Comparative Genomics of Trace Elements: Emerging Dynamic View of Trace Element Utilization and Function

Yan Zhang and Vadim N. Gladyshev*

Department of Biochemistry and Redox Biology Center, University of Nebraska, Lincoln, Nebraska 68588-0664

Received December 31, 2008

Contents

1. Introduction

Biological trace elements are needed in minute quantities for proper growth, development, and physiology of organisms.[1−3] These micronutrients include iron (Fe), zinc (Zn), copper (Cu), molybdenum (Mo), tungsten (W), nickel (Ni), cobalt (Co), manganese (Mn), chromium (Cr), vanadium (V), selenium (Se), iodine (I), and possibly other trace elements. Although used in small amounts (generally less than 100 mg/day) as opposed to macrominerals (e.g., calcium, magnesium, phosphorus, sulfur, potassium, and sodium) that are required in larger quantities, trace elements provide proteins with unique coordination, catalytic, and electron transfer properties. These properties are employed by organisms in key functions in a variety of pathways, resulting in the dependence of organisms on various trace elements.[4]

The majority of trace elements are metal ions. Except for Fe and Zn, which are thought to be utilized by all organisms,[5,6] most other biological metals, including Cu, Mn, Mo, Ni, and Co, are used in a wide range but not in all organisms in both prokaryotes and eukaryotes. The major metalloid micronutrient, Se, is also involved in a variety of important metabolic processes, but not used by all organisms.[7−9]

Because of the important roles these trace elements play in cells, efficient mechanisms are required to maintain and regulate uptake of these micronutrients and their concentration, utilization, and storage, especially for those elements whose soluble forms are present in trace amounts in natural environments. High-affinity import systems have been reported for several trace elements in both prokaryotes and eukaryotes. The ATP-binding cassette (ABC) transport systems are the most frequently used transporters, such as ZnuABC for Zn,[10] MntABC for Mn,[11] ModABC for Mo,[12] and NikABCDE for Ni.[13] Non-ABC transporters were also identified, e.g., ZupT for Zn and other divalent metal cations,[14] MntH for Mn and Fe,[15] Ctr1 for Cu,[16] and NiCoT for Ni and Co.[17] Some metal ions, such as Fe, Zn, and Mn, could also be transported via unspecific cation influx systems.[18,19] Although a high-affinity transport system for Se has not been identified thus far, its uptake (in the form of selenate or selenite) could be supported by the sulfate transport system.[20] On the other hand, excessive uptake of

* Corresponding author, Telephone: (402) 472-4948. Fax: (402) 472-7842. E-mail: vgladyshev1@unl.edu.

10.1021/cr800557s CCC: $71.50
Published on Web 05/21/2009

Yan Zhang was born in Hangzhou, China, in 1974. He received his B.S. degree from Zhejiang Medical University in 1997 and his Ph.D. degree in biomedical engineering from the Chinese Academy of Medical Sciences and Peking Union Medical College (Beijing, China) in 2002. Since 2002, he has been in Professor Vadim Gladyshev's laboratory at the University of Nebraska—Lincoln, where he worked on *in silico* identification of selenoproteins in prokaryotes and investigation of the evolution of selenium utilization. More recently, he carried out research on comparative genomics of trace elements, which include molybdenum, copper, nickel, cobalt, and selenium, in all three domains of life. He also investigated selenoprotein gene expression in mouse brain and carried out other computational biology projects. His main themes of interest focus on the understanding of trace element utilization and evolution on the basis of comparative genomics and other bioinformatics approaches.

Vadim N. Gladyshev is a Charles Bessey Professor of Biochemistry, University of Nebraska—Lincoln. He received his B.S./M.S. and Ph.D. degrees in biochemistry from Moscow State University and did postdoctoral work at the National Institutes of Health before joining the faculty at University of Nebraska—Lincoln in 1998. He is also a Director of the Redox Biology Center, an NIH-supported Center of Biomedical Research Excellence. His laboratory works at the interface of selenium and redox biology.

certain metals (e.g., Fe or Cu) can result in overload disorders because of high toxicity of these metals. Moreover, some trace elements may interact and could potentially interfere with the essential functions of each other. For example, excessive Zn can induce signs of Cu deficiency.[21,22] Therefore, homeostasis of trace elements within the cell should be carefully maintained by mechanisms regulating their uptake in order to provide adequate levels while preventing accumulation to toxic levels.[1,3]

Utilization of trace elements is generally rather complex. First, cells require metal ions as cofactors of metalloproteins. Most metal ions are directly incorporated into their cognate sites in proteins involved in various cellular pathways. However, some metal ions have to become part of prosthetic groups, cofactors, or complexes prior to insertion of these moieties into target proteins. For example, Mo and Co are the main functionalities in molybdopterin (MPT, or Mo cofactor, Moco) and vitamin B_{12} (or cobalamin, a group of closely related polypyrrole compounds such as cyanocobalamin, methylcobalamin, and deoxyadenosyl cobalamin), respectively.[23−25] Second, several metals, such as Fe, Cu, and Mo, play important roles in redox reactions, but these and other metals are also involved in a variety of other enzymatic reactions. Third, the number of metalloprotein families that are dependent on one or more metal varies greatly. For instance, over 300 protein families have been identified that require Zn,[26,27] whereas less than 10 protein families are known to be dependent on Ni.[28] In contrast to metals, Se is mainly used in the form of the amino acid, selenocysteine (Sec, known as the 21st amino acid), found in a variety of proteins (selenoproteins),[29] and selenouridine, a modified tRNA nucleoside that tunes codon−anticodon interactions in prokaryotes.[30]

In recent years, dramatic advances in genomics have resulted in the generation of complete genomic sequences for a large number of organisms from the three domains of life. Computational analyses of protein sequences and structures on a genomic scale revealed a significant number of proteins that bind trace elements. Thus, identification of all or almost all trace element-containing proteins in genomic databases can greatly assist in understanding of utilization of different micronutrients. Unfortunately, due to the lack of reliable approaches, it is currently not possible to identify complete sets of trace element-containing proteins in organisms. The only exception may be Se, for which complete or nearly complete sets of selenoproteins in both prokaryotes and eukaryotes, including humans, have been reported.[31−33] In recent years, several comparative and functional genomic analyses have been carried out for trace elements, including Zn, Ni, Co, Cu, and Mo.[34−41] These preliminary studies improved our understanding of their biological and biomedical roles, as well as their occurrence and trends in the utilization of these trace elements.

