FEBS Letters

journal homepage: www.FEBSLetters.org

Review

Protein disulfide engineering

Alan A. Dombkowski [a,*], Kazi Zakia Sultana [b], Douglas B. Craig [a]

[a] Carman and Ann Adams Department of Pediatrics, Wayne State University School of Medicine, Detroit, MI 48201, USA
[b] Department of Computer Science & Engineering, Chittagong University of Engineering & Technology, Chittagong 4349, Bangladesh

ABSTRACT

Improving the stability of proteins is an important goal in many biomedical and industrial applications. A logical approach is to emulate stabilizing molecular interactions found in nature. Disulfide bonds are covalent interactions that provide substantial stability to many proteins and conform to well-defined geometric conformations, thus making them appealing candidates in protein engineering efforts. Disulfide engineering is the directed design of novel disulfide bonds into target proteins. This important biotechnological tool has achieved considerable success in a wide range of applications, yet the rules that govern the stabilizing effects of disulfide bonds are not fully characterized. Contrary to expectations, many designed disulfide bonds have resulted in decreased stability of the modified protein. We review progress in disulfide engineering, with an emphasis on the issue of stability and computational methods that facilitate engineering efforts.

© 2013 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Disulfide bonds in proteins

A protein disulfide bond is a covalent link between the sulfur atoms of the thiol groups (–SH) in two cysteine residues. The disulfide (also called an SS-bond, disulfide bridge, or crosslink) is formed upon oxidation of the two thiols, thus linking the two cysteines and their respective main peptide chains by the covalent disulfide bond. Conversely, a disulfide bond can be disrupted by a reductive reaction (e.g. using dithiothreitol).

$$R_1\text{-SH} + R_2\text{-SH} \leftrightarrow R_1\text{-S—S-}R_2 + 2H^+ + 2\ e^-$$

Disulfide bonds are found predominantly in secreted extracellular proteins. The redox environment within the cytosol preserves cysteine sulfhydryls in a reduced state. Disulfide bonds rapidly form outside of the cell in the presence of oxygen.

Most disulfide bonds in proteins secreted from eukaryotic cells are formed in the endoplasmic reticulum, which offers an oxidizing environment as well as chaperones and disulfide isomerases to ensure correct protein folding and disulfide connectivity [1].

In proteins, disulfide bonds are a configuration of six atoms, $C_\alpha—C_\beta—S_\gamma—S'_\gamma—C'_\beta—C'_\alpha$, linking two cysteine residues. The seminal work of Janet Thornton in 1981 characterized the features and bond geometry of disulfides by analyzing the atomic coordinates of 55 disulfide bonds that existed in protein structures available at the time [2]. Nearly twenty years later, disulfide bond features were further detailed by Petersen et al. using 351 disulfide bonds in 131 non-homologous protein structures [3]. These analyses revealed the distribution of bond angles and distances found in naturally occurring disulfides, and this work has provided the basis of most models for disulfide engineering. Two important bond angles are the $C\alpha\text{-}C\beta\text{-}S\gamma$ and $C\beta\text{-}S\gamma\text{-}S\gamma$ (Fig. 1), typically found near 114° and 105° respectively, although a range of values with clear peaks are observed. Petersen et al. reported the χ_3 torsion angle, which represents rotation of the $C\beta$ atoms about the S–S bond, to have peaks at +100° (right-handed) and −80° (left-handed) based on 351 disulfide bonds. Our own analysis, recently performed on 1505 native disulfide bonds, indicates that the χ_3 peaks are −87° and +97° ([69]). These values are slightly different from the often-cited ±90°. This torsion angle is critical to the stability of a disulfide bond, and deviations from optimal values can produce an energy strain by several kcal/mol [4,5]. The χ_1 torsion angle, defined by the N–$C\alpha$–$C\beta$–$S\gamma$ bonds, has peaks at ±60° and ±180°, and the χ_2 torsion angle, defined by the $C\alpha\text{-}C\beta\text{-}S\gamma\text{-}S\gamma$ bonds, has peaks at ±60°.

The contribution of disulfide bonds to the stability of proteins was investigated in early experiments that measured the change in protein stability upon elimination of one or more native disulfide bonds. By methodically disrupting the two native disulfide bonds in ribonuclease T1 it was demonstrated that each crosslink contributed ~3.5 kcal/mol to the thermodynamic stability of the

* Corresponding author. Address: Division of Clinical Pharmacology and Toxicology, Children's Hospital of Michigan, 3901 Beaubien Blvd., Detroit, MI 48201, USA. Fax: +1 313 745 5441.
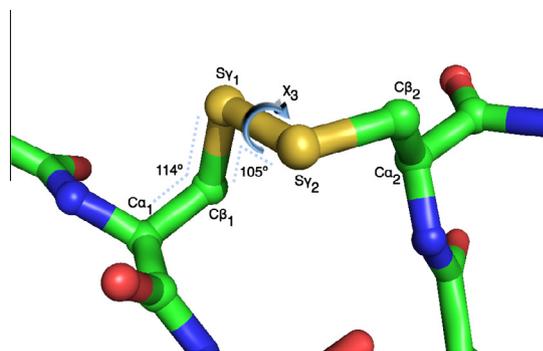E-mail address: domski@wayne.edu (A.A. Dombkowski).

