Chapter 18

# From Biology to Biotechnology: Disulfide Bond Formation in *Escherichia coli*

Bradley J. Landgraf, Guoping Ren, Thorsten Masuch, Dana Boyd and Mehmet Berkmen

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67393

**Abstract**

Disulfide bonds formed between a pair of oxidized cysteines are important to the structural integrity and proper folding of many proteins. Accordingly, Nature has evolved several systems for the genesis and maintenance of such bonds. Beginning with the discovery of protein disulfide isomerase, which provided the first evidence for enzyme-catalyzed disulfide-bond formation, many years of research have resulted in the explication of the complex network of electron transport pathways needed for this process. Herein, we take a historical approach in describing the elucidation of disulfide-bond formation in E. coli. We frame this topic in the context of genome sequencing eras. The first section describes the discovery of eukaryotic protein disulfide isomerase and the subsequent research that followed from the early 1960s to the early 1990s, a time period we have named the pre-genomic sequencing era. The second section details the renaissance in research on disulfide-bond formation in the periplasm of prokaryotes, fueled by bacterial genetic screens and the development of genomic sequencing technology. Accordingly, we have named this section the genomic sequencing era, which ranges from the early 1990s to approximately 2010. The final section outlines the use of bacterial genetic screens to select for new oxidoreductase enzymes and their potential uses in biotechnological and pharmaceutical applications. This era we have dubbed the post-genomic sequencing era, and we envision it to represent the future of research on oxidative folding.

**Keywords:** disulfide bond, thiol, redox, oxidation, reduction, Dsb, SHuffle

## 1. Introduction: covalent bonds in proteins

The amino acids comprising a protein are covalently linked by peptide bonds, which collectively form the "backbone" of the protein. These bonds are the most prevalent covalent links

between amino acids in proteins and form the foundations that support protein secondary and tertiary structures. In addition to peptide bonds, disulfide bonds are the second most common covalent bonds between amino acids. Disulfide bonds are sulfur-sulfur bonds formed within a protein when the thiol (-SH) groups of two cysteine residues are each oxidized, resulting in the net loss of two electrons to an oxidizing agent. Other names for these bonds include S-S bonds, disulfide bridges, disulfide crosslinks, and simply disulfides/disulphides. Disulfide bonds play three main roles in proteins as signal relays (signaling disulfide bonds), as active site residues in enzymatic catalysis (catalytic disulfide bonds), and as structural supports (structural disulfide bonds). Signaling disulfide bonds act as environmental redox readouts, such as those observed in the OxyR transcription factor (reactive oxygen species sensing), the ArcAB system (senses changes in respiratory growth conditions), and the antisigma factor RsrA (activates cytoplasmic oxidative stress response) [1–3]. Catalytic disulfide bonds act as active site residues in oxidoreductases such as DsbA and DsbC, which catalyze the oxidation and isomerization of disulfide bonds, respectively [4–6]. While signaling and catalytic disulfide bonds are important, the primary functions of structural disulfide bonds are to facilitate protein folding and to stabilize protein tertiary structure, which will be the focus of this chapter.

## 2. Disulfide-bond research during the pre-genomic sequencing era

The pre-genomic sequencing era focuses on the research surrounding disulfide-bond formation conducted in the early 1960s through the early 1980s. The work of Anfinsen, Straub, and others on the disulfide-bond formation in Ribonuclease A (RNase A) provided the initial impetus behind studying this process and is described below.

The roles of structural disulfide bonds in protein folding and stability are well exemplified and characterized in the protein RNase A, which contains a total of four disulfide bonds (Cys26-Cys84, Cys40-Cys95, Cys58-Cys110, and Cys65-Cys74) that contribute to the "legendary" stability of RNase A [7]. Indeed, the classic isolation and purification protocol of active RNase A from bovine pancreas calls for extremely harsh conditions in the context of proteins: treatment of pancreatic tissue with 0.25 *N* sulfuric acid at 5°C for 1 day, followed by ammonium sulfate precipitation and boiling in 20% saturated ammonium sulfate at pH 3.0 for 5–10 min [8]. While these purification conditions disrupt noncovalent interactions, the covalent nature of disulfide bonds allows them to survive such treatments, thereby imparting structural stability to the protein. Disulfide bonds also constrain the number of conformations an unfolded protein can adopt, which destabilizes the unfolded state relative to the folded state [9]. In fact, the stability of properly folded RNase A has been estimated to be 8 kcal/mol greater than the unfolded state [10], and the four disulfide bonds collectively contribute an estimated 19 kcal/mol to this stability [11]. While it is clear that structural disulfide bonds play important roles in maintaining the overall integrity of a protein, especially in the case of RNase A, they also serve to aid in the initial folding and refolding of a protein.

Seminal experiments conducted in the early 1960s on RNase A led to the idea that all of the information required to form the correct structures and disulfide bonds of small, globular

http://dx.doi.org/10.5772/67393

proteins is contained within its amino acid sequence [12–17]. Evidence supporting this idea—now known as Anfinsen's dogma—was obtained from *in vitro* experiments in which RNase A was denatured and completely reduced to yield an unstructured polypeptide chain containing eight cysteine residues. Removal of denaturant and reducing agent and subsequent incubation of the enzyme open to atmosphere resulted in relatively rapid and spontaneous formation of disulfide bonds—formed by oxidation via molecular oxygen—followed by much slower recovery of RNase activity. Under optimized conditions, the halftime required for recovery of RNase activity was 20 min. This apparent lag phase between disulfide-bond formation and recovery of activity suggested that spontaneous disulfide-bond formation resulted in incorrect cysteine pairing and that the recovery of activity required additional time for disulfide interchange and proper formation to occur [16]. This also suggested that a system for oxidizing and reshuffling disulfide bonds existed *in vivo*, since the 20-min recovery of activity observed *in vitro* was far too slow to be operative in the cell.

Shortly after demonstrating the spontaneous formation of disulfide bonds in reduced RNase A, Anfinsen and colleagues identified an enzyme, isolated from rat liver microsomes, that stimulated RNase A reactivation, with recovery of activity requiring a halftime of 4.5 min. This discovery confirmed the existence of an *in vivo* system capable of catalyzing the formation of the correct disulfide bonds to yield the active enzyme [18]. In addition to the rat liver, the same enzyme had been identified by Venetianer and Straub in chicken, pig, and pigeon pancreatic tissue [17]. Further studies in which the microsomal enzyme was incubated with reduced egg white lysozyme as a substrate—which contains four disulfide bonds required for its activity—demonstrated similar recovery of activity. Due to the recovery of activity in RNase A from both bovine pancreas and rat liver and in egg white lysozyme, the substrate specificity of the enzyme appeared to be rather low. As a result, this enzyme was speculated to be a general mediator of "sulfhydryl-disulfide exchange" *in vivo*. The catalyzed reaction is driven by the release of free energy attained upon folding substrates from unorganized conformations to their more stable native structures [18–20]. Furthermore, the enzyme catalyzing sulfhydryl-disulfide exchange enzyme was shown to possess a DTT-sensitive disulfide bond that was likely to be involved in catalysis [19]. In 1973, nearly 6 years after its identification, the microsome-associated enzyme that catalyzed sulfhydryl-disulfide exchange was given a name: protein disulfide isomerase (PDI).

In the early 1980s, several studies demonstrated a correlation between PDI activity and the synthesis of disulfide-bond containing proteins in specialized cells and tissues [21, 22]. However, it was not until 1983—nearly 10 years after its naming—that PDI was purified to homogeneity and biochemically characterized [23]. Homogenous PDI was shown to catalyze the reduction of disulfide bonds in insulin *in vitro*. This result demonstrated that PDI could catalyze both disulfide-bond formation and reduction, and led some to question whether the name 'protein disulfide isomerase' was a misnomer. Based on the physiological evidence surrounding PDI—its distribution in tissue with great abundance of disulfide-bonded proteins [22]; its localization in the endoplasmic reticulum, where many disulfide-bonded proteins are synthesized; and its broad substrate specificity—in conjunction with findings showing PDI to be a relatively poor reductase, the name PDI was retained and is still in use [24].

The advent of genome sequencing and PCR in the later 1980s caused a shift from eukaryotic PDI studies to research centered on bacterial disulfide-bond formation, which is detailed in the following section. It should be noted that Anfinsen's idea that the amino acid sequence of a protein encodes all of the information necessary for its proper folding was not fully correct. Even though Anfinsen shared the 1972 Nobel Prize in Chemistry with Stanford Moore and William H. Stein, the following decades of his and others' research showed that disulfide-bond formation and protein folding are, in fact, catalyzed processes *in vivo*. The work surrounding RNase A refolding and the elucidation of PDI serves as an example wherein the true answers to fundamental questions often require far more research to unravel their complexities.

## 3. Genome sequencing enables a great leap forward in bacterial redox biology research

The eponymous Sanger DNA sequencing method was developed by Frederick Sanger and colleagues in 1977 [25]. This method is based on selective incorporation of chain-terminating dideoxynucleotides by DNA polymerases during *in vitro* DNA replication [26]. Sanger sequencing was the most widely utilized DNA sequencing technology until relatively recently. Gene sequencing became reasonably attainable with the 1986 release of a fully automated DNA sequencer made by Applied Biosystems. Around the same time, Kary Mullis of Cetus Corporation developed polymerase chain reaction (PCR) technology, which led to the first commercial PCR enzyme and thermal cycler systems available to scientists in 1987 [27]. Together, Sanger sequencing and the development of PCR technology ushered in the gene sequencing era and revolutionized molecular biology.