In this review, we discuss the use of five biological trace elements (Mo, Ni, Co, Cu, and Se) in the three domains of life from the perspective of comparative genomics. The reason we chose these elements is that they are widely used by organisms but are characterized by a limited number of user proteins. In addition, unlike Fe and Zn, they are not used by all organisms. Therefore, the analysis of their utilization may provide important information with regard to fundamental issues of function and evolution of trace elements. We will begin with a brief introduction of the strategies to analyze the utilization of trace elements. After that, each of these trace elements will be discussed in detail, including their transport systems, cofactor biosynthesis pathways, and proteins that use them. We will then proceed with the molecular basis for comparative genomics and recent comparative studies on trace element utilization. We will finish with the interactions among trace elements and unique patterns of their utilization.

2. Comparative genomic analysis of trace element utilization

Comparative genomics is a research area that examines the relationships of genome structure and function, as well as its genes and other functional elements, across species.[42−44] It takes advantage of the information provided by the evolutionary signatures of selection to understand the func-

tion and processes that act on genomes and their components. The wide use of comparative genomics has been covered in detail in many reviews and benchmark studies,[42−46] and it is not our intention to discuss this subject in detail here. While comparative genomics is still a young field, it holds great promise in providing insights into many aspects of the evolution of ancestral and modern species, which in turn helps aid the understanding of pathways and other biological processes in currently living organisms. These considerations include trace elements whose unique chemical, redox, and coordination properties are exploited in a variety of ways in biological systems.[47]

In the past decade, complete genome sequencing transformed the way science is done. These advances also led to ever increasing information about sequences, structures, and functions of metalloproteins as well as their occurrence across thousands of organisms. It is now possible to compile sets of trace element-dependent pathways that an organism adopts. Questions can now be addressed regarding the extent of the use of trace elements in individual organisms, as well as in groups of organisms, based exclusively on computational analyses.

However, the prediction power of these methods is not uniform across trace elements. For example, unlike Se, whose insertion into proteins (in the form of Sec) could be reliably predicted on the basis of several specific features (details will be discussed in section 6.3.1), a precise approach has not been developed for the identification of metalloproteins, partially because of overlapping signatures for different metals or the uncertainty of metal-binding ligands. On the other hand, analysis of sequence and structural properties of many metal-containing proteins and the conservation of metal-binding ligands in these proteins resulted in the development of numerous metal-binding motifs/patterns, which can assist in the analysis of metal utilization and help identify additional metal-binding proteins. Furthermore, comparative genomics and phylogenetic analyses provide important information regarding the function and evolution of trace element utilization. Finally, searches for trace element utilization traits can be assisted with the analyses of factors involved in metal transport or biosynthesis of metal-containing cofactors utilized by proteins. In general, the procedure of comparative genomic analysis of a trace element could be divided into three major steps (Figure 1).

2.1. Step 1: Identification of Trace Element-Dependent Proteins

Occurrence of trace element-dependent proteins (e.g., metalloproteins, selenoproteins) is the most important indicator of utilization of the corresponding trace elements. Therefore, the first step in the analysis may be to collect all known trace element-dependent proteins.

Metalloproteins that use Mo, Ni, Co or Cu belong to diverse classes of proteins, with the inserted metal atoms providing catalytic, regulatory, and structural roles critical to protein function. A previous examination of the Protein Data Bank (PDB) showed that almost one-fourth of structurally characterized proteins contain a coordinated metal atom.[48] Based on sequence and structural information of a variety of known metalloproteins, several bioinformatics tools and databases were developed that provide predictions for sequences of interest. These resources, for example, include Pfam (the protein families database),[49] PROSITE (a database of protein domains, families, and functional sites),[50] PRINTS (a protein fingerprint database),[51] CDD (a collection of multiple sequence alignments and derived database search models at NCBI),[52] BLOCKS (a block database),[53] ProDom (a comprehensive set of protein domain families automatically generated from the SWISS-PROT and TrEMBL databases),[54] COG (clusters of orthologous groups),[55] and MDB (a metalloprotein database).[56] For metalloprotein prediction, some of these tools contain sequence patterns, which are the occurrences of specific amino acids (ligands) in protein sequences, whereas other tools use position-specific scoring matrices (PSSM, or profiles), which describe similarity among distant sequences. However, these databases and tools do not include all metal-binding motifs and could only help identify a partial set of metal-binding proteins with known structures or conserved metal-binding residues. Moreover, some metalloproteins may bind different metals with the same ligands. In addition, the protein folding location and other factors may also regulate metal selection in organisms.[57]

Occurrence of genes involved in high-affinity metal transport, metal-containing cofactor (such as Moco and B_{12}) biosynthesis, and other regulatory pathways (such as chaperones and repressors) may provide complementary information regarding metal utilization and should be analyzed in parallel. Thus, a metal utilization trait (i.e., occurrence of at least one protein that utilizes this metal) could also be verified by the requirement for the presence of high-affinity transport systems and/or cofactor biosynthesis pathways. Identification of selenoproteins (both known and new) is easier, as the searches may be based on specific features of Sec insertion machinery. Several bioinformatics tools have been developed for prediction of selenoproteins, and it appears that they are capable of detecting all or almost all selenoprotein families in sequenced organisms (details will be discussed in section 6.3).

2.2. Step 2: Identification of Orthologs of Query Proteins

The next step may be to identify orthologs of selected proteins in the sequenced genomes of archaea, bacteria, and eukaryotes. A list of fully sequenced organisms from the three domains of life is available at NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). At the time of prepation of this article (November 2008), more than 700 species were available (excluding various strains of the same organism).