Fig. 1. Disulfide bond geometry. Analysis of protein structures with native disulfide bonds reveals consistent geometric relationships that can be utilized to predict sites for disulfide engineering. The χ_3 torsion angle is a key geometric consideration for disulfide formation, with peaks near $-87°$ and $+97°$.

protein, consistent with a range of 2.3–5.2 kcal/mol reported in other experiments [6,7]. Furthermore, the increase in stability associated with each disulfide bridge was determined to be due to the loss of conformational entropy of the unfolded state. A disulfide bond restricts motion of the unfolded, random coil to a far greater extent than restrictions imposed upon the folded protein. Based on experimental evidence, the contribution of a disulfide bond to the change in conformational entropy of a protein was modeled by the expression:

$$\Delta S = -2.1 - \left(\frac{3}{2}\right) R \ln(n) \qquad (1)$$

where n is the number of residues in the loop formed by the crosslink, and R is the universal gas constant [6]. The model provides a reasonable approximation, consistent with experimental observations for proteins with simple two-state unfolding, i.e. those exhibiting reversible denaturation. Later work using "loop permutation analysis" of T4 lysozyme suggests that the sequence composition of the loop may also be a factor in determining the energy contribution of a disulfide [8]. Additionally, evidence indicates that an introduced disulfide bond can contribute to an enthalpic change in free energy [9–11]. The effect of a disulfide bond on stability is more challenging to predict in proteins that do not conform to a two-state model and unfold along intermediate states that act as a kinetic barrier.

2. Disulfide bond engineering

Given that native disulfide bonds provide considerable stability to proteins, it logically follows that the addition of a novel crosslink into a protein might provide an increase in stability. There are many biomedical and industrial applications where increased stability of a protein is beneficial. For example, it is desirable to increase the stability and half-life of protein-based therapeutics. Likewise, increasing the stability of industrial enzymes can greatly improve their yield and expand their operational range, including environments with temperature and pH unfavorable to unmodified enzymes. Engineered disulfide bonds have been widely used in attempts to improve protein stability, and these efforts have frequently been successful in creating novel disulfides. However, not all engineered disulfides produce an increase in stability. Contrary to expectations, many cases of destabilizing disulfides have been reported. The factors that determine if an engineered disulfide bond will increase or decrease stability of a protein are not well characterized, but some general guidelines can be drawn from disulfide engineering experiments. Recent computational modeling has attempted to further address this issue.

Early investigations to explore the stabilizing effect of engineered disulfides were performed using phage T4 lysozyme, subtilisin, dihydrofolate reductase, and λ repressor [12–14]. T4 lysozyme is a small, 164 amino acid protein having no native disulfide bonds and two unpaired cysteines. In 1984 Perry and Wetzel reported the introduction of a novel disulfide bond into T4 lysozyme between one of the native cysteines (Cys97) and an Ile3 → Cys mutant [15]. Criteria for selecting potential sites for insertion of a disulfide bond were: (1) "wide" linear separation along the primary sequence; and (2) a Cβ-Cβ distance ⩽5.5 Å. The former constraint is based on the concept that the entropic contribution of a disulfide is dependent on the length of the loop formed by the crosslink. Under oxidizing conditions the resultant Cys3–Cys97 disulfide bond significantly increased the half-life of T4 lysozyme activity after incubation at elevated temperatures. Under reducing conditions the mutant protein had the same stability as the wild type, demonstrating that the stabilizing effect was due to the disulfide crosslink and not the amino acid mutation from Ile to Cys.

The initial disulfide engineering success in T4 lysozyme was soon followed by a number of experiments that further characterized the effects of various engineered disulfides in the protein. The Matthews group found that the Cys21-Cys142 disulfide spanning the hinge-like active site cleft of the protein provided an increase in the melting temperature (T_m) of 11 °C [16]. This result and a series of engineered disulfides in T4 lysozyme by the same group led to the observation that disulfide bridges most effective in increasing stability are those: (1) located in flexible regions of the protein; and (2) create a large loop [17,18]. Flexibility is believed to allow the protein backbone to adjust and accommodate the disulfide crosslink with preferable geometry and without introducing strain. Investigating the effect of multiple disulfide bonds, a combination of three engineered disulfides concurrently introduced into T4 lysozyme was shown to increase the T_m by 23.4 °C, and the change was roughly additive of the individual contribution of each disulfide [19]. The Matthews group noted destabilizing crosslinks were associated with rigid structure. Examples of destabilizing disulfides in T4 lysozyme are Cys90–Cys122 and Cys127–Cys154, and disulfide formation in these mutants changed the T_m by −0.5 and −2.4 °C, respectively. Both are located within the most rigid region of the protein and created short loop lengths (32 and 27) in comparison to stabilizing mutants (94, 122, 155) [17].

3. Disulfide engineering and kinetic stability

In addition to altering thermodynamic stability, engineered disulfide bonds can disrupt the rate of folding and unfolding, and this may prove advantageous in efforts to increase protein stability. Considerable evidence indicates that disulfide bridges that inhibit early events in the unfolding process confer kinetic stability to the protein. The extensive experiments of the Fersht lab with mutants of barnase demonstrated that protein modifications to stabilize structural elements that unfold prior to or in the rate-limiting step of unfolding can reduce the rate of unfolding, thus providing kinetic stability [20]. This effect was noted in the protein folding simulations of Abkevich and Shakhnovich [21]. Another group reported that it is possible to stabilize a protein by immobilizing the region of the protein that unfolds first [22].

The concept of kinetic stabilization has been explored more recently through computational simulations of protein unfolding where results also show that the location of an engineered disulfide with respect to the order of unfolding of a protein determines if the disulfide bond will provide stability. Using the GeoFold algorithm, Ramakrishnan et al. modeled the unfolding of several proteins in which disulfide bonds had previously been engineered, including T4 lysozyme, barnase, and dihydrofolate reductase [23].

They found that disulfide bonds in regions that unfold early provide an increase in stability, while a disulfide located in a region that unfolds late in the unfolding process destabilizes the structure. Recently, a computational approach was utilized to identify potential sites for disulfide engineering in phytase using molecular dynamic simulations that revealed segments of the protein associated with temperature-induced unfolding [24]. The identified regions were then analyzed with a geometric model to determine specific mutation sites appropriate for disulfide formation. Several disulfide bonds were designed with this approach and resulted in increased thermodynamic stability and dramatically slowed inactivation kinetics. These reports illustrate the importance of kinetic considerations and why simplified energy models limited to two-state equilibrium thermodynamics may sometimes prove inconsistent with experimental measurements of the change in stability resulting from engineered disulfide bonds in proteins.