With the ability to sequence genes, in conjunction with the already rich field of bacterial genetics and its corresponding techniques, the stage was set for identifying genes involved in redox biology. Along these lines, a genetic selection in *Escherichia coli* designed to identify factors involved in protein translocation led to the serendipitous discovery of mutations in the *dsbA* gene that affected disulfide-bond formation [28]. The DsbA protein was isolated and demonstrated to catalyze disulfide-bond reduction using insulin as a substrate *in vitro* [28, 29]. Later studies revealed DsbA to be a potent and sequential oxidant [30]. Specifically, DsbA forms disulfide bonds between sequential cysteines in proteins as they are translocated to the periplasm [31] (**Figure 1**). Collectively, these studies identified DsbA as the first periplasmic protein involved in disulfide-bond formation and paved the way for elucidating the disulfide-bond forming machinery in *E. coli*.

A second protein involved in disulfide-bond formation was identified through genetic screens of resistance or sensitivity to strong reducing agents. In these screens, Tn*10* insertion mutants sensitive to DTT and benzylpenicillin were mapped to a second gene, which was named *dsbB* [30, 32]. The *dsbB* gene product was later confirmed to specifically oxidize DsbA [33]. Since then, research in several laboratories has elucidated the electron transfer pathway through which approximately 40% of cell envelope proteins in *E. coli* obtain disulfide bonds [34–38]

http://dx.doi.org/10.5772/67393

(see **Figure 1**). Specifically, the DsbA protein transfers disulfide bonds to substrate proteins in the periplasm by accepting electrons from the substrate's cysteine residues. As a result, the cysteine residues of DsbA become reduced and the protein must be oxidized for it to catalyze another round of disulfide bond transfer [28, 29]. This oxidation reaction is carried out by DsbB, an inner membrane protein with two pairs of redox-active cysteines [30, 32]. The electrons received by DsbB in its oxidation of DsbA are transferred to the pool of quinones within the inner membrane [37, 39–43]. Then, the reduced quinones are recycled by cytochrome and terminal oxidases of the electron transport chain [42, 44–46]. Together, DsbA and DsbB act as the oxidation system for disulfide-bond formation in the periplasm (**Figure 1**). These two proteins form one part of the periplasmic disulfide-bond forming pathway; additional proteins, DsbC and DsbD, among others, play downstream roles in the fidelity of native disulfide bonds.

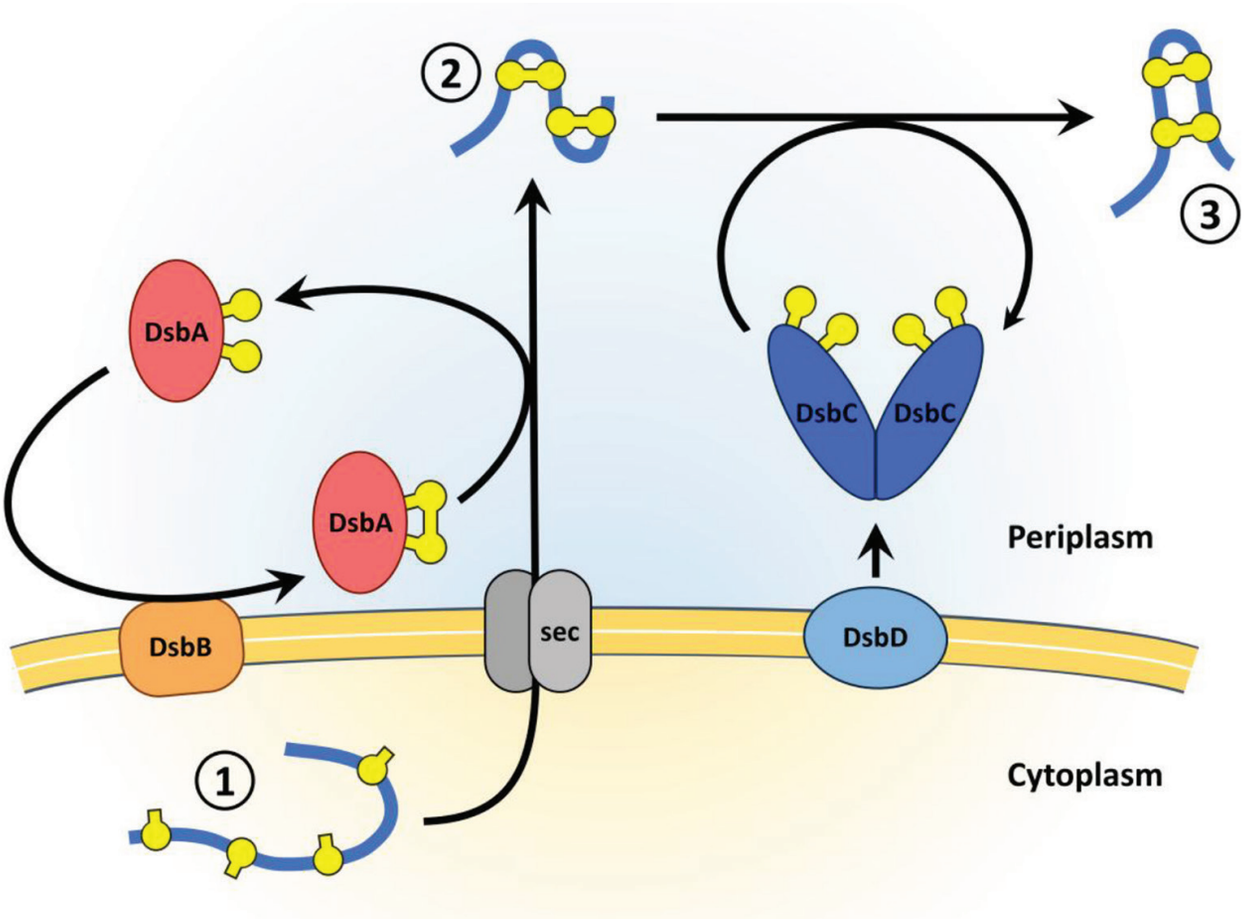


**Figure 1.** The disulfide-bond-forming pathways in the periplasm of E. coli**.** A protein containing four cysteines in their reduced (free thiol) states is translocated into the periplasm by the SecYEG translocon. (1). Oxidized DsbA catalyzes disulfide-bond formation either as the protein is translocated or after, resulting in sequential disulfide bonds in this protein. DsbA is then oxidized to its active state by DsbB. DsbB is oxidized by ubiquinone or menaquinone under aerobic or anaerobic conditions, respectively (not shown). (2). If the disulfide bonds formed by DsbA are misoxidized, reduced DsbC catalyzes their isomerization to yield the properly folded protein. (3). DsbD then reduces DsbC to its active state. DsbD is reduced by an electron cascade originating from NADPH and mediated by thioredoxin reductase and thioredoxin in the cytoplasm (not shown).

The misoxidation of substrates by DsbA necessitates the existence of a system capable of isomerizing incorrect disulfide bonds to their correct linkages in prokaryotes. In *E. coli*, the isomerization of disulfide bonds in proteins is catalyzed by DsbC. The *dsbC* gene was discovered in 1994, shortly after the discovery of the *dsbB* gene, using the same genetic selection approach [47]. The *dsbC* gene product was characterized and was shown to contain two cysteines that reside in the CXXC motif generally found in oxidoreductases. Subsequently, DsbC was demonstrated to catalyze disulfide-bond isomerization of substrates containing nonconsecutive disulfide bonds [48–51] (**Figure 1**). This substrate preference of DsbC was illustrated with two nearly identical *E. coli* proteins, phytase (AppA) and glucose-1-phosphatase (Agp), which differ by the former containing a nonconsecutive disulfide bond, while the latter has only consecutive disulfide bonds. AppA was shown to be dependent on DsbC for proper folding into its active conformation, whereas Agp exhibited no dependence on DsbC until a nonconsecutive disulfide bond placed similarly as that found in phytase was introduced [48]. To date, no exceptions to the substrate preference of DsbC have been found. Taken together, these results suggested that DsbC is a protein disulfide isomerase that catalyzes the rearrangement of misoxidized disulfide bonds, in particular, the rearrangement of nonconsecutive disulfide bonds. Thus, DsbA and DsbC work in parallel in maintaining the correct disulfide bonds in the periplasmic *E. coli* proteome. DsbA catalyzes disulfide-bond formation as the protein is translocated into the periplasm, resulting in the formation of consecutive disulfide bonds. In those proteins requiring nonconsecutive disulfide bonds, DsbC catalyzes the isomerization of misoxidized bonds to yield active enzymes. The exact details of substrate recognition and the *in vivo* mechanism of isomerization catalyzed by DsbC have yet to be elucidated. However, preliminary evidence suggests that certain correctly oxidized proteins may result not only from oxidation and isomerization but also from iterative cycles of reduction and oxidation by DsbA and DsbC [52]. Another protein, DsbG, shares 28% sequence identity with DsbC and exhibits protein disulfide isomerase activity, albeit on a more narrow scope of yet-to-be-identified substrates [5, 53, 54].