To analyze for the occurrence of orthologs of query proteins, a set of sequences obtained in step 1 can be used as initial seeds to search for homologous sequences in various organisms via BLAST programs (such as BLASTP and TBLASTN).[58] Distant homologs could be further identified using iterative BLAST searches in each phylum or with the help of PSI-BLAST. Orthologous proteins could then be defined using multiple approaches, such as conserved domain (COG/Pfam/CDD) searches, bidirectional best hits (BBH), and genomic context (i.e., operon or gene neighborhood) analyses and phylogenetic analyses. If metal-binding ligands are known, conservation of these residues in the orthologs should be investigated to assess the ability to bind metal. For some proteins, it is difficult to identify metal-dependent orthologs or determine metal preference. For example, Ni and Co share common transport systems in archaea and bacteria, and members of these transporter families in some organisms often do not have a clearly discernible function.[39]

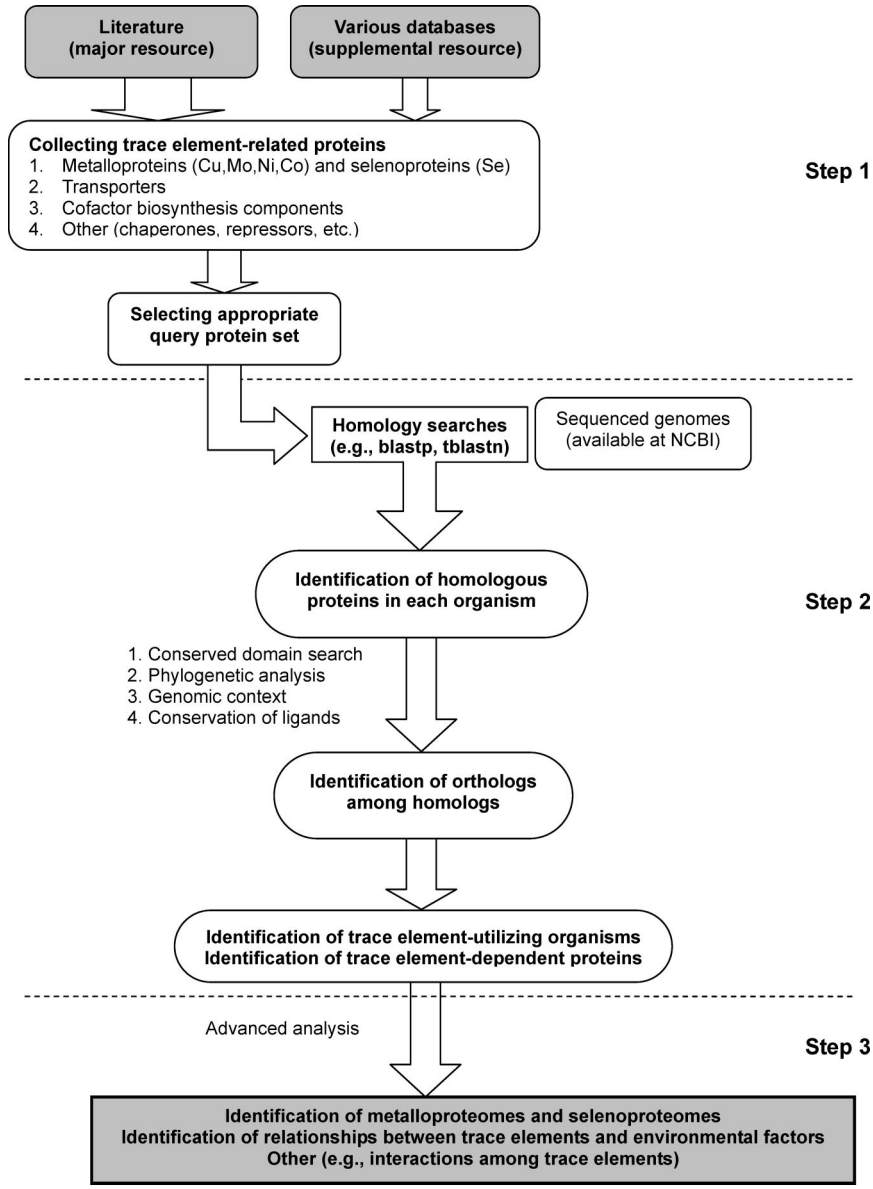

Figure 1. Schematic diagram for comparative genomic analyses of trace element utilization. This process can be divided into three steps. Details are discussed in the text.

Occurrence of Moco and vitamin B_{12} biosynthesis could be verified by the presence of proteins involved in the corresponding pathways (see sections 3.1 and 4.3 for details). Occurrence of the Sec biosynthesis pathway could be easily identified based on the co-occurrence of components involved in this process (see section 6.2). The presence of the utilization trait of a trace element X in an organism could be verified by the requirement for occurrence of at least one predicted X-specific transporter (if available) or X-containing cofactor biosynthesis pathway (if available) or at least one X-dependent protein.

A potential disadvange of this approach is that only proteins strictly specific for a particular metal must be selected, which may result in incomplete analysis of metal utilization in some organisms. However, regarding the four metals discussed here, most of the metal-utilizing proteins are strictly dependent on their primary metal. Therefore, such a comparative genomic approach may indeed reveal a general (even if incomplete) picture of metal utilization in organisms.

2.3. Step 3: Comparative and Functional Analyses of Trace Element Utilization

Functional analysis of trace element utilization by means of comparative genomics is one of the important goals of the field, which enhances our understanding of general mechanisms and evolutionary dynamics of trace elements used in different organisms, phyla, or even kingdoms. Based on the data derived from the two previous steps and subsequent analyses, additional questions could be addressed, such as identification of the relationship between trace element utilization and environmental factors, characterization of metalloproteomes and selenoproteomes, and interactions or other common features among trace elements. Initial studies have recently been reported for these trace elements on the basis of comparative genomics. In the following sections, we will focus on each of the five trace elements and discuss unique and common features of their utilization.

3. Molybdenum

Mo plays a critical role in several metabolic pathways and forms part of the active sites of several metalloenzymes in bacteria, archaea, and eukaryotes.[59] These enzymes execute reactions in the biogeochemical cycles of carbon, nitrogen, and sulfur of the Earth and occupy key positions in several metabolic pathways.[60] Except for the Fe−Mo cofactor in nitrogenase,[61] Mo is complexed by a pterin moiety, thereby generating the molybdenum cofactor (Moco) in Mo-dependent enzymes (molybdoenzymes).[59,60] Some microorganisms utilize tungsten that is also coordinated by pterin.[62] The close chemical and physical similarities of tungstate and molybdate are based on similar atomic and ionic radii, electronegativity, and coordination properties.[63,64] Although Mo use is much more widespread than that of tungsten, it is often impossible to distinguish these elements based on sequence analysis. Therefore, in this review, the term Moco refers to the utilization of both metals.