Ultimately, computational methods to predict the stabilizing effect of an engineered disulfide should consider both thermodynamic contributions (entropic and enthalpic) as well as kinetic alterations that affect folding and unfolding rates [25]. Consideration of kinetic effects is challenging but warranted in proteins where the folded and unfolded conformations are not in rapid equilibrium and stability is governed by kinetics, i.e. those with irreversible denaturation indicative of transition barriers. A considerable number of native proteins appear to be stabilized kinetically, although the portion of the proteome that fall into this category is unclear [26,27]. An excellent review on protein kinetic stability is available, including discussion of laboratory methods to determine kinetic properties [26].

4. A focus on flexibility

Numerous disulfide engineering reports have followed the early T4 lysozyme experiments, and many have characterized the resultant change in protein stability. These reports have enabled comparison across a diverse set of proteins to further define features that may predict if an engineered disulfide bridge will stabilize or destabilize the target protein. Dani et al. performed an analysis of previously engineered disulfide bonds where the effect on protein stability had been characterized and published, and the wild type crystal structure was available in the Protein Data Bank (PDB) [28]. A total of 47 disulfides in 24 protein structures were included in the study. Each novel disulfide bond was classified as stabilizing (30), destabilizing (14), or no change (3). The authors investigated the relationship between several structural features of each disulfide bond and the change in stability. In summary they reported: (1) stabilizing mutations were most often found in regions of medium to high mobility; (2) stabilizing mutations were more likely to be near the protein surface; (3) stabilizing mutations were associated with longer loop lengths (>25 residues); and (4) the introduced disulfide bond should not cause steric overlap. The authors suggested that inclusion of these criteria in disulfide engineering would improve the likelihood of a stabilizing mutation.

A recurring theme noted in disulfide engineering experiments is the observation that engineered disulfide bonds linking regions of relatively high mobility are those most likely to confer stability to the protein. The meta-analysis of Dani et al. substantiates what was reported by the Matthews group more than 20 years ago based on their T4 lysozyme experiments: disulfide bonds that crosslink flexible regions of the protein are most likely to increase stability [17]. It should be emphasized that this general rule is not entirely accurate, reflecting limited knowledge of the physicochemical determinants of the stabilizing effect of a novel disulfide bridge. However, a number of studies provide additional support for the relationship. It is unclear if the association between flexible regions

and stabilizing disulfide bonds is because these regions can shift to accommodate the inserted crosslink or because the regions represent structural elements involved in the early stages of unfolding. Evidence suggests that both explanations may be valid. Disulfide engineering experiments with T4 lysozyme have demonstrated a considerable backbone shift in response to a disulfide inserted into a flexible region [29]. In addition, a number of studies have stabilized proteins by crosslinking a region of the structure involved in local unfolding [24,30–33]. The approach of targeting flexible regions with engineered disulfide bonds to improve protein stability is discussed as a component of the recently reviewed concept of rigidifying flexible sites (RFS), including a useful summary of computational methods to identify flexible regions in a protein [34].

A recent disulfide engineering experiment to improve the stability of *Candida antarctica* lipase B (CalB) effectively utilized the association between flexible regions and stabilizing crosslinks [35]. Lipase B is an important industrial enzyme used in many applications, including use as a catalyst in biodiesel production. It is desirable to improve the thermal stability of CalB to enable increased enzymatic rates at higher temperatures. The authors used two geometry-based computational methods to select potential sites for insertion of a disulfide bond into the protein, and then focused on regions with high mobility as reflected by the crystallographic B-factor. The B-factor (temperature factor, units of Å^2) is a measure of the uncertainty of electron density for atoms in a protein structure determined with X-ray crystallography [4]. It reflects atomic localization and can reveal dynamic motion as well as disorder in the crystal lattice. A high B-factor indicates low electron density (smearing) and low confidence in atomic position, often associated with regions of protein flexibility.

In the CalB study, potential disulfides that satisfied geometric constraints of the computational models were then ranked by the sum of the B-factor for the two residues involved in each crosslink. Corresponding mutants were individually created for the top four disulfides and one lower ranking candidate. The disulfide that produced the greatest improvement in stability (A162C-K308C) was the bond linking the residue pair having the greatest B-factor sum. This disulfide improved the enzyme half-life more than four fold. Fig. 2 shows the structure of CalB and the location of the A162C-K308C crosslink. An interactive model is also available in the Supplementary material. The inserted disulfide bond links a peripheral domain having relatively high mobility to the core
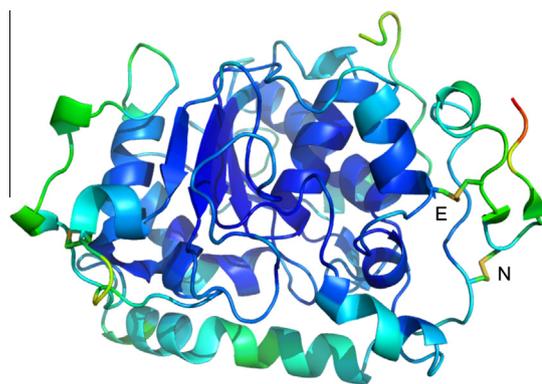

Fig. 2. An engineered disulfide bond (E) increases the thermostability of lipase B. The disulfide was inserted into a region of relatively high mobility as determined by B factors. The mutant protein structure was created with Disulfide by Design 2.0, based on the wild type structure (PDB: 1TCA) and the A162C-K308C disulfide bond designed in Le et al. [35]. One of three native disulfide bonds is also indicated (N). The protein is colored by the crystallographic *B*-factor. Blue represents a low *B* value (low mobility), with green and red indicating higher mobility. Graphical rendering was performed with PyMol v1.5. A 3D model is available in the Supplementary material.

region of the protein. Using molecular dynamics simulations to model the rigidifying effect of the inserted disulfides, the authors determined that the change in flexibility in the regions linked by each bond was anti-correlated to the change in stability. In other words, greater reduction of motion due to the disulfide bond resulted in a larger increase in stability. However, this observation is counter to the report of Pjura et al. who noted that a stabilizing disulfide inserted into a flexible region of T4 lysozyme did not result in additional rigidity of the structure [29].