Like DsbA, DsbC has a dedicated redox protein partner, named DsbD, which is responsible for maintaining it in its reduced state (**Figure 1**). The *dsbD* gene was discovered using the same genetic screens that led to the discoveries of both DsbB and DsbC [55]. The *dsbD* gene product consists of three domains: an N-terminal periplasmic domain, a transmembrane domain, and a C-terminal periplasmic thioredoxin-like domain that shares approximately 45% sequence homology with eukaryotic PDIs [56]. Each of the domains of DsbD contains a pair of conserved cysteine residues that are redox active and essential for its function [56]. To maintain DsbC in its reduced state, DsbD channels reducing equivalents that are mediated through a cascade of disulfide-bond reductions starting with the reduction of thioredoxin reductase by NADPH [57, 58]. Thioredoxin reductase reduces thioredoxin, which then reduces the cysteine pair in the transmembrane domain of DsbD [58, 59]. This reduced cysteine pair initiates the sequential reduction of disulfide bonds in the C-terminal and N-terminal DsbD domains, respectively [59]. The reduced N-terminal domain cysteines then reduce DsbC (**Figure 1**). Reduction of DsbC occurs only when it is dimeric [60, 61]. This substrate preference likely stems from the tertiary structure of the N-terminal domain of DsbD, which adopts a immunoglobulin-like fold and places the active site in the antigen-binding-like region [62]. The tertiary

structure promotes the binding of the DsbC dimer and occludes the binding of the monomeric DsbA and DsbB proteins, thereby separating the oxidative and reductive pathways [63].

The formation of disulfide bonds is essential to the structural integrity and folding of proteins that are vital in many biological processes. *E. coli* and other prokaryotes have evolved a complex network of electron transport chains and quality control systems to facilitate and ensure proper disulfide-bond formation in the form of the Dsb proteins described above. The discovery of these Dsb proteins, and the subsequent revival of interest in disulfide-bond formation in eukaryotes, would not have been realized without the powerful combination of well-designed, selective genetic screens to produce mutants and the ability to sequence the resulting mutated genes. With the advent of next-generation sequencing, we should expect further elucidation of the biological and chemical processes that we do not yet understand or have yet to be discovered.

## 4. Disulfide-bond research in the post-genomic sequencing era

Since 2008, the cost of genome sequencing has declined faster than predicted by Moore's Law [64]. Currently, the cost of sequencing a genome is ~$1500, and the lofty $1000/genome goal is within reach. Due to the radical drop in DNA sequencing costs, a multitude of laboratories and private and government institutions have completed the sequencing of approximately 30,000 bacterial genomes [65]. This wealth of data is currently being used for a variety of biotechnological and clinical purposes including diagnostics, public health benefits, and biosurveillance/epidemiological studies [66, 67]. Accordingly, we have termed this time period as the "post-genomic sequencing era" to represent research that uses sequenced genomes, metagenomes, and environmental samples to search for novel enzymes and pathways and to predict the redox biology of bacteria.

### 4.1. Hunting for new disulfide-bond forming enzymes in the genomic landscape

One of the first examples of the use of sequenced genomes to predict and identify novel disulfide-bond forming pathways was conducted by Todd Yeates and colleagues [68–70]. They hypothesized that organisms rich in disulfide-bonded proteins would have a propensity to encode for proteins with an even number of cysteine residues, since an odd number might cause formation of aberrant disulfide bonds. This conjecture was based on the observation that the predicted open reading frames (ORFs) of the hyperthermophilic *Pyrobaculum aerophilum* and *Aeropyrum pernix* species are strongly biased toward an even number of cysteines [70]. Since then, they have expanded their analysis to show that hyperthermophilic members of the Crenarchaeota branch all contain a multitude of disulfide-bonded proteins [68]. Mass spectrometric analysis of the proteome of *Sulfolobus solfataricus* revealed the majority of cysteines to be disulfide bonded [71], and several disulfide-bonded proteins were identified using 2D gel analysis of lysates of *P. aerophilum* [72]. The presence of a high number of disulfide bond-containing proteins in hyperthermophilic Crenarchaeota suggested these bacteria possess an undiscovered method of disulfide-bond maintenance. Indeed, experimental

evidence of such a system was obtained from the *in vitro* characterization of protein disulfide oxidoreductases (PDO) from *Pyrococcus furiosus* [73], *Aquifex aeolicus* [74], *A. pernix* [75], and *S. solfataricus* [76]. PDOs have been shown to be functional homologs of PDI and DsbC, that exhibit reduction, oxidation, and isomerization of disulfide bonds. Although there is growing evidence that the cytoplasm of Crenarchaeota is more amenable to disulfide-bond formation, the exact mechanism and the enzymes involved remain to be elucidated *in vivo*.

The method of predicting redox biology of organisms by simply analyzing the cysteine content of the predicted ORFs from sequenced genomes was expanded to all prokaryotic organisms with known genome sequences. By separating the predicted proteome into two subgroups—proteins predicted to be exported and those that remained in the cytoplasm—this bioinformatic method was further developed to predict whether the periplasmic space was oxidizing or reducing [77]. This method led to the observation that some bacteria predicted to have an oxidizing periplasm encode a homolog of DsbA but lack a homolog of its partner DsbB. A closer look at these strains revealed that the DsbA homolog in *Mycobacterium* was a fusion protein to vitamin K epoxide reductase (VKOR) [77]. Characterization of bacterial VKOR homologs confirmed that VKOR can indeed functionally replace DsbB in certain organisms [78, 79]. To our knowledge, this was the first use of genomic data to mine for new oxidoreductases, leading to the discovery of VKOR as a functional homolog of DsbB.

### 4.2. Selecting for new oxidoreductases using living bacteria

The advent of modern biomolecular tools, in conjunction with classical bacterial genetic screens, has led to the discovery of novel enzymes, yielded many new insights into biochemical pathways, and elucidated molecular mechanisms. The discovery that disulfide bonds are not formed spontaneously but are, in fact, formed catalytically by the enzyme DsbA was a serendipitous discovery using a blue/white screen for secretion defects [28]. The malF-lacZ fusion has been used to not only discover DsbA [28] but also mutants of DsbA with various kinetic properties [31]. Since then, many other genetic screens have been developed to specifically detect the activity of an oxidoreductase in *E. coli*. These screens, described briefly below, allow for the selection of gene products whose activities permit the growth of strains in the absence of a *dsb* component. Characterization of mutant strains revealed insight into the molecular machinery of disulfide-bond formation and highlighted the plasticity of the dsb machinery. A few key mutations could convert a dedicated reductase into an oxidase or create novel pathways to maintain cell viability.

FlgI is a protein component of the flagellar machinery and requires a disulfide bond for its correct folding and activity [80]. Strains that have a functional disulfide-bond forming pathway are motile, while those with defects in disulfide-bond formation are not. By simply spotting bacteria incapable of forming disulfide bonds on dilute agar, researchers are able to screen and select for bacteria that have gained the ability to form disulfide bonds, since they become motile and swim away from the center. This phenotype has been used to characterize and select for new disulfide bond oxidases, such as selecting for mutant thioredoxins possessing a new mechanism of disulfide-bond formation in the periplasm [81]. In another approach, researchers screened a multicopy plasmid library of *E. coli* and selected a rhodanese protein

(PspE) with a single cysteine that can promote disulfide-bond formation in a strain completely lacking the *dsb* pathway [82].

Heavy metals such as copper or cadmium can oxidize thiol groups in periplasmic proteins, resulting in misfolding of proteins containing cysteines and, in some cases, leading to death [83]. DsbC can reduce and refold proteins that were misoxidized by such metals and is therefore necessary to protect cells from copper and cadmium-induced oxidative damage. This phenotype was used to select for strains containing mutant DsbG proteins that have gained the ability to isomerize misoxidized proteins [84]. In another heavy metal screen, cells lacking the *dsbA* gene were screened for cadmium resistance to select for mutant DsbB that can bypass the need for DsbA [85]. The mutant DsbB proteins were able to oxidize DsbC and thus promote disulfide-bond formation.

A blue/white screen was developed using a mutant alkaline phosphatase (*phoA**) that required DsbC for its correct folding and activity. Unlike DsbC, DsbG cannot isomerize misoxidized PhoA*. Mutants of *dsbG* were selected for their gained ability to isomerize PhoA*, resulting in the first *in vivo* screen that directly detected disulfide-bond isomerization of a single protein. This screen permitted the identification of key residues that converted a sulfenic acid reductase (DsbG) into a disulfide-bond isomerase whose activity increased the cells' resistance to copper. Searching the genomes of sequenced prokaryotes, homologs of DsbG were discovered to naturally have the key residues identified through the *phoA** screen. Interestingly, these naturally existing homologs were also capable of protecting cells against copper toxicity. Thus, through the identification of these key residues, activities of homologs can be predicted and tested [86].

The study of disulfide-bond formation has grown and matured significantly since the discovery of DsbA in 1991 [28]. Subsequently, the Dsb pathway in the model organism *E. coli* has been studied in great detail both *in vivo* and *in vitro,* and many novel and interesting mutants and suppressors have been identified using various *in vivo* screens. These new enzymes should have applications in both biotechnology and the pharmaceutical industry as detailed in the next section.