3.1. Mo Uptake and Moco Biosynthesis

Synthesis of Moco and conversion of apoenzymes to the active forms of molybdoenzymes depend on transport of molybdate anion, activation of molybdate, and finally incorporation of the activated Mo into molybdopterin.[65,66]

Identification of Mo (or W) transport systems and the Moco biosynthesis pathway are essential for characterization of the Mo utilization trait. In *Escherichia coli*, the high-affinity molybdate ABC transporter (ModABC, products of *modABC* genes) consists of ModA (molybdate-binding protein), ModB (membrane integral channel protein), and ModC (cytoplasmic ATPase).[67,68] As mentioned above, due to the close chemical and physical similarities between molybdate and tungstate, the latter was previously thought to be only unselectively cotransported or cometabolized with molybdate by the ModABC system. However, a new class of molybdate/tungstate transport system (WtpABC) and a highly specific tungstate ABC transporter (TupABC) have recently been reported.[69,70] WtpA (the periplasmic component of the WtpABC transport system) contains a ModA-like domain (COG0725, ModA), whereas TupA (the periplasmic component of the TupABC transport system) contains a different domain (COG4662, TupA), which has a low similarity to ModB. They showed different anion affinities compared to ModA. TupA specifically binds tungstate, whereas WtpA has a higher affinity for tungstate than ModA and TupA, and the affinity for molybdate is similar to that of ModA. Crystal structures showed that the residues involved in molybdate binding in *E. coli* ModA (1AMF) and tungstate binding in *Archaeoglobus fulgidus* WtpA (2ONS) were partially overlapped.[71,72] Therefore, it is now clear that tungstate can also be selectively transported into some prokaryotic cells by these two selective ABC-type transporters and may be selectively incorporated into some Moco-dependent enzymes.[73] Alignment of ModA and WtpA sequences from different species showed that the residues involved in substrate binding are not strictly conserved, not only between WtpA and ModA, but also within the same family (Figure 2).

In *E. coli*, the *modABC* operon is regulated by a repressor protein, ModE, which may also control the transcription of genes coding for molybdopterin synthesis and molybdoenzymes.[74−77] *E. coli* ModE is composed of an N-terminal DNA-binding domain (ModE_N) and a C-terminal molybdate-binding domain.[74,75,78] The latter contains a tandem repeat of the Mo-binding protein (Mop; the C-terminal domain is also referred to as the Di-Mop domain).[75] The ModABC-ModE systems are widely distributed in organisms but are not ubiquitous.[79−82] Several recent studies showed ModE-like variants in some prokaryotes (including separate ModE_N, Mop/Di-Mop proteins as well as their additional fusion forms), suggesting the presence of more complex ModE-related ModABC regulation in these organisms.[41,82,83] At the same time, regulation of WtpABC and TupABC transporters is unclear. However, the observation that the *modE* gene is located close or next to *tupABC* or *wtpABC* operons suggests that the two transport systems are regulated, at least in some organisms, by ModE-related mechanisms.[41]

In contrast to the well-characterized molybdate transport in prokaryotes, eukaryotic molybdate transport and its regulation are poorly understood. In 2007, a high-affinity molybdate transport system, MOT1, which belongs to the sulfate transporter superfamily, was first characterized in *Arabidopsis thaliana* and *Chlamydomonas reinhardtii*.[84,85] More recent studies showed that *A. thaliana* MOT1 is strongly expressed in the roots and is localized to the mitochondria instead of the plasma membrane of root cells.[86]

Moco is synthesized by an evolutionarily conserved multistep pathway in all three domains of life. Details have been described in many articles and reviews.[60,87−91] Although some steps in the Moco biosynthesis remain incompletely understood, the overall pathway can be divided into three or four steps as outlined in Figure 3: (i) early steps in which a guanosine derivative, most likely GTP, is converted into precursor Z; (ii) transformation of precursor Z into molybdopterin; (iii) metal incorporation into the apo-cofactor; (iv) further activities that are required to generate an active cofactor in some organisms. For example, most enzymes from eubacteria contain a dinucleotide form of the cofactor (molybdopterin guanine dinucleotide, MGD) in which a second nucleotide, such as GMP or CMP, is linked to the organic component of the cofactor. In *E. coli*, the proteins required for biosynthesis and regulation of the pterin cofactor are encoded in the *moa-mog* operon.[92,93] The *moa* and *moe* operons encode proteins involved in the biosynthesis of the mononucleotide form of the pterin cofactor, and the *mob* operon encodes pterin guanine dinucleotide synthase that adds GMP to the Mo-complexed pterin cofactor. In eukaryotes, six proteins (Cnx1−3 and Cnx5−7, as designated in plants) are involved in Moco biosynthesis.[93−97] These proteins are homologous to their counterparts in bacteria. Thus, the *moa-mog* genes could be used for identification of Moco biosynthesis in prokaryotes whereas *cnx* genes could be used in eukaryotes.

3.2. Molybdoenzymes

Molybdoenzymes represent the metabolic pathways that Mo is involved in. As mentioned above, Moco-containing enzymes catalyze important redox reactions in the global carbon, nitrogen, and sulfur cycles. To date, more than 50 molybdoenzymes, mostly of bacterial origin, have been described.[60,88,89] On the basis of sequence similarity and spectroscopic properties, these enzymes can be divided into four families: xanthine oxidase (XO), sulfite oxidase (SO), dimethylsulfoxide reductase (DMSOR), and aldehyde:ferredoxin oxidoreductase (AOR).[98] Each family includes a variety of subfamilies based on substrate preferences. Table 1 includes the majority of known Moco-containing enzymes.