While the B-factor can reveal flexible protein regions, there are some limitations to this approach. Because of differences in resolution, refinement procedures, and crystal contacts, B-factors can vary greatly between protein structures [4,36–39]. Therefore, it is preferable to use relative or normalized B-factors that reveal mobility within a target protein. Additionally, static crystallographic B-factors may not indicate structural regions prone to early unfolding. This is evident from disulfide engineering experiments that used molecular dynamics simulations to reveal protein regions of high mobility that were not apparent from crystallographic B-factors [30,33].

Limited proteolysis may be useful in protein engineering experiments by helping to identify protein regions involved in early unfolding or having high mobility. Studies using limited proteolysis have demonstrated that protein regions of local disorder and early unfolding can be identified with proteolytic probes [40–42]. Sites of limited proteolysis are well correlated with sites of high flexibility, as measured by the protein B-factor. Limited proteolysis, perhaps coupled with mild denaturation, might be a feasible strategy to identify regions involved in the initial steps of unfolding for the purpose of selecting candidate sites for kinetic stabilization. Such regions could then be targeted using computational methods to identify optimal residue pairs for mutation to cysteine (methods discussed below).

The analysis of Dani et al. suggested that in addition to targeting regions of high mobility, as indicated by the B-factor, stabilizing disulfide bonds were also likely to occur close to the protein surface. However, residue burial is correlated with residue flexibility. An analysis of more than 100 protein structures reported a bimodel distribution of normalized B-factors, with low values corresponding to buried residues [36]. Conversely, high B-factors were associated with surface exposed residues. Other studies have found a strong correlation between residue depth and the B-factor [8,37]. Fig. 2 is illustrative of this relationship. The periphery of the protein has higher B-factors, indicated by green, yellow, and red coloring, while the core region of the protein is predominantly blue reflecting low B-factors.

5. Computational methods for disulfide engineering

Early disulfide engineering studies often used simple distance criteria to identify potential residue pairs for mutation to cysteines with the goal of creating a novel disulfide bond. Growing interest in disulfide engineering spurred the development of computational methods to assist with site selection. The effort was greatly assisted by the rapidly growing repository of protein structures in the PDB. Most computational methods for disulfide engineering rely on a geometric model of native disulfide bonds. The general paradigm is to analyze atomic coordinates of a large set of native disulfide bonds to establish geometric parameters for the location and orientation of cysteine atoms involved in disulfide crosslinks. From these parameters an algorithm then assesses the atomic coordinates of each possible residue pair in a target structure to determine if they meet the geometric criteria for disulfide bond formation, assuming they were mutated to cysteines. Typically, a scoring or ranking method is provided to identify the best candidates.

Perhaps the earliest computational method for disulfide engineering was Proteus, developed by Pabo and Suchanek in 1986 [43]. Using backbone coordinates, the authors derived a database from 90 native disulfide bonds. The atomic coordinates of prospective residue pairs in the target protein were compared to the database to determine if they conformed to native geometry, determined by root-mean-square deviation (RMSD). Proteus was limited by the restriction of requiring that the mainchain atoms of the residues in a potential disulfide must align with a relatively small number of known examples from the PDB. To overcome this limitation Hazes and Dykstra devised the SSBOND algorithm [44]. The method creates Cβ and Sγ atoms for each residue in the target structure, starting from mainchain atomic coordinates. Multiple Sγ positions are considered based on rotation around the Cα–Cβ bond. The algorithm determines if the two Sγ from a residue pair are properly oriented to form a disulfide bond. An energy minimization was included to optimize putative bond lengths and angles, and an energy value was derived for each prospective bond. A similar algorithm based on native disulfide geometry (MODIP) was created by Sowdhamini et al. [45]. MODIP provides a grade (A, B, C) for each candidate disulfide, with "A" indicating the best geometry. The MODIP model was later refined using data from a larger set of PDB structures [28]. The newer model also includes consideration of steric interactions between the modeled cysteines and surrounding residues. Additionally, the rating of candidate disulfides was modified to provide a discrete value from 1 to 10, with 10 representing the best grade. The MODIP algorithm has been used successfully in wide range of applications.

Disulfide by Design (DbD) was created based on computational methods developed for protein structure prediction [46,47]. Using a geometric model derived from native disulfide bonds, the algorithm very accurately and rapidly estimates the χ_3 torsion angle based on the Cβ–Cβ distance. A test of the method achieved an R^2 value of 0.995 between predicted and actual angles, and the chirality is predicted correctly for 96% of disulfide bonds. Based on the wide range observed in native disulfides, the Cα-Cβ-Sγ angle is allowed some tolerance in DbD, whereas this value is fixed in the above algorithms. DbD calculates an energy value for each prospective disulfide to enable ranking, and mutant PDB files can be created for selected disulfides. DbD 2.0 was recently implemented with a number of new features, including an updated energy model based on an expanded set of PDB structures ([69]). The new application is web-accessible and provides graphical tools for secondary and tertiary structural analysis of prospective disulfides (http://cptweb.cpt.wayne.edu/DbD2/). These include the ability to analyze B-factors of candidate disulfide regions to help identify those that may confer increased stability to the protein.

The value of a directed, computational approach to disulfide engineering is evident in a report of a series of engineered crosslinks in alkaline phosphatase [48]. Five potential disulfides were identified using a simple distance approach, and three disulfides were identified using a computational method that considers side chain and bond geometry. There was one disulfide common to the two selection methods. After the corresponding site-directed mutations were made, all three of the disulfides predicted by the computational methods were confirmed with experimental methods. Only two of the five disulfides identified by simple distance approach were confirmed, including the one bond predicted by both methods.