### 4.3. Biotechnological applications of disulfide-bonded proteins

Both the pharmaceutical and the biotechnological industries are extremely interested in disulfide-bonded proteins. Most eukaryotic cell surface and secreted proteins are rich in disulfide bonds due to the increased stability they confer, making these proteins attractive candidates as therapeutics (also known as biologics). For example, the first recombinant biologic was the hormone insulin, which was introduced by Eli Lilly in 1982, and the most profitable biologic is the antibody Humira (adalimumab), both of which are disulfide-bonded proteins [87]. Between 1982 and 2013, approximately 100 recombinant protein therapeutics have been approved by the FDA, of which more than one-third are disulfide-bonded proteins (in particular monoclonal antibodies) [88].

Currently, antibodies represent the fastest growing category of biologics. Their specificity to therapeutic targets, ability to induce or inhibit immune response, and favorable pharmacokinetic profiles within the human body make them attractive therapeutics. The first therapeutic

monoclonal antibody product, Orthoclone OKT3 (muromonab-CD3), was FDA approved in 1986. Since then, research and development of biologics has led to many successful therapeutics, with projected sales expected to reach nearly $125 billion by 2020 [89] (see **Table 1** for top 11 best-selling biologics in 2013 [90]). The production of antibodies for therapeutic applications is a well-established pipeline dominated by the use of Chinese hamster ovary (CHO) cells or hybridomas. However, identifying, characterizing, and engineering therapeutic antibodies are still expensive, time-consuming, and effortful endeavors, leaving room for these aspects of biologic development to be streamlined.

| Name | Lead company | Molecule type | Approved indication(s) | 2013 worldwide sales ($ millions) |
|---|---|---|---|---|
| Humira *(adalimumab)* | AbbVie | mAb | RA, juvenile RA, Crohn's disease, PA, psoriasis, ankylosing spondylitis, UC | 10,659 |
| Enbrel *(etanercept)* | Amgen | Protein | RA, psoriasis, ankylosing spondylitis, PA juvenile RA | 8739 |
| Lantus (*insulin glargine*) | Sanofi | Peptide | Diabetes mellitus type I, diabetes mellitus type II | 7593 |
| Rituxan (*rituximab*) | Roche | mAb | RA, chronic, lymphocytic leukemia/small cell lymphocytic lymphoma, non-Hodgkin's lymphoma, antineutrophil cytoplasmic antibodies-associated vasculitis, indolent non-Hodgkin's lymphoma, diffuse large B-cell lymphoma | 7500 |
| Remicade (*infliximab*) | Johnson & Johnson | mAb | RA, Crohn's disease, psoriasis, UC, ankylosing spondylitis, PA | 6962 |
| Avastin (*bevacizumab*) | Roche | mAb | Colorectal cancer, non-small cell lung cancer, renal cell cancer, brain cancer (malignant glioma; AA and GBM) | 6747 |
| Herceptin (*trastuzumab*) | Roche | mAb | Breast cancer, gastric cancer | 6558 |
| Gleevec (*imatinib*) | Novartis | Small molecule | Chronic myelogenous leukemia, gastrointestinal stromal tumor, acute lymphocytic leukemia, hypereosinophilic syndrome, mastocytosis, dermatofibrosarcoma protuberans, myelodysplastic syndrome, myeloproliferative disorders | 4693 |
| Neulasta (*pegfilgrastim*) | Amgen | Protein | Neutropenia/leukopenia | 4392 |
| Copaxone (*glatiramer acetate*) | Teva Pharmaceutical | Peptide | Multiple sclerosis | 4356 |
| Revlimid (*lenalidomide*) | Celgene | Small molecule | Multiple myeloma, myelodysplastic syndrome, mantle cell lymphoma | 4281 |

Abbreviations: mAb, monoclonal antibody; RA, rheumatoid arthritis; PA, psoriatic arthritis; UC, ulcerative colitis; AA, anaplastic astrocytoma; GBM, glioblastoma multiforme. Adapted from Ref. [90].

**Table 1.** The top 10 best selling biologics in 2013. Of these 11 biologics, five are antibody-based therapeutics, indicated by the mAb under molecule type.

The use of *E. coli* as the most popular host for recombinant engineering of proteins stems from the bevy of powerful genetic tools available, its cost effectiveness, and the short time frames required for both its growth and genetic experiments. The periplasm, where conditions favor oxidized proteins, was the clear compartment to express various antibody fragments, including full-length antibodies, versus the reducing conditions of the cytoplasm [91, 92]. The periplasmic space remains an attractive alternative for the production of disulfide-bonded proteins whose presence/activities may be toxic when expressed in the cytoplasm [93]. However, translocation of the target protein across the inner membrane to the periplasm can be problematic and may require extensive optimization of both the expression conditions and the targeting signal sequence. Furthermore, the lack of ATP in the periplasm makes it an energy-poor environment for proteins that require ATP-dependent chaperones for their folding. The cytoplasm is therefore a more suitable compartment for high-yielding protein production. It also obviates the problem of crossing the membrane and is rich in ATP, chaperones, and folding factors.

With the introduction of the Δ*trxB*, *gor* engineered strains of *E. coli* [94, 95], it is now possible to not only express various antibody fragments but also full-length antibodies in the cytoplasm [96]. Yet, the lack of N-linked glycosylation in *E. coli* has hampered its use in the production of therapeutic immunoglobulins (IgGs), although a few examples of *E. coli*-produced therapeutic antibody fragments can be found, such as the Fab' fragment named Lucentis (ranibizumab) against age-related macular degeneration [97]. The discovery of mutations in the Fc portion of IgG that circumvent the dependency on glycosylation for effective interaction with its cognate Fcγ receptor [98] opened the path to potential therapeutic applications of *E. coli*-expressed IgG [96]. Though *E. coli* is currently not as established as CHO or hybridoma cell lines for the production of therapeutic IgG, it is slowly becoming a more common host for the production of antibodies. Other *E. coli*-based technologies, such as phage display, have had extensive use in the discovery and engineering of antibodies, both for the biotech and the pharmaceutical industries. The use of phage display technology to identify novel antibodies of therapeutic targets, such as the HIV virus coat protein, was first described in 1991 [99]. Since then, phage display has been used to develop novel antibody-based applications. For example, the antibody Humira went through extensive engineering using this technique to create an effective biologic [100].

### 4.4. Engineering disulfide bonds

One key feature of disulfide bonds is their ability to increase the thermostability of proteins by decreasing the number of conformations a protein can attain and thus lowering the conformational entropy of a protein. Secreted proteins leave the protective environment of the cell cytoplasm, and they are rich in disulfide bonds which help to increase their extracellular half-lives. These enzymes are of significant utility in the biotech industry where high-temperature processes are often used. In some cases, disulfide bonds have been introduced into such enzymes to increase their thermostability [101]. Early investigations into the effects of engineered disulfide bonds were performed on phage lambda repressor [102], T4 lysozyme [103], and subtilisin [104], and later were expanded to antibodies [105] and other proteins used in the biotechnology industry. For example, the disulfide bond engineered into the extracellular ribonuclease (barnase) from *Bacillus amyloliquefaciens* unfolds 20 times slower than wild type and 170 times

slower than the reduced protein [106]. It is also possible to engineer an interchain disulfide bond within two subunits to bring together the activities of two distinct enzymes [107].

In addition to engineering disulfide bonds into proteins, the reactivity of disulfide-bond forming proteins can also be altered to provide new functionalities. For example, chimeras were created by fusing the disulfide-bond oxidase DsbA to the dimerization domain and $\alpha$-helical linker derived from the bacterial proline *cis*/*trans* isomerase FkpA. These chimeras were capable of catalyzing the *in vivo* isomerization of misoxidized disulfide bonds with similar efficiency as that of DsbC [108]. The DsbA-FkpA chimeras also conferred modest resistance to $CuCl_2$, which is dependent on disulfide-bond isomerization. This resistance allowed for the selection of DsbA-FkpA mutants which were found to contain a single amino acid variation in the active site of DsbA from CPHC to CPYC. Substitution of histidine with tyrosine made the active site more DsbC-like (CGYC), which could partially explain the gain of DsbC-like isomerization activity. Interestingly, DsbC is not a substrate for the DsbA-DsbB oxidation system and does not exhibit disulfide-bond oxidase activity. However, the DsbA-FkpA chimeras exhibited both oxidase and isomerization functionalities, and *dsbA* deletion strains were partially complemented by the presence of the DsbB-dependent DsbA-FkpA chimeras [108].