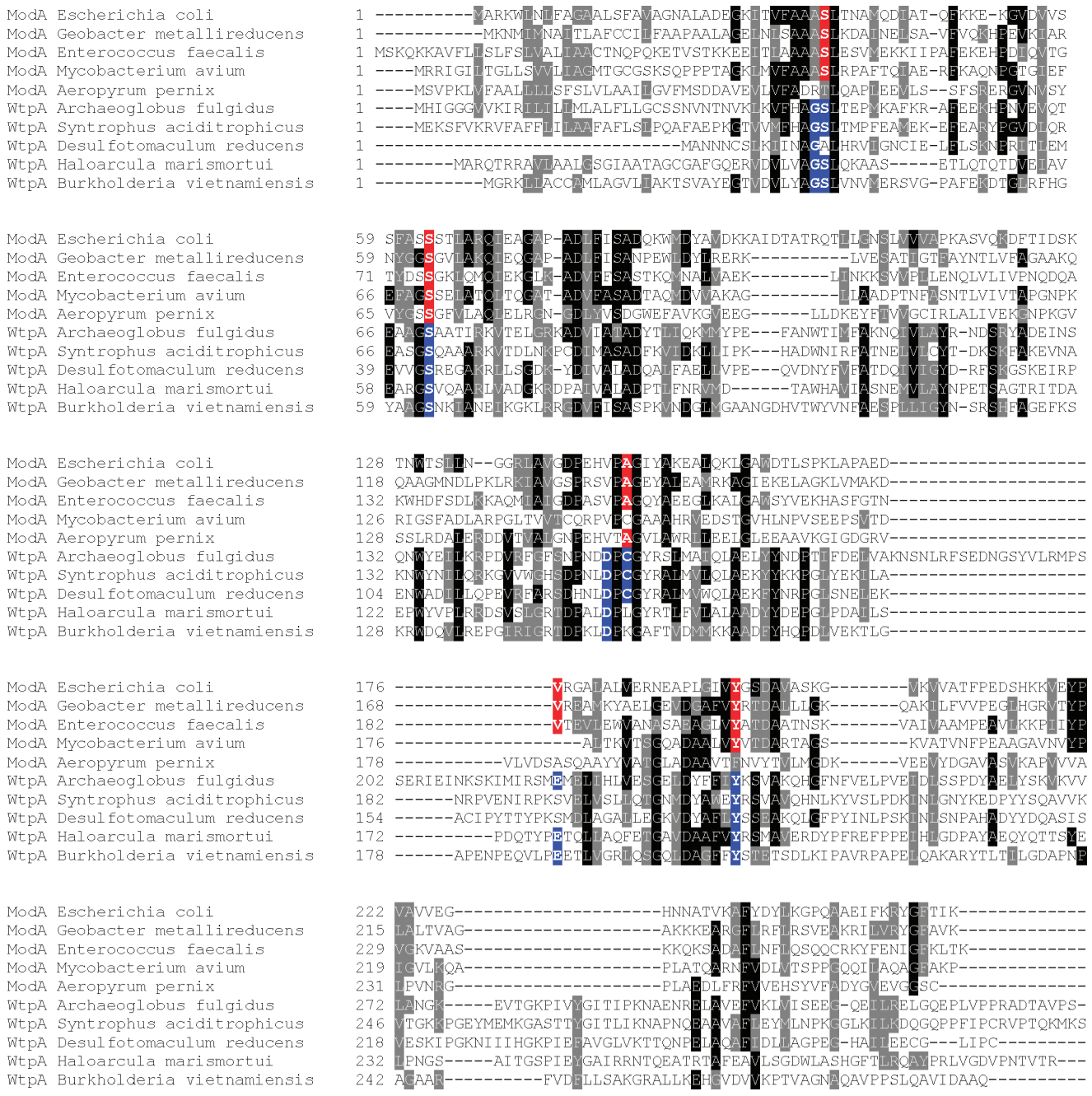

Figure 2. Multiple sequence alignment of ModA and WtpA. Based on the resolved crystal structures,[71,72] residues involved in molybdate binding in *E. coli* ModA are shown in a red background and residues involved in tungstate binding in *Archaeoglobus fulgidus* WtpA in a blue background. Some of these residues overlap between ModA and WtpA proteins. Other conserved residues are shown in white on black or gray.

All four families occur in prokaryotes, but only a limited number of enzymes belonging to the XO and SO families occur in eukaryotes.[60] Members of the XO family are broadly distributed, with many eukaryotic, prokaryotic, and archaeal representatives. On the other hand, SO proteins are largely found in eukaryotes while members of the DMSOR family have only been found in bacteria and archaea. In eukaryotes, the SO family includes nitrate reductase (NR) and SO, whereas the XO family is mainly represented by xanthine dehydrogenase (XDH) and aldehyde oxidase (AO). XDH, AO, and SO are typical of many eukaryotes analyzed so far, whereas NR is only present in autotrophic organisms (e.g., plants, algae, and fungi), as this enzyme is required for nitrate assimilation.

Members of the XO family catalyze oxidative hydroxylation of a wide range of aldehydes and aromatic heterocycles.[88] The overall reaction mechanism of these enzymes is typically broken down into reductive and oxidative half-reactions of the catalytic cycle, and the Mo center participates in the former. The major enzymes in this family include AO (catalyzes the oxidation of a variety of aromatic and nonaromatic heterocycles and aldehydes), XDH (a key enzyme of purine degradation that oxidizes hypoxanthine to xanthine and xanthine to uric acid), and bacterial aldehyde oxidoreductase.

The SO family consists of sulfite oxidase and assimilatory nitrate reductase. Members of this family catalyze net oxygen atom transfer to or from a heteroatom lone electron pair

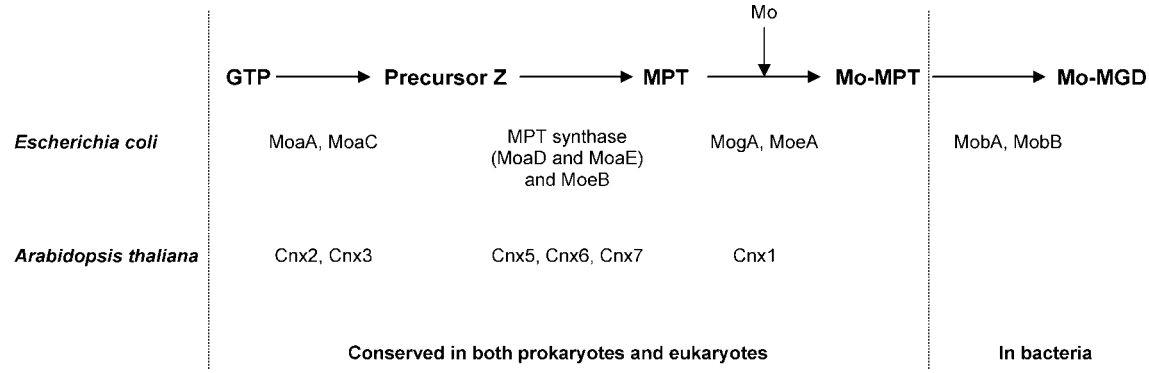

Figure 3. Biosynthesis of molybdenum cofactor. The pathway of Moco synthesis can be divided into three or four steps. The proteins from *E. coli* and *A. thaliana* catalyzing the respective steps are depicted and their names are given. MPT, molybdopterin; MGD, molybdopterin guanine dinucleotide.