The error rate of computational methods for disulfide engineering is not well characterized. An accurate estimate of the false positive rate is difficult to determine, as it would require a large number of predictions over a wide range of proteins, followed by experimental confirmation. In an attempt to estimate error rates of our algorithm, we performed an informal survey of authors of publications that used Disulfide by Design to engineer disulfide

bonds [46]. Of 41 laboratories contacted, we received eleven responses describing twelve projects. One additional group had reported the details in a publication [49]. We define "success" as an engineered disulfide bond that has been experimentally confirmed to form within the target protein. Seven projects achieved 100% success out of all attempted disulfides, four had ∼50% success, and two groups reported between 30 and 40% success. The lowest rate (33%) was obtained in a study that attempted six *inter*-molecular disulfide bonds based on a protein–protein interaction model derived in part through molecular dynamics simulation [50]. The average number of attempted disulfides per project was two, and the average project success rate was 75%. Tabulating across projects, a total of 26 successful disulfides were achieved in 41 attempts, reflecting that the groups with a lower success rate tended to perform a greater number of attempted disulfides. The informal survey did not take into account the predicted energy (quality) of the attempted bonds or other experimental factors.

With regard to false negatives, it has been suggested that geometric models employed in disulfide prediction algorithms may be overly restrictive given the dynamic motion of protein chains and their ability to tolerate some modifications without strain [51]. This indeed may be the case, as there are numerous reports of successfully engineered disulfides that violate geometric constraints of computational models. An estimate of the false negative rate can be obtained from Table 1 of Dani et al. [28]. The authors applied their geometric method for disulfide prediction (MODIP) to analyze the sites of 47 engineered disulfide bonds that had previously been reported, covering 23 different protein structures. Their algorithm did not predict nearly half of the disulfide bonds. The same rate was obtained using our own computational method (DbD). The false negatives occurred primarily due to violations of disulfide model bond distance and angle constraints. Comparison of crystal structures of T4 lysozome having engineered disulfides to wild type enzyme reveals that the backbone regions joined by a disulfide bond can move considerably closer together (∼2.5 Å) to accommodate the inserted disulfide [29]. A computational approach developed to include this consideration is BridgeD [52]. This method uses multiple structural models of the target protein to account for backbone flexibility. The model structure set can be assembled when multiple crystal or nuclear magnetic resonance imaging (NMR) structures are available, or from structures derived with homology modeling. The algorithm includes optimization of nearby residue side chains to adjust for the inserted disulfide bond. The authors report that this method improves the ability to identify putative disulfide bonds in cases where optimal geometry is not available.

As noted in the preceding sections, molecular dynamics (MD) simulations have been a valuable component of some disulfide engineering experiments. MD is a computational technique used to simulate the motion of atoms in a molecule as a function of time [53]. A "force field" (potential energy function) defining bonded and non-bonded atomic interactions is used to calculate the forces acting on each atom, and then Newton's laws dictate the resultant atomic motion (trajectory). The process is repeated for millions of iterations, with each step covering femtoseconds. The total time simulated for a protein is typically in the nanoseconds but can extend into microseconds, and recently the millisecond range, for small proteins modeled with advanced hardware [54]. A comprehensive review of MD for protein simulations is available in [55].

A common strategy for the application of MD in disulfide engineering is to first perform MD simulations with elevated temperatures to identify flexible protein regions or segments that unfold early. Then potential disulfide bonds intended to restrict motion of the identified regions are designed using geometric methods, described above. This two-step approach was used effectively by

Pikkemaat et al. to enhance the kinetic stability of haloalkane dehalogenase [33], by Badieyan et al. to increase the stability of cellulase C [30], and recently by Sanchez-Romero et al. to engineer several disulfide bonds that increased the thermodynamic and kinetic stability of phytase [24].

6. Beyond stability: other disulfide engineering applications

While the goal of most disulfide engineering experiments has focused on improving protein stability, the introduction of novel disulfides can be an effective tool in a range of other studies. Engineered disulfide bonds have been used to study protein dynamics and also to create functional modifications in proteins. A number of studies have reported using engineered disulfides to restrict protein domain motion to enable the characterization of dynamic processes. For example, insulin-degrading enzyme (IDE) is a protease that degrades a number of polypeptides, including insulin and amyloid-β. IDE is composed of four domains that are structured to open and close around a catalytic chamber. It was hypothesized that in the absence of a suitable substrate IDE exists in a stable closed state. To confirm this and to better understand the catalytic mechanism of IDE, three disulfides were engineered [56]. The mutant IDEs were locked in a closed conformation through oxidant-induced disulfide bond formation. Two of the three mutant IDEs showed inactivity (the third evidently did not form a disulfide), which could then be restored by breaking the disulfide with a reducing agent. In the reduced state all mutants showed a significant increase in proteolytic activity. Beyond characterizing the dependency of IDE on structural dynamics, this work offers hope that a modified IDE may facilitate the clearance of amyloid plaques.

Similarly, the hypothesized dynamics of presequence protease PreP was confirmed by engineering four mutant versions of PreP, each with cysteine double mutants placed at different points along the two halves of the proposed hinged structure [49]. Again, under oxidizing conditions the protease became catalytically inactive when the structure was effectively locked in a closed conformation that prevented the substrate from reaching the active site. Under reducing conditions proteolytic activity was shown to remain intact. Other examples of protein dynamics studied with the aid of engineered disulfides are: characterization of nascent protein folding near the ribosome [57], the mechanism by which drug transporters function [58], and a study of the dynamic mechanism of ATP-dependent unfoldase [59].

Mechanistic studies of amyloid-β aggregation in Alzheimer's disease (AD) have been greatly advanced through disulfide engineering experiments. The aggregation of amyloid-β peptide (Aβ) is believed to follow several paths that include formation of soluble oligomers, globulamers, protofibrils, and fibrils. It was hypothesized that a conformational rearrangement is required for Aβ to transform from soluble oligomers to fibrils [60]. The former are comprised of Aβ in a β-hairpin conformation with *intra*molecular hydrogen bonds between anti-parallel B strands, while Aβ in fibrils maintain a parallel β-sheet conformation with *inter*molecular hydrogen bonds. To investigate this hypothesis, Sandberg et al. engineered a disulfide bond that locked Aβ peptide into the β-hairpin conformation by linking two opposing, antiparallel β-strands (Aβcc) [60]. They demonstrated that oxidized Aβcc with an intact disulfide bond does not form amyloid fibrils, thus confirming the mechanistic model. A number of other reports have utilized engineered disulfides to assist in the study of amyloid-β and its oligomerization [61–64].