### 4.5. Dsb enzymes as novel antimicrobial targets

Many pathogenic bacteria, including *Vibrio cholerae, Pseudomonas aeruginosa, Salmonella enterica, Helicobacter pylori, Bordetella pertussis,* and *E. coli,* among others, make use of periplasmic disulfide-bonded proteins that act as virulence factors or function in processes related to their pathogenicity [109–114]. These virulence factors and pathogenic functions rely on the Dsb proteins, in particular DsbA, for proper folding. As a consequence, disruptions in the redox and isomerization activities of the dsb system partially or fully attenuate the pathogenicity of these bacterial species [115]. Specifically, maturation of toxins of *V. cholerae, B. pertussis,* and *E. coli* requires the formation of DsbA-dependent disulfide bonds [109, 113, 116, 117]. Strains of these bacteria lacking *dsbA* synthesize misfolded, misassembled, and/or unstable toxin proteins that are severely impaired or nonfunctional. Along these lines, Δ*dsbA* strains of *S. enterica* and *E. coli* lack the flagellin (FliC) protein, which is a primary constituent of the filaments of their flagella. FliC does not contain any disulfide bonds. However, due to the hierarchical assembly of the flagellum machinery, which requires several proteins with disulfide bonds to precede FliC in its biogenesis, it is thought that FliC simply is not translated or that it cannot be assembled into the organelle due to the missing disulfide bonds and/or disulfide-bonded proteins [50, 118]. As a result, these Δ*dsbA* strains of *S. enterica* and *E. coli* bacteria are nonmotile and their pathogenicity is severely attenuated. Additionally, the loss of disulfide bonds in Δ*dsbA* strains of *V. cholerae* and *E. coli* affects their ability to adhere to eukaryotic cells and/or form biofilms due to defects in their pili, thereby limiting their infectivity [109, 119, 120]. All together, these studies showed that the Dsb enzymes, especially DsbA, play crucial roles in the pathogenicity of several species of bacteria, making these enzymes logical targets for novel antibiotic development.

Indeed, some research has focused on the development of small molecule inhibitors of Dsb enzymes and their homologs (reviewed in Refs. [121, 122]). In humans, blood coagulation

involves the activity of the enzyme vitamin K epoxide reductase (VKOR), which is inhibited by the anticoagulant drug warfarin (Coumadin). Interestingly, *Mycobacterium tuberculosis (Mtb)* and other bacteria do not encode for a DsbB protein, but instead encode a homolog of VKOR. Although DsbB and VKOR exhibit little sequence similarity, they appear to be functionally similar, since VKOR can replace DsbB in both *E. coli* and cyanobacterial *ΔdsbB* strains [77, 123]. Warfarin was shown to both inhibit *Mtb*VKOR activity and bacterial growth. Furthermore, mutations in the VKOR protein from warfarin-resistant *Mtb* mutants were mapped to nearly identical locations in mutant VKORs from patients who require higher effective doses of warfarin, indicating the drug likely inhibits bacterial and human VKORs in similar manners [34]. These findings, in conjunction with the severe growth defects observed in *Mtb*VKOR homolog deletion strains, suggested that stronger inhibitors of *Mtb*VKOR could be used as effective antituberculosis agents [34].

Promising small molecule inhibitors of bacterial Dsb proteins have been identified using fragment-based lead discovery (FBLD) [124]. FBLD identifies small molecule fragments that weakly bind to a target of interest. Through many rounds of iterative combinations of such fragments and high-throughput screening, candidate molecules with higher binding affinities for the target are created, leading to possible drug candidates. Using a detergent-solubilized *Ec*DsbB immobilized onto sepharose resin and $^{1}$H NMR, 1071 fragments were tested for both binding to and inhibition of *Ec*DsbB, yielding eight fragments exhibiting $IC_{50}$ values of 7–170 μM. The eight fragments were divided into two groups based on their molecular scaffolds and hypothesized mechanisms of inhibition: blocking of quinone binding and blocking of both quinone and *Ec*DsbA binding to DsbB [124]. A further study improved the $IC_{50}$ value of a candidate molecule to 1.1 μM through additional rounds of FBLD. This molecule inhibited both *Ec*DsbA and DsbB through covalent modification of active site cysteine residue in each protein with a propionyl group, thereby abrogating their ability to form disulfide bonds. The molecule also exhibited a degree of selectivity for DsbA and DsbB proteins, since it was shown to have no effect on human thioredoxin activity [125].

Through the use of high-throughput blue/white screening, six additional small molecule inhibitors of *Ec*DsbB were identified from a pool of approximately 52,000 compounds. These six molecules all contained a pyridazinone ring and exhibited a degree of selectivity for *Ec*DsbB, since they were unable to inhibit the *Mtb*VKOR homolog described above. Interestingly, the molecules inhibited DsbB enzymes from other Gram-negative pathogens, including *V. cholerae, Haemophilus influenzae, Salmonella typhimurium, Klebsiella pneumoniae, Francisella tularensis, Acinetobacter baumannii*, and *P. aeruginosa*, to varying degrees [126].

In addition to small molecules, larger peptides capable of inhibiting the formation of the DsbA-DsbB complex have been developed. Using the crystal structure of the DsbA-DsbB complex [127], a peptide of seven amino acids corresponding to a loop of DsbB involved in docking with DsbA was identified and found to bind to *Ec*DsbA with low micromolar affinity ($K_d$ = 13.1 ± 0.4 μM). Further engineering of this peptide resulted in a new peptide with greater affinity ($K_d$ = 5.7 ± 0.4 μM) that also exhibited fairly potent inhibition of *Ec*DsbA oxidase activity ($IC_{50}$ = 8.8 ± 1.1 μM) [128]. The studies described herein clearly show that the DsbA-DsbB protein system is an attractive and tractable target for novel antibiotic

development. While the inhibitors described above exhibit relatively weak binding affinities, the resulting phenotypes observed support their disruption of disulfide-bond formation in the cell. These "first-generation" molecules can serve as a foundation from which more potent compounds can be identified and developed.

## 5. Future directions

There are 4306 predicted *E. coli* K12 protein sequences present in the UniProt proteome database (http://www.uniprot.org/proteomes/) [129]. An initial analysis of their compartmentalization within the cell using the prediction software TOPCONS2 (http://topcons.cbr.su.se/pred/) [130] allowed us to putatively assign each of these proteins to one of three subcellular compartments: cytoplasmic, transmembrane in the inner membrane (referred to as transmembrane hereafter), or secreted. To hone in on proteins exhibiting possible oxidoreductase activity, the CXXC motif was used as a signature to identify 406 proteins, which showed that approximately 10% of all predicted *E. coli* proteins contain this motif, thereby demonstrating its relative ubiquity. Of these 406 proteins, ~75% are cytoplasmic, ~18% are transmembrane, and ~7% are secreted (see **Table 2**). The pool of non-CXXC-containing proteins comprises the remaining 3900 proteins, of which ~63% are cytoplasmic, ~23% are transmembrane, and ~14% are secreted (omitted from **Table 2**). The transmembrane and secreted compartments have a lower fraction of CXXC-containing proteins in keeping with the exclusion of cysteine residues from these compartments in aerobes [77]. A comparison of the non-CXXC sequence pool with the CXXC sequence pool shows a slight enrichment of CXXC proteins in the cytoplasm (~75%) versus non-CXXC cytoplasmic proteins (~63%). The distribution of CXXC and non-CXXC proteins in the transmembrane is similar (18 and 23%, respectively); however, about twice as many non-CXXC proteins are secreted (14%) compared to CXXC proteins. Approximately 22% (90 of 406) of CXXC proteins are annotated in the UniProt data as binding metal ions or as iron-sulfur cluster-containing proteins. While 46% of all CXXC proteins have been functionally characterized, the remaining majority (54%) should be characterized to develop a better understanding of the reactions they catalyze, how those identified to be oxidoreductases may contribute to the redox biology of bacteria, and to identify novel targets for therapeutics.

| Compartment | Number of proteins | | Contain CXXC | | Known function | | Unknown function | | Metal binding | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cytoplasm | 2755 | 64% | 305 | 75% | 147 | 78% | 158 | 72% | 79 | 88% |
| Transmembrane | 970 | 23% | 72 | 18% | 33 | 18% | 39 | 18% | 7 | 8% |
| Secreted | 581 | 13% | 29 | 7% | 8 | 4% | 21 | 10% | 4 | 4% |
| Total | 4306 | 100% | 406 | 100% | 188 | 100% | 218 | 100% | 90 | 100% |

Secreted refers to proteins in the periplasm and secreted outside of the cell. Compartment location was predicted using topological and signal sequence input data on the TOPCONS server. Gene ontology (GO) codes EXP and IDA were used to identify proteins with experimentally verified function from the UniProt database; those lacking these codes were defined as having unknown function. GO codes were also used to identify CXXC proteins annotated to bind metals [129].

**Table 2.** The *E. coli* proteome separated by compartment, the presence of CXXC motifs, and known function.

## 6. Conclusions

While more than 20 years of research have elucidated many of the Dsb proteins and their functions, more questions surrounding these proteins remain to be answered: What are the precise mechanisms by which PDI and DsbC catalyze disulfide-bond isomerization *in vivo*? How are electrons transported across the inner membrane by DsbD? What are the redox states and midpoint potentials of the cytoplasm of Crenarchaeota? Additionally, most of the characterization of Dsb proteins has been done in *E. coli*, which is not an appropriate model for all bacteria, e.g., *M. tuberculosis, Staphylococcus aureus,* and *Listeria monocytogenes*, so further characterization of the Dsb protein networks in other organisms is needed. Along these lines, Dsb proteins from pathogenic bacteria represent possible targets for antibiotic/vaccine development. Since several Dsb proteins have been structurally characterized, it is now possible to develop antibiotics by structure-guided design. While broad-spectrum antibiotic molecules are unlikely to be developed, again due to the diversity of Dsb proteins/networks within bacterial species, those targeting specific pathogenic species are not out of reach.