Table 1. Moco-Containing Enzymes

family	protein
xanthine oxidase	xanthine oxidase
	xanthine dehydrogenase
	aldehyde oxidase
	aldehyde oxidoreductase
	CO dehydrogenase
	quinoline-2-oxidoreductase
	isoquinoline 1-oxidoreductase
	quinoline-4-carboxylate-2-oxidoreductase
	quinaldine-4-oxidoreductase
	quinaldic acid 4-oxidoreductase
	nicotinic acid hydroxylase
	6-hydroxynicotinate hydroxylase
	nicotine dehydrogenase
	picolinate hydroxylase
	(2*R*)-hydroxycarboxylate oxidoreductase
sulfite oxidase	sulfite oxidase
	nitrate reductase (assimilatory)
dimethylsulfoxide reductase	dimethylsulfoxide reductase
	biotin sulfoxide reductase
	trimethylamine-*N*-oxide reductase
	nitrate reductase (dissimilatory)
	formate dehydrogenase
	formylmethanofuran dehydrogenase
	polysulfide/thiosulfate/arsenate reductase
	arsenite oxidase
aldehyde:ferredoxin oxidoreductase	aldehyde:ferredoxin oxidoreductase

rather than hydroxylation of a carbon center. Sulfite oxidase, the name-giving enzyme for members of the SO family, catalyzes the oxidation of sulfite to sulfate, the final step in the degradation of sulfur-containing amino acids.[60] The assimilatory nitrate reductase catalyzes the reduction of nitrate to nitrite and is responsible for the first step in the uptake and utilization of nitrate.[60]

The DMSOR family consists of a number of Moco-binding enzymes, all from bacterial and archaeal sources, exhibiting substantial sequence homology that justifies their grouping into a single family. Some of these enzymes possess Mo as their sole redox-active center.[88] Furthermore, the DMSOR family is characterized by the coordination of two pyranopterin-ene-1,2-dithiolate ligands in their active sites, which is distinctive among the other Mo enzymes.[99] Among those, DMSOR (a soluble protein found in the periplasmic space of bacteria that catalyzes reductive deoxygenation of dimethyl sulfoxide to dimethyl sulfide), formate dehydrogenase (catalyzes the oxidation of formate to bicarbonate), dissimilatory (or respiratory) nitrate reductase (couples the reduction of nitrate to nitrite at the expense of menaquinol to generate a transmembrane proton gradient), and trimethylamine-*N*-oxide (TMAO) reductase (catalyzes the reduction of TMAO to trimethylamine) are prominent. Recent phylogenetic analyses of two DMSOR subfamilies, arsenite oxidase and respiratory arsenate reductase, revealed different evolutionary histories although both enzymes are known to mediate the bioenergetic use of arsenics. The emergence of arsenite oxidase that is responsible for the biological oxidation of arsenite to arsenate is probably prior to the Archaea/Bacteria split, whereas respiratory arsenate reductase appeared to originate in the domain Bacteria after the Bacteria/Archaea divergence.[100]

AOR was the first enzyme that was structurally characterized as a protein containing a Moco-type cofactor. The AOR from *Pyrococcus furiosus* contains a molybdopterin-based tungsten cofactor that is analogous to the Mo cofactor.[62] This enzyme has been proposed to be the primary enzyme responsible for the interconversion of aldehydes and carboxylates in archaea.[101]

3.3. Comparative Genomics of Mo Utilization

Although Mo is an important transition metal (essential in many organisms), almost all previous studies focused on the identification of Mo uptake systems, Moco biosynthesis pathways, and Mo-dependent enzymes in individual organisms. In contrast, comparative analyses of Mo utilization have been lacking. An early investigation of the ModABC-ModE system in deltaproteobacteria revealed that all analyzed deltaproteobacteria have ModABC transporters, whereas the full-length ModE was only observed in a few of them.[83] To compare the Mo utilization trait among sequenced organisms, we applied the comparative genomic approaches described in section 2 to examine the occurrence and dynamics of Mo utilization in bacteria, archaea, and eukaryotes at the level of (i) Mo transport and Moco utilization trait, and (ii) Mo-dependent enzymes.[41] This study provided a first glance at Mo utilization in the three domains of life and showed its wide occurrence, yet limited use of this metal in individual organisms. Distribution of Mo-utilizing organisms and molybdoenzyme families is shown in Figure 4. Although nitrogenase does not use Moco as a cofactor (it uses the Fe−Mo cofactor), more than 97% of nitrogenase-containing organisms possess Moco biosynthesis pathways and Moco-dependent enzymes.[41]

Comparative genomics of Mo utilization showed several unique features of this trait: (i) Most prokaryotes and all higher eukaryotes (such as plants and vertebrates) utilize Mo whereas many lower eukaryotes including most parasites and

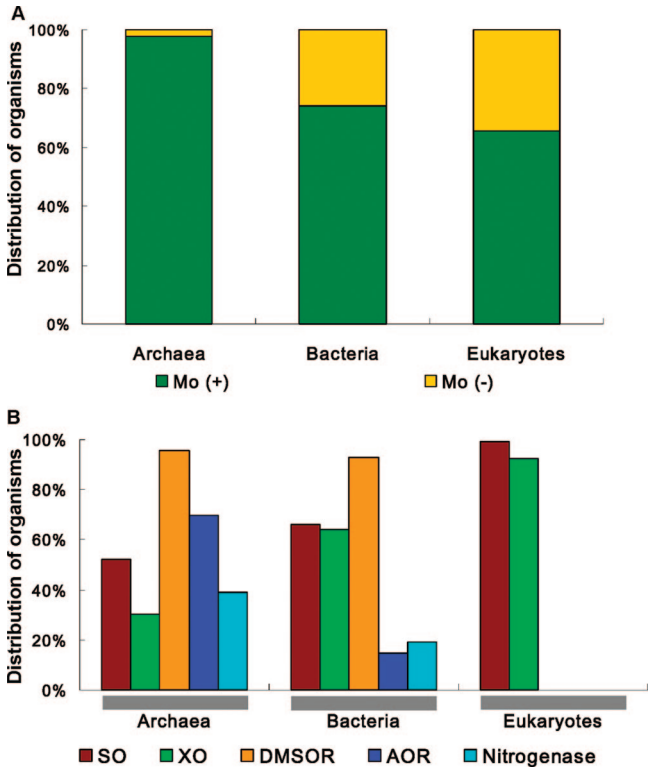

Figure 4. Distribution of Mo utilization in the three domains of life. (A) Distribution of Mo-utilizing organisms among those with completely sequenced genomes. All organisms were classified into two groups: Mo (+), i.e., containing the Mo utilization trait; Mo (−), i.e., lacking the Mo utilization trait. (B) Occurrence of molybdoenzymes in Mo-utilizing organisms. SO, sulfite oxidase; XO, xanthine oxidase; DMSOR, dimethylsulfoxide reductase, AOR, aldehyde:ferredoxin oxidoreductase.