Engineered disulfides have also been used to modify or improve the function of a protein. For example, the bioluminescence properties of firefly luciferase have made it ideal for a wide variety of applications ranging from a reporter of transcription activity to

the imaging of whole cell populations. In addition to increasing the thermostability of luciferase, engineered disulfide bonds have resulted in mutants with decreased pH-sensitivity and a shift in bioluminescent color [65–67]. These modifications extend the potential application of this important protein for use as a reporter of gene expression in high throughput screening.

Another example using disulfide engineering to create a functional modification with important biotechnological implications is the enhanced interaction of monomeric streptavidin and a biotinylated peptide tag. Monomeric M4 streptavidin reversibly binds to biotin, but is unsuitable for immobilization of the ligand, a property previously only found in natural tetrameric streptavidin. A computational method for disulfide engineering was used to identify a site in the binding pocket of monomeric M4 streptavidin and a paired residue in the biotinylated peptide for mutation to cysteines [68]. Upon binding the complex formed a covalent disulfide bond, thereby immobilizing the tagged ligand and providing infinite interaction affinity. The immobilization could easily be disrupted with reducing agents and biotin. This novel biomaterial has potential application in reversible and reusable immobilization devices such as biosensors and protein chips.

7. Future directions

Disulfide engineering has proven to be valuable in a wide range of research and has great potential in many additional applications. It is clear that a better understanding of the determinants of the stabilizing effect of an inserted disulfide bond will improve prediction methods and facilitate expanded use of this technology. Recent work has shown that kinetic effects are important and may enable the design of disulfide bonds that provide kinetic stabilization. Enhanced kinetic stability is an appealing property in many biotechnological applications [26]. Considerable evidence indicates that disulfide bonds inserted into protein regions involved in early unfolding can provide substantial stability. While B-factors may sometimes reveal these regions, the most reliable modeling approach entails the use of molecular dynamics simulations. These methods require specialized computer hardware (e.g. computing clusters) and detailed knowledge of computational methods. This impedes practical use by many investigators. There is a need for fast, efficient algorithms implemented in a manner that allows use by non-specialists and are designed to enable identification of early unfolding regions or intermediate states of proteins. Advancements in this area will likely involve heuristic methods and parallel those in the protein folding field.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2013.11.024.

References

[1] Gidalevitz, T., Stevens, F. and Argon, Y. (2013) Orchestration of secretory protein folding by ER chaperones. Biochim. Biophys. Acta 1833 (11), 2410–2424.

[2] Thornton, J.M. (1981) Disulphide bridges in globular proteins. J. Mol. Biol. 151 (2), 261–287.

[3] Petersen, M.T., Jonson, P.H. and Petersen, S.B. (1999) Amino acid neighbours and detailed conformational analysis of cysteines in proteins. Protein Eng. 12 (7), 535–548.

[4] Creighton, T.E. (1993) Proteins structures and molecular properties, second ed, W.H. Freeman and Company, New York.

[5] Qian, W. and Krimm, S. (1993) Energetics of the disulfide bridge: an ab initio study. Biopolymers 33 (10), 1591–1603.

[6] Pace, C.N., Grimsley, G.R., Thomson, J.A. and Barnett, B.J. (1988) Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J. Biol. Chem. 263 (24), 11820–11825.

[7] Tidor, B. and Karplus, M. (1993) The contribution of cross-links to protein stability: a normal mode analysis of the configurational entropy of the native state. Proteins 15 (1), 71–79.

[8] Zhang, T., Bertelsen, E. and Alber, T. (1994) Entropic effects of disulphide bonds on protein stability. Nat. Struct. Biol. 1 (7), 434–438.

[9] Pecher, P. and Arnold, U. (2009) The effect of additional disulfide bonds on the stability and folding of ribonuclease A. Biophys. Chem. 141 (1), 21–28.

[10] Betz, S.F. (1993) Disulfide bonds and the stability of globular proteins. Protein Sci. 2 (10), 1551–1558.

[11] Zavodszky, M., Chen, C.W., Huang, J.K., Zolkiewski, M., Wen, L. and Krishnamoorthi, R. (2001) Disulfide bond effects on protein stability: designed variants of Cucurbita maxima trypsin inhibitor-V. Protein Sci. 10 (1), 149–160.

[12] Sauer, R.T., Hehir, K., Stearman, R.S., Weiss, M.A., Jeitler-Nilsson, A., Suchanek, E.G. and Pabo, C.O. (1986) An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of lambda repressor. Biochemistry 25 (20), 5992–5998.

[13] Wells, J.A. and Powers, D.B. (1986) In vivo formation and stability of engineered disulfide bonds in subtilisin. J. Biol. Chem. 261 (14), 6564–6570.

[14] Wetzel, R., Perry, L.J., Baase, W.A. and Becktel, W.J. (1988) Disulfide bonds and thermal stability in T4 lysozyme. Proc. Natl. Acad. Sci. USA 85 (2), 401–405.

[15] Perry, L.J. and Wetzel, R. (1984) Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation. Science 226 (4674), 555–557.

[16] Jacobson, R.H., Matsumura, M., Faber, H.R. and Matthews, B.W. (1992) Structure of a stabilizing disulfide bridge mutant that closes the active-site cleft of T4 lysozyme. Protein Sci. 1 (1), 46–57.

[17] Matsumura, M., Becktel, W.J., Levitt, M. and Matthews, B.W. (1989) Stabilization of phage T4 lysozyme by engineered disulfide bonds. Proc. Natl. Acad. Sci. USA 86 (17), 6562–6566.