As more disulfide-bonded proteins are characterized, our knowledge of the stability and structures these bonds confer, their likelihood of scrambling in mulitply disulfide-bonded proteins, and their relative redox potentials will grow. This will allow researchers to better predict native disulfide bonds from sequence data and better engineer disulfide bonds in proteins for desirable physicochemical properties, which will benefit both the biotechnological and pharmaceutical industries, especially in the development and production of antibodies. Ideally, both industries should aim to produce antibodies as quickly, cheaply, and effectively as possible. The engineering of bacterial strains to overproduce correctly folded antibodies and/or engineering antibodies themselves for desired properties represents a technically challenging but incredibly useful advancement in the field of oxidative protein folding. Future research in these areas should lead to great innovations in both the biotechnological and pharmaceutical industries that will improve the health and increase the knowledge of humankind.

## Author details

Bradley J. Landgraf[1], Guoping Ren[1], Thorsten Masuch[1], Dana Boyd[2] and Mehmet Berkmen[1]*

*Address all correspondence to: berkmen@neb.com

1 New England Biolabs / Ruhr-Universität Bochum, Germany

2 Department of Microbiology and Immunology, Harvard Medical School, Boston, MA, USA

## References

[1] Lee C, Lee SM, Mukhopadhyay P, Kim SJ, Lee SC, et al. 2004. Redox regulation of OxyR requires specific disulfide bond formation involving a rapid kinetic reaction path. Nat. Struct. Mol. Biol. 11: 1179-85.

[2] Alvarez FA, Georgellis D. 2010. *In vitro* and *in vivo* analysis of the ArcB/A redox signaling pathway. Methods Enzymol. 471: 205-26.

[3] Kang JG, Paget MS, Seok YJ, Hahn MY, Bae JB, et al. 1999. RsrA, an anti-sigma factor regulated by redox change. EMBO J. 18: 4292-8.

[4] Jander G, Martin NL, Beckwith J. 1994. Two cysteines in each periplasmic domain of the membrane protein DsbB are required for its function in protein disulfide bond formation. EMBO J. 13: 5121-7.

[5] Andersen CL, Matthey-Dupraz A, Missiakas D, Raina S. 1997. A new *Escherichia coli* gene, *dsbG*, encodes a periplasmic protein involved in disulphide bond formation, required for recycling DsbA/DsbB and DsbC redox proteins. Mol. Microbiol. 26: 121-32.

[6] Gon S, Faulkner MJ, Beckwith J. 2006. *In vivo* requirement for glutaredoxins and thioredoxins in the reduction of the ribonucleotide reductases of *Escherichia coli*. Antioxid. Redox Signal. 8: 735-42.

[7] Klink TA, Woycechowsky KJ, Taylor KM, Raines RT. 2000. Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. Eur. J. Biochem. 267: 566-72.

[8] Kunitz M, McDonald MR. 1953. Biochem. Prep. 3: 9.

[9] Flory PJ. 1956. Theory of elastic mechanisms in fibrous proteins. J. Am. Chem. Soc. 78: 5222-35.

[10] Pfeil W. 1981. The problem of the stability of globular proteins Mol. Cell. Biochem. 40: 3-28.

[11] Konishi Y, Ooi T, Scheraga HA. 1982. Regeneration of ribonuclease A from the reduced protein. 2. Energetic analysis. Biochemistry 21: 4741-8.

[12] White FH, Jr. 1960. Regeneration of enzymatic activity by air oxidation of reduced ribonuclease with observations on thiolation during reduction with thioglycolate. J. Biol. Chem. 235: 383-9.

[13] Haber E, Anfinsen CB. 1961. Regeneration of enzyme activity by air oxidation of reduced subtilisin-modified ribonuclease. J. Biol. Chem. 236: 422-4.

[14] Anfinsen CB, Haber E. 1961. Studies on the reduction and re-formation of protein disulfide bonds. J. Biol. Chem. 236: 1361-3.

[15] Anfinsen CB, Haber E, Sela M, White FH, Jr. 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. U. S. A. 47: 1309-14.

[16] Goldberger RF, Epstein CJ, Anfinsen CB. 1963. Acceleration of reactivation of reduced bovine pancreatic ribonuclease by a microsomal system from rat liver. J. Biol. Chem. 238: 628-35.

[17] Venetianer P, Straub FB. 1963. The enzymic reactivation of reduced ribonuclease. Biochim. Biophys. Acta (BBA) Spec. Sec. Enzymol. Subj. 67: 166-8.

[18] Goldberger RF, Epstein CJ, Anfinsen CB. 1964. Purification and properties of a microsomal enzyme system catalyzing the reactivation of reduced ribonuclease and lysozyme. J. Biol. Chem. 239: 1406-10.

[19] De Lorenzo F, Goldberger RF, Steers E, Jr., Givol D, Anfinsen B. 1966. Purification and properties of an enzyme from beef liver which catalyzes sulfhydryl-disulfide interchange in proteins. J. Biol. Chem. 241: 1562-7.

[20] Fuchs S, De Lorenzo F, Anfinsen CB. 1967. Studies on the mechanism of the enzymic catalysis of disulfide interchange in proteins. J. Biol. Chem. 242: 398-402.

[21] Brockway BE, Forster SJ, Freedman RB. 1980. Protein disulphide-isomerase activity in chick-embryo tissues. Correlation with the biosynthesis of procollagen. Biochem. J. 191: 873-6.

[22] Roth RA, Koshland ME. 1981. Role of disulfide interchange enzyme in immunoglobulin synthesis. Biochemistry 20: 6594-9.

[23] Lambert N, Freedman RB. 1983. Structural properties of homogeneous protein disulphide-isomerase from bovine liver purified by a rapid high-yielding procedure. Biochem. J. 213: 225-34.

[24] Lambert N, Freedman RB. 1983. Kinetics and specificity of homogeneous protein disulphide-isomerase in protein disulphide isomerization and in thiol-protein-disulphide oxidoreduction. Biochem. J. 213: 235-43.

[25] Sanger F, Coulson AR. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. 94: 441-8.

[26] Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 74: 5463-7.

[27] Saiki R, Scharf S, Faloona F, Mullis K, Horn G, et al. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-4.

[28] Bardwell JC, McGovern K, Beckwith J. 1991. Identification of a protein required for disulfide bond formation *in vivo.* Cell 67: 581-9.

[29] Akiyama Y, Kamitani S, Kusukawa N, Ito K. 1992. In vitro catalysis of oxidative folding of disulfide-bonded proteins by the *Escherichia coli dsbA* (*ppfA*) gene product. J. Biol. Chem. 267: 22440-5.

[30] Bardwell JC, Lee JO, Jander G, Martin N, Belin D, Beckwith J. 1993. A pathway for disulfide bond formation *in vivo*. Proc. Natl. Acad. Sci. U. S. A. 90: 1038-42.

[31] Kadokura H, Tian H, Zander T, Bardwell JC, Beckwith J. 2004. Snapshots of DsbA in action: detection of proteins in the process of oxidative folding. Science 303: 534-7.

[32] Missiakas D, Georgopoulos C, Raina S. 1993. Identification and characterization of the *Escherichia coli* gene *dsbB*, whose product is involved in the formation of disulfide bonds in vivo. Proc. Natl. Acad. Sci. U. S. A. 90: 7084-8.

[33] Kishigami S, Kanaya E, Kikuchi M, Ito K. 1995. DsbA-DsbB interaction through their active site cysteines. Evidence from an odd cysteine mutant of DsbA. J. Biol. Chem. 270: 17072-4.

[34] Dutton RJ, Wayman A, Wei JR, Rubin EJ, Beckwith J, Boyd D. 2010. Inhibition of bacterial disulfide bond formation by the anticoagulant warfarin. Proc. Natl. Acad. Sci. U. S. A. 107: 297-301.

[35] Kadokura H, Beckwith J. 2002. Four cysteines of the membrane protein DsbB act in concert to oxidize its substrate DsbA. EMBO J. 21: 2354-63.

[36] Inaba K, Takahashi YH, Fujieda N, Kano K, Miyoshi H, Ito K. 2004. DsbB elicits a red-shift of bound ubiquinone during the catalysis of DsbA oxidation. J. Biol. Chem. 279: 6761-8.

[37] Inaba K, Takahashi YH, Ito K. 2005. Reactivities of quinone-free DsbB from *Escherichia coli.* J. Biol. Chem. 280: 33035-44.

[38] Grauschopf U, Fritz A, Glockshuber R. 2003. Mechanism of the electron transfer catalyst DsbB from *Escherichia coli.* EMBO J. 22: 3503-13.

[39] Zeng H, Snavely I, Zamorano P, Javor GT. 1998. Low ubiquinone content in *Escherichia coli* causes thiol hypersensitivity. J. Bacteriol. 180: 3681-5.

[40] Regeimbal J, Bardwell JC. 2002. DsbB catalyzes disulfide bond formation de novo. J. Biol. Chem. 277: 32706-13.

[41] Regeimbal J, Gleiter S, Trumpower BL, Yu CA, Diwakar M, et al. 2003. Disulfide bond formation involves a quinhydrone-type charge-transfer complex. Proc. Natl. Acad. Sci. U. S. A. 100: 13779-84.

[42] Bader M, Muse W, Ballou DP, Gassner C, Bardwell JC. 1999. Oxidative protein folding is driven by the electron transport system. Cell 98: 217-27.