fungi lost the ability to use this metal. (ii) In prokaryotes, the ModABC transport system is the most frequently used Mo transporter, which is present in more than 90% Mo-utilizing bacteria; the occurrence of the other two systems (WtpABC and TupABC) is much more restricted. In contrast, WtpABC is the most common transporter in archaea, whereas ModABC systems show a restricted use in these organisms. In eukaryotes, MOT1 is the only characterized Mo transporter. However, it was found that most Mo-utilizing organisms (including all animals) lack this transporter family, suggesting the presence of currently unknown Mo transport systems in these organisms. (iii) Eukaryotes appear to have fewer molybdoenzyme families than prokaryotes: only two molybdoenzyme families, SO and XO, were detected. DMSOR and SO families were the most widespread molybdoenzymes in prokaryotes and eukaryotes, respectively. (iv) ModE-related regulation of Mo uptake occurred in less than 30% of Moco-utilizing organisms, suggesting the presence of novel or unspecific regulatory pathways for molybdate uptake in these organisms. In addition, the two secondary transporters, TupABC and WtpABC, may also be regulated by a ModE-type system in some organisms. (v) A link between Mo and Se utilization in prokaryotes was identified wherein the Sec utilization trait was largely a subset of the Mo trait, suggesting that Sec utilization is heavily dependent on Mo utilization. (vi) Host-associated organisms (mostly intracellular symbionts and parasites) tend to reduce Mo utilization, perhaps due to limited bioavailability of this trace element or availability of corresponding pathways of the host.[41]

Table 2. Ni- and Co(B_{12})-Dependent Enzymes

Ni-dependent proteins	Co(B_{12})-dependent proteins
urease	Adenosylcobalamin-dependent
Ni−Fe hydrogenase	isomerase
carbon monoxide	methylmalonyl-CoA mutase
dehydrogenase	isobutyryl-CoA mutase
acetyl−coenzyme A	ethylmalonyl-CoA mutase
decarbonylase/synthase	glutamate mutase
superoxide dismutase SodN	methyleneglutarate mutase
methyl-coenzyme M reductase	D-lysine 5,6-aminomutase
(uses F_{430} as a cofactor)	diol dehydratase
	glycerol dehydratase
	ethanolamine ammonia lyase
	B_{12}-dependent ribonucleotide reductase
	Methylcobalamin-dependent
	methyltransferase
	methionine synthase (MetH)
	other methyltransferases Mta, Mtm, Mtb, Mtt, Mts, and Mtv
	B_{12}-dependent reductive dehalogenase CprA

A general evolutionary model of Mo utilization in the three domains of life emerges on the basis of the above findings. First, although Moco is essential for most organisms due to the common role of molybdoenzymes in catalysis of important redox reactions in the global carbon, nitrogen, and sulfur cycles, some organisms or even complete clades/lineages evolved alternative mechanisms and lost both the Moco biosynthesis pathway and Moco-containing enzymes. Second, out of the four molybdoenzyme families which include more than 50 subfamilies, only SO and XO (including NR, SO, XDH, and AO subfamilies) span all three domains of life. This observation suggests that SO and XO families may have evolved in the last universal common ancestor. Third, the loss of the Moco utilization trait happened independently in the early ancestors of different clades, whereas horizontal gene transfer (HGT) is unlikely to play a significant role for acquisition of Moco utilization. Fourth, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and all other sequenced yeasts lost the ability to use Mo. The absence of Mo-dependent NR yeast suggests either that Mo-dependent reduction of nitrate to nitrite is unnecessary for these organisms or that alternative Mo-independent mechanisms have evolved.

4. Nickel and Cobalt

Ni and Co are used in several enzymes involved in diverse biological processes. Ni is a component of metalloenzymes involved in energy and nitrogen metabolism,[102,103] whereas Co is mainly used in the form of vitamin B_{12}, a cofactor involved in methyl group transfer and rearrangement reactions, but is also found in some noncorrin cobalt-dependent enzymes such as [Co] nitrile hydratase.[104] Table 2 shows the list of known Ni- and B_{12}-dependent enzymes.

4.1. Ni and Co Uptake

Synthesis of Ni and Co enzymes is dependent on high-affinity uptake of these metal ions from the environment. It is interesting that both transition metals use similar transport systems. Therefore, identification of substrate preference of members of each transporter family is important for comparative genomics analyses of Ni and Co utilization. A

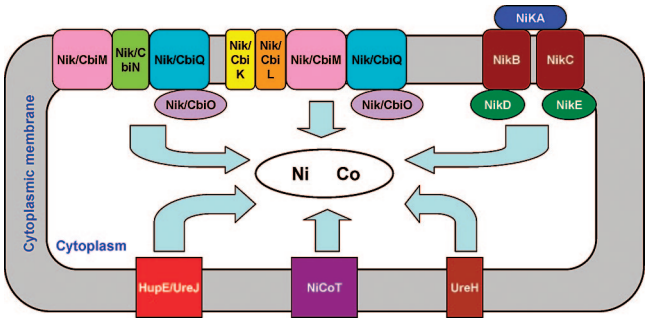

Figure 5. Schematic representation of Ni/Co transport systems. The Ni/Co transport systems include NikABCDE, Nik/CbiMNQO, Nik/CbiKMLQO, NiCoT, HupE/UreJ, and UreH.

schematic representation of known Ni/Co transport systems is shown in Figure 5.