[18] Baase, W.A., Liu, L., Tronrud, D.E. and Matthews, B.W. (2010) Lessons from the lysozyme of phage T4. Protein Sci. 19 (4), 631–641.

[19] Matsumura, M., Signor, G. and Matthews, B.W. (1989) Substantial increase of protein stability by multiple disulphide bonds. Nature 342 (6247), 291–293.

[20] Clarke, J. and Fersht, A.R. (1993) Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. Biochemistry 32 (16), 4322–4329.

[21] Abkevich, V.I. and Shakhnovich, E.I. (2000) What can disulfide bonds tell us about protein energetics, function and folding: simulations and bioninformatics analysis. J. Mol. Biol. 300 (4), 975–985.

[22] Mansfeld, J., Vriend, G., Van den Burg, B., Eijsink, V.G. and Ulbrich-Hofmann, R. (1999) Probing the unfolding region in a thermolysin-like protease by site-specific immobilization. Biochemistry 38 (26), 8240–8245.

[23] Ramakrishnan, V., Srinivasan, S.P., Salem, S.M., Matthews, S.J., Colon, W., Zaki, M. and Christopher, B. (2012) Geofold: topology-based protein unfolding pathways capture the effects of engineered disulfides on kinetic stability. Proteins 80 (3), 920–934.

[24] Sanchez-Romero, I., Ariza, A., Wilson, K.S., Skjot, M., Vind, J., De Maria, L., Skov, L.K. and Sanchez-Ruiz, J.M. (2013) Mechanism of protein kinetic stabilization by engineered disulfide crosslinks. PLoS One 8 (7), e70013.

[25] Shaw, A. and Bott, R. (1996) Engineering enzymes for stability. Curr. Opin. Struct. Biol. 6 (4), 546–550.

[26] Sanchez-Ruiz, J.M. (2010) Protein kinetic stability. Biophys. Chem. 148 (1–3), 1–15.

[27] Braselmann, E., Chaney, J.L. and Clark, P.L. (2013) Folding the proteome. Trends Biochem. Sci. 38 (7), 337–344.

[28] Dani, V.S., Ramakrishnan, C. and Varadarajan, R. (2003) MODIP revisited: re-evaluation and refinement of an automated procedure for modeling of disulfide bonds in proteins. Protein Eng. 16 (3), 187–193.

[29] Pjura, P.E., Matsumura, M., Wozniak, J.A. and Matthews, B.W. (1990) Structure of a thermostable disulfide-bridge mutant of phage T4 lysozyme shows that an engineered cross-link in a flexible region does not increase the rigidity of the folded protein. Biochemistry 29 (10), 2592–2598.

[30] Badieyan, S., Bevan, D.R. and Zhang, C. (2012) Study and design of stability in GH5 cellulases. Biotechnol. Bioeng. 109 (1), 31–44.

[31] Mansfeld, J., Vriend, G., Dijkstra, B.W., Veltman, O.R., Van den Burg, B., Venema, G., Ulbrich-Hofmann, R. and Eijsink, V.G. (1997) Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. J. Biol. Chem. 272 (17), 11152–11156.

[32] Melnik, B.S., Povarnitsyna, T.V., Glukhov, A.S., Melnik, T.N., Uversky, V.N. and Sarma, R.H. (2012) SS-stabilizing proteins rationally: intrinsic disorder-based design of stabilizing disulphide bridges in GFP. J. Biomol. Struct. Dyn. 29 (4), 815–824.

[33] Pikkemaat, M.G., Linssen, A.B., Berendsen, H.J. and Janssen, D.B. (2002) Molecular dynamics simulations as a tool for improving protein stability. Protein Eng. 15 (3), 185–192.

[34] Yu, H. and Huang, H. (2013) Engineering proteins for thermostability through Rigidifying Flexible Sites (RFS). Biotechnol. Adv., http://dx.doi.org/10.1016/j.biotechadv.2013.10.012.

[35] Le, Q.A., Joo, J.C., Yoo, Y.J. and Kim, Y.H. (2012) Development of thermostable Candida antarctica lipase B through novel in silico design of disulfide bridge. Biotechnol. Bioeng. 109 (4), 867–876.

[36] Parthasarathy, S. and Murthy, M.R. (1997) Analysis of temperature factor distribution in high-resolution protein structures. Protein Sci. 6 (12), 2561–2567.

[37] Sonavane, S., Jaybhaye, A.A. and Jadhav, A.G. (2013) Prediction of temperature factors from protein sequence. Bioinformation 9 (3), 134–140.

[38] Zhang, H., Zhang, T., Chen, K., Shen, S., Ruan, J. and Kurgan, L. (2009) On the relation between residue flexibility and local solvent accessibility in proteins. Proteins 76 (3), 617–636.

[39] Smith, D.K., Radivojac, P., Obradovic, Z., Dunker, A.K. and Zhu, G. (2003) Improved amino acid flexibility parameters. Protein Sci. 12 (5), 1060–1072.

[40] Fontana, A., de Laureto, P.P., Spolaore, B. and Frare, E. (2012) Identifying disordered regions in proteins by limited proteolysis. Methods Mol. Biol. 896, 297–318.

[41] Fontana, A., de Laureto, P.P., Spolaore, B., Frare, E., Picotti, P. and Zambonin, M. (2004) Probing protein structure by limited proteolysis. Acta Biochim. Pol. 51 (2), 299–321.

[42] Hubbard, S.J. (1998) The structural aspects of limited proteolysis of native proteins. Biochim. Biophys. Acta 1382 (2), 191–206.

[43] Pabo, C.O. and Suchanek, E.G. (1986) Computer-aided model-building strategies for protein design. Biochemistry 25 (20), 5987–5991.

[44] Hazes, B. and Dijkstra, B.W. (1988) Model building of disulfide bonds in proteins with known three-dimensional structure. Protein Eng. 2 (2), 119–125.

[45] Sowdhamini, R., Srinivasan, N., Shoichet, B., Santi, D.V., Ramakrishnan, C. and Balaram, P. (1989) Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis. Protein Eng. 3 (2), 95–103.