[43] Bader MW, Xie T, Yu CA, Bardwell JC. 2000. Disulfide bonds are generated by quinone reduction. J. Biol. Chem. 275: 26082-8.

[44] Collet JF, Bardwell JC. 2002. Oxidative protein folding in bacteria. Mol. Microbiol. 44: 1-8.

[45] Kobayashi T, Kishigami S, Sone M, Inokuchi H, Mogi T, Ito K. 1997. Respiratory chain is required to maintain oxidized states of the DsbA-DsbB disulfide bond formation system in aerobically growing *Escherichia coli* cells. Proc. Natl. Acad. Sci. U. S. A. 94: 11857-62.

[46] Kobayashi T, Ito K. 1999. Respiratory chain strongly oxidizes the CXXC motif of DsbB in the *Escherichia coli* disulfide bond formation pathway. EMBO J. 18: 1192-8.

[47] Missiakas D, Georgopoulos C, Raina S. 1994. The *Escherichia coli dsbC* (*xprA*) gene encodes a periplasmic protein involved in disulfide bond formation. EMBO J. 13: 2013-20.

[48] Berkmen M, Boyd D, Beckwith J. 2005. The nonconsecutive disulfide bond of *Escherichia coli* phytase (AppA) renders it dependent on the protein-disulfide isomerase, DsbC. J. Biol. Chem. 280: 11387-94.

[49] Bessette PH, Qiu J, Bardwell JC, Swartz JR, Georgiou G. 2001. Effect of sequences of the active-site dipeptides of DsbA and DsbC on *in vivo* folding of multidisulfide proteins in *Escherichia coli.* J. Bacteriol. 183: 980-8.

[50] Hiniker A, Bardwell JC. 2004. *In vivo* substrate specificity of periplasmic disulfide oxidoreductases. J. Biol. Chem. 279: 12967-73.

[51] Joly JC, Swartz JR. 1997. *In vitro* and *in vivo* redox states of the *Escherichia coli* periplasmic oxidoreductases DsbA and DsbC. Biochemistry 36: 10067-72.

[52] Shouldice SR, Cho SH, Boyd D, Heras B, Eser M, et al. 2010. *In vivo* oxidative protein folding can be facilitated by oxidation-reduction cycling. Mol. Microbiol. 75: 13-28.

[53] Chung J, Chen T, Missiakas D. 2000. Transfer of electrons across the cytoplasmic membrane by DsbD, a membrane protein involved in thiol-disulphide exchange and protein folding in the bacterial periplasm. Mol. Microbiol. 35: 1099-109.

[54] Bessette PH, Cotto JJ, Gilbert HF, Georgiou G. 1999. *In vivo* and *in vitro* function of the *Escherichia coli* periplasmic cysteine oxidoreductase DsbG. J. Biol. Chem. 274: 7784-92.

[55] Missiakas D, Schwager F, Raina S. 1995. Identification and characterization of a new disulfide isomerase-like protein (DsbD) in *Escherichia coli.* EMBO J. 14: 3415-24.

[56] Stewart EJ, Katzen F, Beckwith J. 1999. Six conserved cysteines of the membrane protein DsbD are required for the transfer of electrons from the cytoplasm to the periplasm of *Escherichia coli.* EMBO J. 18: 5963-71.

[57] Rietsch A, Belin D, Martin N, Beckwith J. 1996. An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli.* Proc. Natl. Acad. Sci. U. S. A. 93: 13048-53.

[58] Rietsch A, Bessette P, Georgiou G, Beckwith J. 1997. Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. J. Bacteriol. 179: 6602-8.

[59] Katzen F, Beckwith J. 2000. Transmembrane electron transfer by the membrane protein DsbD occurs via a disulfide bond cascade. Cell 103: 769-79.

[60] Goldstone D, Haebel PW, Katzen F, Bader MW, Bardwell JC, et al. 2001. DsbC activation by the N-terminal domain of DsbD. Proc. Natl. Acad. Sci. U. S. A. 98: 9551-6.

[61] Haebel PW, Goldstone D, Katzen F, Beckwith J, Metcalf P. 2002. The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: crystal structure of the DsbC-DsbDalpha complex. EMBO J. 21: 4774-84.

[62] Goulding CW, Sawaya MR, Parseghian A, Lim V, Eisenberg D, Missiakas D. 2002. Thiol-disulfide exchange in an immunoglobulin-like fold: structure of the N-terminal domain of DsbD. Biochemistry 41: 6920-7.

[63] Rozhkova A, Stirnimann CU, Frei P, Grauschopf U, Brunisholz R, et al. 2004. Structural basis and kinetics of inter- and intramolecular disulfide exchange in the redox catalyst DsbD. EMBO J. 23: 1709-19.

[64] Hayden EC. 2014. Technology: the $1000 genome. Nature 507: 294-5.

[65] Land M, Hauser L, Jun S-R, Nookaew I, Leuze MR, et al. 2015. Insights from 20 years of bacterial genome sequencing. Funct. Integr. Genom. 15: 141-61.

[66] Köser CU, Holden MTG, Ellington MJ, Cartwright EJP, Brown NM, et al. 2012. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. New Engl. J. Med. 366: 2267-75.

[67] Harris SR, Cartwright EJP, Török ME, Holden MTG, Brown NM, et al. 2013. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study. Lancet Infect. Dis. 13: 130-6.

[68] Jorda J, Yeates TO. 2011. Widespread disulfide bonding in proteins from thermophilic archaea. Archaea 2011: 409156.

[69] Beeby M, O'Connor BD, Ryttersgaard C, Boutz DR, Perry LJ, Yeates TO. 2005. The genomics of disulfide bonding and protein stabilization in thermophiles. PLoS Biol. 3: e309.

[70] Mallick P, Boutz DR, Eisenberg D, Yeates TO. 2002. Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. Proc. Natl. Acad. Sci. U. S. A. 99: 9679-84.

[71] Heinemann J, Hamerly T, Maaty WS, Movahed N, Steffens JD, et al. 2014. Expanding the paradigm of thiol redox in the thermophilic root of life. Biochim. Biophys. Acta 1840: 80-5.

[72] Boutz DR, Cascio D, Whitelegge J, Perry LJ, Yeates TO. 2007. Discovery of a thermophilic protein complex stabilized by topologically interlinked chains. J. Mol. Biol. 368: 1332-44.

[73] Pedone E, Ren B, Ladenstein R, Rossi M, Bartolucci S. 2004. Functional properties of the protein disulfide oxidoreductase from the archaeon *Pyrococcus furiosus*: a member of a novel protein family related to protein disulfide-isomerase. Eur. J. Biochem. 271: 3437-48.

[74] Pedone E, D'Ambrosio K, De Simone G, Rossi M, Pedone C, Bartolucci S. 2006. Insights on a new PDI-like family: structural and functional analysis of a protein disulfide oxidoreductase from the bacterium Aquifex aeolicus. J. Mol. Biol. 356: 155-64.

[75] D'Ambrosio K, Pedone E, Langella E, De Simone G, Rossi M, et al. 2006. A novel member of the protein disulfide oxidoreductase family from *Aeropyrum pernix* K1: structure, function and electrostatics. J. Mol. Biol. 362: 743-52.

[76] Pedone E, Limauro D, D'Alterio R, Rossi M, Bartolucci S. 2006. Characterization of a multifunctional protein disulfide oxidoreductase from *Sulfolobus solfataricus. FEBS J.* 273: 5407-20.

[77] Dutton RJ, Boyd D, Berkmen M, Beckwith J. 2008. Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. Proc. Natl. Acad. Sci. U. S. A. 105: 11933-8.

[78] Wang X, Dutton RJ, Beckwith J, Boyd D. 2011. Membrane topology and mutational analysis of *Mycobacterium tuberculosis* VKOR, a protein involved in disulfide bond formation and a homologue of human vitamin K epoxide reductase. Antioxid. Redox Signal. 14: 1413-20.

[79] Li W, Schulman S, Dutton RJ, Boyd D, Beckwith J, Rapoport TA. 2010. Structure of a bacterial homolog of vitamin K epoxide reductase. Nature 463: 507-12.

[80] Dailey FE, Berg HC. 1993. Mutants in disulfide bond formation that disrupt flagellar assembly in *Escherichia coli.* Proc. Natl. Acad. Sci. U. S. A. 90: 1043-7.

[81] Masip L, Pan JL, Haldar S, Penner-Hahn JE, DeLisa MP, et al. 2004. An engineered pathway for the formation of protein disulfide bonds. Science 303: 1185-9.

[82] Chng SS, Dutton RJ, Denoncin K, Vertommen D, Collet JF, et al. 2012. Overexpression of the rhodanese PspE, a single cysteine-containing protein, restores disulphide bond formation to an *Escherichia coli* strain lacking DsbA. Mol. Microbiol. 85: 996-1006.

[83] Rensing C, Mitra B, Rosen BP. 1997. Insertional inactivation of *dsbA* produces sensitivity to cadmium and zinc in *Escherichia coli.* J. Bacteriol. 179: 2769-71.

[84] Hiniker A, Ren G, Heras B, Zheng Y, Laurinec S, et al. 2007. Laboratory evolution of one disulfide isomerase to resemble another. Proc. Natl. Acad. Sci. U. S. A. 104: 11670-5.

