino acid sequence can guide the initial stages of the folding process and hence admit a very limited number of disulfide species on the pathway.

Oxidative folding of the fully reduced EGF [29] proceeds through 1-disulfide intermediates and accumulates rapidly as a single stable 2-disulfide intermediate (designated as EGF-II), which represents up to more than 85% of the total protein along the folding pathway. Among the five 1-disulfide intermediates that have been structurally characterized, only one is native, and nearly all of them are bridges by neighbouring cysteines. Extensive accumulation of EGF-II indicates that it accounts for the major kinetic trap of EGF folding. EGF-II contains two of the three native disulfide bonds of EGF, Cys(14)-Cys(31) and Cys(33)-Cys(42). However, formation of the third native disulfide (Cys(6)-Cys(20)) for EGF-II is slow and does not occur directly. Kinetic analysis reveals that an important route for EGF-II to reach the native structure is via rearrangement pathway through 3-disulfide scrambled isomers. Epidermal growth factor (EGF) [29] forms

both non-native three-disulfide isomers as well as a predominant species with two native disulfides (EGF-II).

3. Conclusions, scope and limitations

The graph representations described here provide a simple method to visualise folding pathways as studied by experimental methods.

The picture emerging from these representations confirm that the folding pathways of oxidative folding are contiguous routes that connect the fully reduced state to the native state. If we try to reconcile this picture with the three-dimensional energy landscape of oxidative folding, the energy of the protein will be a function of which disulfide bonds are present and of the extent of conformational folding (Figure 7). The protein molecules will have folded successfully when they reach the lowest energy point, which represents the native species, both in terms of disulfides and conformation. The non-native disulfide intermediates lie in the local energy minima, from where they have to be re-activated to reach the native state. This picture suggests a qualitative explanation for the observation that non-native disulfide intermediates can be the necessary steps of the folding pathways. The cited case of pro-BPTI is an indirect proof for that. In the analysis of a small 3-disulfide peptide, AAI we found that a disulfide intermediate with no native disulfide bonds is in fact the most abundant species [30,31].

The current approach is limited by the fact that the 3D images of oxidative folding pathways cannot be generated fully automatically. (A drawing application that uses the Tulip package (www.tulip-software.org) is available from VA, vilagos@nucleus.szbk.u-szeged.hu). A further plausible improvement would include colouring of the folding states by quantitative properties and look for correlations between the coloured areas of the network and the experimentally determined folding pathways.

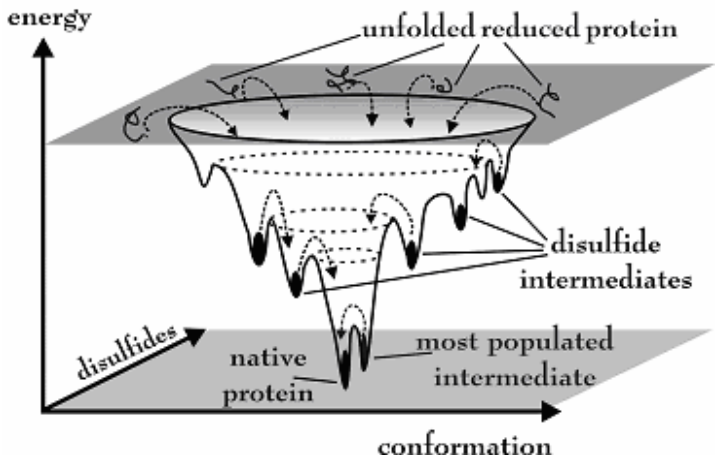


Figure 7. The energy landscape of oxidative protein folding [7]. The energy of the protein is displayed as a function of which disulfide bonds are present and the extent of conformational folding. The local minima represent non-native disulfide intermediates, which are kinetic traps.