[46] Dombkowski, A.A. (2003) Disulfide by design: a computational method for the rational design of disulfide bonds in proteins. Bioinformatics 19 (14), 1852–1853.

[47] Dombkowski, A.A. and Crippen, G.M. (2000) Disulfide recognition in an optimized threading potential. Protein Eng. 13 (10), 679–689.

[48] Asgeirsson, B., Adalbjornsson, B.V. and Gylfason, G.A. (2007) Engineered disulfide bonds increase active-site local stability and reduce catalytic activity of a cold-adapted alkaline phosphatase. Biochim. Biophys. Acta 1774 (6), 679–687.

[49] Johnson, K.A., Bhushan, S., Stahl, A., Hallberg, B.M., Frohn, A., Glaser, E. and Eneqvist, T. (2006) The closed structure of presequence protease PreP forms a unique 10000 Angstroms3 chamber for proteolysis. EMBO J. 25 (9), 1977–1986.

[50] Guo, Q., Jureller, J.E., Warren, J.T., Solomaha, E., Florian, J. and Tang, W.J. (2008) Protein–protein docking and analysis reveal that two homologous bacterial adenylyl cyclase toxins interact with calmodulin differently. J. Biol. Chem. 283 (35), 23836–23845.

[51] Meinhold, D., Beach, M., Shao, Y., Osuna, R. and Colon, W. (2006) The location of an engineered inter-subunit disulfide bond in factor for inversion stimulation (FIS) affects the denaturation pathway and cooperativity. Biochemistry 45 (32), 9767–9777.

[52] Pellequer, J.L. and Chen, S.W. (2006) Multi-template approach to modeling engineered disulfide bonds. Proteins 65 (1), 192–202.

[53] Karplus, M. and Petsko, G.A. (1990) Molecular dynamics simulations in biology. Nature 347 (6294), 631–639.

[54] Shaw, D.E., Maragakis, P., Lindorff-Larsen, K., Piana, S., Dror, R.O., Eastwood, M.P., Bank, J.A., Jumper, J.M., Salmon, J.K., Shan, Y., et al. (2010) Atomic-level characterization of the structural dynamics of proteins. Science 330 (6002), 341–346.

[55] Adcock, S.A. and McCammon, J.A. (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. Chem. Rev. 106 (5), 1589–1615.

[56] Shen, Y., Joachimiak, A., Rosner, M.R. and Tang, W.J. (2006) Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. Nature 443 (7113), 870–874.

[57] Hoffmann, A., Becker, A.H., Zachmann-Brand, B., Deuerling, E., Bukau, B. and Kramer, G. (2012) Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. Mol. Cell 48 (1), 63–74.

[58] Seeger, M.A., von Ballmoos, C., Eicher, T., Brandstatter, L., Verrey, F., Diederichs, K. and Pos, K.M. (2008) Engineered disulfide bonds support the functional rotation mechanism of multidrug efflux pump AcrB. Nat. Struct. Mol. Biol. 15 (2), 199–205.

[59] Glynn, S.E., Nager, A.R., Baker, T.A. and Sauer, R.T. (2012) Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. Nat. Struct. Mol. Biol. 19 (6), 616–622.

[60] Sandberg, A., Luheshi, L.M., Sollvander, S., Pereira de Barros, T., Macao, B., Knowles, T.P., Biverstal, H., Lendel, C., Ekholm-Petterson, F., Dubnovitsky, A., et al. (2010) Stabilization of neurotoxic Alzheimer amyloid-beta oligomers by protein engineering. Proc. Natl. Acad. Sci. USA 107 (35), 15595–15600.

[61] Hard, T. (2011) Protein engineering to stabilize soluble amyloid beta-protein aggregates for structural and functional studies. FEBS J. 278 (20), 3884–3892.

[62] O'Nuallain, B., Freir, D.B., Nicoll, A.J., Risse, E., Ferguson, N., Herron, C.E., Collinge, J. and Walsh, D.M. (2010) Amyloid beta-protein dimers rapidly form stable synaptotoxic protofibrils. J. Neurosci. 30 (43), 14411–14419.

[63] Yamaguchi, T., Yagi, H., Goto, Y., Matsuzaki, K. and Hoshino, M. (2010) A disulfide-linked amyloid-beta peptide dimer forms a protofibril-like oligomer through a distinct pathway from amyloid fibril formation. Biochemistry 49 (33), 7100–7107.

[64] Muller-Schiffmann, A., Andreyeva, A., Horn, A.H., Gottmann, K., Korth, C. and Sticht, H. (2011) Molecular engineering of a secreted, highly homogeneous, and neurotoxic abeta dimer. ACS Chem. Neurosci. 2 (5), 242–248.

[65] Imani, M., Hosseinkhani, S., Ahmadian, S. and Nazari, M. (2010) Design and introduction of a disulfide bridge in firefly luciferase: increase of thermostability and decrease of pH sensitivity. Photochem. Photobiol. Sci. 9 (8), 1167–1177.

[66] Nazari, M. and Hosseinkhani, S. (2011) Design of disulfide bridge as an alternative mechanism for color shift in firefly luciferase and development of secreted luciferase. Photochem. Photobiol. Sci. 10 (7), 1203–1215.

[67] Nazari, M., Hosseinkhani, S. and Hassani, L. (2013) Step-wise addition of disulfide bridge in firefly luciferase controls color shift through a flexible loop: a thermodynamic perspective. Photochem. Photobiol. Sci. 12 (2), 298–308.

[68] Wu, S.C., Ng, K.K. and Wong, S.L. (2009) Engineering monomeric streptavidin and its ligands with infinite affinity in binding but reversibility in interaction. Proteins 77 (2), 404–412.

[69] Craig, D.B., and Dombkowski, A.A. (2013) Disulfide by design 2.0: a web-based tool for disulfide engineering in proteins. Bioinformatics. 14 (1), 346. [Epub ahead of print] PubMed PMID: 24289175. http://dx.doi.org/10.1186/1471-2105-14-346.