[85] Pan JL, Sliskovic I, Bardwell JC. 2008. Mutants in DsbB that appear to redirect oxidation through the disulfide isomerization pathway. J. Mol. Biol. 377: 1433-42.

[86] Chatelle C, Kraemer S, Ren G, Chmura H, Marechal N, et al. 2015. Converting a sulfenic acid reductase into a disulfide bond isomerase. Antioxid. Redox Signal. 23: 945-57.

[87] Zhang L, Chou CP, Moo-Young M. 2011. Disulfide bond formation and its impact on the biological activity and stability of recombinant therapeutic proteins produced by *Escherichia coli* expression system. Biotechnol. Adv. 29: 923-9.

[88] Kinch MS. 2015. An overview of FDA-approved biologics medicines. Drug Discov. Today 20: 393-8.

[89] Ecker DM, Jones SD, Levine HL. 2015. The therapeutic monoclonal antibody market mAbs 7: 9-14.

[90] Lawrence S, Lahteenmaki R. 2014. Public biotech 2013[mdash]the numbers. Nat. Biotechnol. 32: 626-32.

[91] Lee YJ, Lee DH, Jeong KJ. 2014. Enhanced production of human full-length immunoglobulin G1 in the periplasm of *Escherichia coli.* Appl. Microbiol. Biotechnol. 98: 1237-46.

[92] Zhou Y, Liu P, Gan Y, Sandoval W, Katakam AK, et al. 2016. Enhancing full-length antibody production by signal peptide engineering. Microb. Cell Fact. 15: 47.

[93] Yoon SH, Kim SK, Kim JF. 2010. Secretory production of recombinant proteins in *Escherichia coli.* Recent Pat. Biotechnol. 4: 23-9.

[94] Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M. 2012. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb. Cell Fact. 11: 56.

[95] Bessette PH, Aslund F, Beckwith J, Georgiou G. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc. Natl. Acad. Sci. U. S. A. 96: 13703-8.

[96] Robinson MP, Ke N, Lobstein J, Peterson C, Szkodny A, et al. 2015. Efficient expression of full-length antibodies in the cytoplasm of engineered bacteria. Nat. Commun. 6: 8072.

[97] Magdelaine-Beuzelin C, Pinault C, Paintaud G, Watier H. 2010. Therapeutic antibodies in ophthalmology: old is new again. mAbs 2: 176-80.

[98] Mazor Y, Van Blarcom T, Mabry R, Iverson BL, Georgiou G. 2007. Isolation of engineered, full-length antibodies from libraries expressed in *Escherichia coli.* Nat. Biotechnol. 25: 563-5.

[99] Burton DR, Barbas CF, 3rd, Persson MA, Koenig S, Chanock RM, Lerner RA. 1991. A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. Proc. Natl. Acad. Sci. U. S. A. 88: 10134-7.

[100] Nixon AE, Sexton DJ, Ladner RC. 2014. Drugs derived from phage display: from candidate identification to clinical practice. mAbs 6: 73-85

[101] Dombkowski AA, Sultana KZ, Craig DB. 2014. Protein disulfide engineering. FEBS Lett. 588: 206-12.

[102] Sauer RT, Hehir K, Stearman RS, Weiss MA, Jeitler-Nilsson A, et al. 1986. An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of lambda repressor. Biochemistry 25: 5992-8.

[103] Wetzel R, Perry LJ, Baase WA, Becktel WJ. 1988. Disulfide bonds and thermal stability in T4 lysozyme. Proc. Natl. Acad. Sci. U. S. A. 85: 401-5.

[104] Wells JA, Powers DB. 1986. *In vivo* formation and stability of engineered disulfide bonds in subtilisin. J. Biol. Chem. 261: 6564-70.

[105] Hagihara Y, Saerens D. 2014. Engineering disulfide bonds within an antibody. Biochim. Biophys. Acta 1844: 2016-23.

[106] Clarke J, Fersht AR. 1993. Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. Biochemistry 32: 4322-9.

[107] Heikoop JC, van den Boogaart P, Mulders JW, Grootenhuis PD. 1997. Structure-based design and protein engineering of intersubunit disulfide bonds in gonadotropins. Nat. Biotechnol. 15: 658-62.

[108] Arredondo S, Segatori L, Gilbert HF, Georgiou G. 2008. *De novo* design and evolution of artificial disulfide isomerase enzymes analogous to the bacterial DsbC. J. Biol. Chem. 283: 31469-76.

[109] Peek JA, Taylor RK. 1992. Characterization of a periplasmic thiol: disulfide interchange protein required for the functional maturation of secreted virulence factors of Vibrio cholerae. Proc. Natl. Acad. Sci. U. S. A. 89: 6210-4.

[110] Barbieri JT, Sun J. 2005. In Reviews of Physiology, Biochemistry and Pharmacology, pp. 79-92. Berlin, Heidelberg: Springer Berlin Heidelberg.

[111] Miki T, Okada N, Danbara H. 2004. Two periplasmic disulfide oxidoreductases, DsbA and SrgA, target outer membrane protein SpiA, a component of the Salmonella pathogenicity island 2 type III secretion system. J. Biol. Chem. 279: 34631-42.

[112] Godlewska R, Dzwonek A, Mikuła M, Ostrowski J, Pawłowski M, et al. 2006. *Helicobacter pylori* protein oxidation influences the colonization process. Int. J. Med. Microbiol. 296: 321-4.

[113] Stenson TH, Weiss AA. 2002. DsbA and DsbC are required for secretion of pertussis toxin by *Bordetella pertussis.* Infect. Immun. 70: 2297-303.

[114] Okamoto K, Nomura T, Fujii Y, Yamanaka H. 1998. Contribution of the disulfide bond of the A subunit to the action of *Escherichia coli* heat-labile enterotoxin. J. Bacteriol. 180: 1368-74.

[115] Lasica AM, Jagusztyn-Krynicka EK. 2007. The role of Dsb proteins of Gram-negative bacteria in the process of pathogenesis. FEMS Microbiol. Rev. 31: 626-36.

[116] Wülfing C, Rappuoli R. 1997. Efficient production of heat-labile enterotoxin mutant proteins by overexpression of dsbA in a degP-deficient *Escherichia coli* strain. Arch. Microbiol. 167: 280-3.

[117] Foreman DT, Martinez Y, Coombs G, Torres A, Kupersztoch YM. 1995. TolC and DsbA are needed for the secretion of STB, a heat-stable enterotoxin of *Escherichia coli.* Mol. Microbiol. 18: 237-45.

[118] Agudo D, Mendoza MT, Castañares C, Nombela C, Rotger R. 2004. A proteomic approach to study *Salmonella typhi* periplasmic proteins altered by a lack of the DsbA thiol: disulfide isomerase. Proteomics 4: 355-63.

[119] Zhang H-Z, Donnenberg MS. 1996. DsbA is required for stability of the type IV pilin of enteropathogenic *Escherichia coli.* Mol. Microbiol. 21: 787-97.

[120] Lee Y, Kim Y, Yeom S, Kim S, Park S, et al. 2008. The role of disulfide bond isomerase A (DsbA) of *Escherichia coli* O157: H7 in biofilm formation and virulence. FEMS Microbiol. Lett. 278: 213-22.

[121] Smith RP, Paxman JJ, Scanlon MJ, Heras B. 2016. Targeting bacterial Dsb proteins for the development of anti-virulence agents. Molecules 21: 811-25.

[122] Heras B, Shouldice SR, Totsika M, Scanlon MJ, Schembri MA, Martin JL. 2009. DSB proteins and bacterial pathogenicity. Nat. Rev. Microbiol. 7: 215-25.

[123] Singh AK, Bhattacharyya-Pakrasi M, Pakrasi HB. 2008. Identification of an atypical membrane protein involved in the formation of protein disulfide bonds in oxygenic photosynthetic organisms. J. Biol. Chem. 283: 15762-70.

[124] Früh V, Zhou Y, Chen D, Loch C, Eiso AB, et al. 2010. Application of fragment based drug discovery to membrane proteins: biophysical identification of ligands of the integral membrane enzyme DsbB. Chem. Biol. 17: 881-91.

[125] Halili MA, Bachu P, Lindahl F, Bechara C, Mohanty B, et al. 2015. Small molecule inhibitors of disulfide bond formation by the bacterial DsbA-DsbB dual enzyme system. ACS Chem. Biol. 10: 957-64.

[126] Landeta C, Blazyk JL, Hatahet F, Meehan BM, Eser M, et al. 2015. Compounds targeting disulfide bond forming enzyme DsbB of Gram-negative bacteria. Nat. Chem. Biol. 11: 292-8.

[127] Inaba K, Murakami S, Suzuki M, Nakagawa A, Yamashita E, et al. 2006. Crystal structure of the DsbB-DsbA complex reveals a mechanism of disulfide bond generation. Cell 127: 789-801.

[128] Duprez W, Premkumar L, Halili MA, Lindahl F, Reid RC, et al. 2015. Peptide inhibitors of the *Escherichia coli* DsbA oxidative machinery essential for bacterial virulence. J. Med. Chem. 58: 577-87.

[129] Consortium TU. 2015. UniProt: a hub for protein information. Nucleic Acids Res. 43: D204-12.

[130] Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res. 43: W401-7.