References

[1] Levinthal C. Are there pathways in protein folding? J. Chim. Phys. 1968;65:44-5.
[2] Anfinsen CB. Principles that govern the folding of protein chains. Science 1973;181:223-30.
[3] Pain RH. Mechanisms of Protein Folding. Oxford, New York: Oxford University Press, 2000.
[4] Dobson CM, Karplus M. The fundamentals of protein folding: bringing together theory and experiment. Curr Opin Struct Biol 1999;9 (1):92-101.
[5] 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 (7):331-9.
[6] Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG. Protein folding funnels: the nature of the transition state ensemble. Fold Des 1996;1 (6):441-50.
[7] Cemazar M. Oxidative folding of a cystine knot protein: the Amarathus alpha-amylase inhibitor. International Centre for Genetic Engineering and Biotechnology. Trieste, Italy: Open University, UK, 2003. pp. 130.
[8] Chang JY. Evidence for the underlying cause of diversity of the disulfide folding pathway. Biochemistry 2004;43 (15):4522-9.
[9] Wedemeyer WJ, Welker E, Scheraga HA. Proline cis-trans isomerization and protein folding. Biochemistry 2002;41 (50):14637-44.
[10] Welker E, Wedemeyer WJ, Narayan M, Scheraga HA. Coupling of conformational folding and disulfide-bond reactions in oxidative folding of proteins. Biochemistry 2001;40 (31):9059-64.
[11] Vishveshwara S, Brinda KV, Kannan N. Protein Structure: Insights from Graph Theory. Journal of Theoretical and Computational Chemistry 2002;1 (1):187-211.
[12] Plaxco KW, Simons KT, Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. J Mol Biol 1998;277 (4):985-94.
[13] Magyar C, Tudos E, Simon I. Functionally and structurally relevant residues of enzymes: are they segregated or overlapping? FEBS Lett 2004;567 (2-3):239-42.
[14] Selvaraj S, Gromiha MM. Importance of hydrophobic cluster formation through long-range contacts in the folding transition state of two-state proteins. Proteins 2004;55 (4):1023-35.
[15] Vendruscolo M, Paci E, Karplus M, Dobson CM. Structures and relative free energies of partially folded states of proteins. Proc Natl Acad Sci U S A 2003;100 (25):14817-21.
[16] Vendruscolo M, Paci E, Dobson CM, Karplus M. Three key residues form a critical contact network in a protein folding transition state. Nature 2001;409 (6820):641-5.
[17] Albert R, Jeong H, Barabasi AL. Error and attack tolerance of complex networks. Nature 2000;406 (6794):378-82.
[18] Dorogovtsev SN, Mendes JFF. Evolution of Networks: From Biological Nets to the Internet and Www (Physics). Oxford, New York: Oxford University Press, 2003.
[19] Scala A, Amaral LAN, Barthelemy M. Small-world networks and the conformation space of a short lattice polymer chain. Europhysics Letters 2001;55 (4):594-600.
[20] Dokholyan NV, Shakhnovich B, Shakhnovich EI. Expanding protein universe and its origin from the biological Big Bang. Proc Natl Acad Sci U S A 2002;99 (22):14132-6.
[21] Cantrell MA, Anderson D, Cerretti DP, Price V, McKereghan K, Tushinski RJ, Mochizuki DY, Larsen A, Grabstein K, Gillis S, et al. Cloning, sequence, and expression of a human granulocyte/macrophage colony-stimulating factor. Proc Natl Acad Sci U S A 1985;82 (18):6250-4.
[22] Werner JM, Breeze AL, Kara B, Rosenbrock G, Boyd J, Soffe N, Campbell ID. Secondary structure and backbone dynamics of human granulocyte colony-stimulating factor in solution. Biochemistry 1994;33 (23):7184-92.
[23] Weissman JS, Kim PS. The pro region of BPTI facilitates folding. Cell 1992;71 (5):841-51.
[24] Creighton TE. The disulfide folding pathway of BPTI. Science 1992;256 (5053):111-4.
[25] Weissman JS, Kim PS. Reexamination of the folding of BPTI: predominance of native intermediates. Science 1991;253 (5026):1386-93.
[26] Hober S, Uhlen M, Nilsson B. Disulfide exchange folding of disulfide mutants of insulin-like growth factor I in vitro. Biochemistry 1997;36 (15):4616-22.
[27] Milner SJ, Carver JA, Ballard FJ, Francis GL. Probing the disulfide folding pathway of insulin-like growth factor-I. Biotechnol Bioeng 1999;62 (6):693-703.
[28] Yang Y, Wu J, Watson JT. Probing the folding pathways of long R(3) insulin-like growth factor-I (LR(3)IGF-I) and IGF-I via capture and identification of disulfide intermediates by cyanylation methodology and mass spectrometry. J Biol Chem 1999;274 (53):37598-604.
[29] Chang JY, Li L, Lai PH. A major kinetic trap for the oxidative folding of human epidermal growth factor. J Biol Chem 2001;276 (7):4845-52.

[30] Cemazar M, Zahariev S, Lopez JJ, Carugo O, Jones JA, Hore PJ, Pongor S. Oxidative folding intermediates with nonnative disulfide bridges between adjacent cysteine residues. Proc Natl Acad Sci U S A 2003;100 (10):5754-9.
[31] Cemazar M, Zahariev S, Pongor S, Hore PJ. Oxidative folding of Amaranthus alpha-amylase inhibitor: disulfide bond formation and conformational folding. J Biol Chem 2004;279 (16):16697-705.