In bacteria, Ni and Co uptake is mediated by ABC systems and several additional transporters.[105,106] The best-characterized ABC-type transporter for Ni is that found in *E. coli*. This multicomponent system consists of a periplasmic Ni-binding protein (NikA), two integral membrane proteins (NikB and NikC), and two ABC proteins (NikD and NikE).[107] In addition, NikA binds divalent Co, Cu, and Fe with at least 10-fold lower affinity.[108] A recent study showed that NikA could bind heme in *E. coli*, suggesting an additional transport function independent of Ni uptake.[109] The NikABCDE system belongs to a large family of ABC transporters, designated the nickel/peptide/opine transporter family (PepT). To date, residues involved in Ni binding have not been well characterized and conflicting results have been reported by various research groups. Cherrier et al. suggested that *E. coli* NikA binds Ni chelated by a small organic molecule, probably butane-1,2,4-tricarboxylate (BTC), and that some residues, including Tyr402, Arg137, Arg97, and His416, form a binding site that is involved in the BTC-Ni-NikA interaction.[110] On the other hand, Addy and co-workers showed that Ni appears to bind *E. coli* NikA without chelators and is coordinated by two histidine residues (His56 and His442, although not conserved in other NikA proteins) at a position distant from the previously characterized binding site.[111] Anyway, the presence of the majority of these residues could be used to help predict NikA proteins from other Ni-unrelated homologs. Besides, distantly related Ni ABC transporters were characterized in *Yersinia* species (designated YntABCDE), revealing complex diversity of Ni ABC-type transporters in prokaryotes.[112]

An additional ABC-like system, encoded by the *cbi/nikMNQO* operon, was recently found to be frequently located next to either B_{12} biosynthesis or urease (a major Ni-dependent enzyme) genes in bacterial genomes and was shown to mediate Ni and/or Co uptake.[113−116] Comparison of operon structures of *cbi/nikMNQO* and occurrences of each component revealed that M, Q, and O gene products are universal components. In contrast, the transmembrane proteins CbiN (Co uptake) and NikN (Ni uptake) are different components with no homology. However, they were predicted to have the same topology with two membrane-spanning segments.[39] Besides, two additional components, NikK and NikL, were proposed to be involved in Ni uptake in the absence of NikN to form an alternative NikKMLQO system.[39]

Three types of secondary Ni/Co transporters were also reported: (i) NiCoT (also designated as HoxN, HupN, NicT, or NhlF in various organisms);[17,117,118] (ii) UreH/SodT;[106] and (iii) HupE/UreJ.[106,119] NiCoTs, a family of prokaryotic and fungal membrane proteins with an eight-transmembrane-segment structure, are widespread among bacteria and found in several thermoacidophilic archaea and several species of fungi.[106,118,120] Subtypes of various NiCoTs have different ion preferences ranging from strict selectivity for Ni to unbiased transport of both ions to strong preference for Co. The preference for a particular metal often correlates with the genomic location of NiCoT genes, which are adjacent to genes for Ni or B_{12} biosynthesis enzymes.[17,114,117,118] The other two families (UreH/SodT and HupE/UreJ) are putative secondary transporters, and members of these families have been shown to mediate Ni transport.[119,121] Close homologs of UreH/SodT also occur in plants, including *A. thaliana* and many other higher plants.[106] Recently, several new types of candidate Co transporters were predicted, including CbtAB, CbtC, CbtD, CbtE, CbtF, CbtG, and CbtX, which show a limited distribution.[83,114] In addition, multifunctional or unspecific transporters might be involved in Ni/Co uptake. For example, CorA proteins are generally associated with the transport of magnesium ions, but some members of the CorA family may also transport Co and Ni.[122]

In eukaryotes, a subfamily of cation-efflux family members (designated TgMTP1) was reported to account for the enhanced ability of Ni hyperaccumulation in plants.[123,124] Although no high-affinity Co uptake system has been reported in eukaryotes, some suppressors of Co toxicity, such as COT1 and GRR1 in *S. cerevisiae*, were characterized, which are involved in decreasing the cytoplasmic concentration of metal ions (including Co and Zn). They were proposed to play an important role in metal homeostasis.[125]

In *E. coli*, a Ni repressor gene, *nikR*, is positioned immediately next to its target, the *nikABCDE* operon.[126] NikR-dependent regulation was also predicted for other Ni transporters, such as NikMNQO and Ni-specific NiCoT, as well as Ni-dependent enzymes such as Ni−Fe hydrogenase.[39,127] These NikRs regulate the transcription of specific genes in response to Ni ion concentrations, utilizing a combination of allostery and coordination geometry. As such, Ni regulation bears striking similarity to the regulation of other metal ions (e.g., Zn and Fe) in prokaryotes.[128−130] The presence of a NikR-binding site that contains an inverted repeat and is always located upstream of Ni-associated proteins could help identify NikR-related regulation.[39] The divergence in regulation between species implies that bacteria have evolved to tailor their Ni regulatory pathways to meet specific needs.

4.2. Ni-Dependent Proteins

Humans are not known to utilize Ni, and Ni is generally viewed as a toxic/carcinogenic metal for mammals. The requirement for Ni in fungi and microorganisms, coupled with the distinct uses of Ni in various species, makes Ni-containing enzymes an attractive target for the development of antimicrobial drugs. The characterization of Ni in the active sites of several enzymes has created an active field exploring the biochemistry of this metal.[28,131−135] In prokaryotes, the major strictly Ni-dependent enzymes include urease, Ni−Fe hydrogenase, carbon monoxide dehydrogenase (CODH), acetyl-coenzyme A decarbonylase/synthase (CODH/ACS), methyl-coenzyme M reductase (MCR), and superoxide dismutase SodN. In addition, some Ni-binding proteins appear to bind other metals in some organisms. For example, glyoxalase I (GlxI) binds Ni in *E. coli*, *Pseudomonas aeruginosa*, and human parasite *Leishmania* (e.g., *L. major*)

and *Trypanosoma* (e.g., *T. cruzi*) species, but Zn in *P. putida*, human, and yeast.[136−139] Thus, such proteins could not be used for comparative genomics of Ni utilization because of the uncertainty of the metals they bind. In eukaryotes, urease is the only characterized strictly Ni-dependent enzyme.[140,141] However, additional Ni-containing compounds or proteins may be present in some organisms, including mammals.[142,143]

Urease is the first characterized Ni-containing protein that has been found in bacteria, fungi, and plants. It catalyzes the hydrolysis of urea to carbon dioxide and ammonia.[144] In plants, urease is a hexamer of identical chains. In bacteria, it consists of either two or three different subunits (α, β, and γ).[145] Despite different quaternary structures and enzyme sizes in plants and bacteria, the protein sequence contains conserved regions of greater than 50% identity. The Ni active site appears to be particularly conserved, as two Ni atoms are associated with each active site of the respective enzymes based on the X-ray structures.[146,147]

Another Ni-dependent protein, hydrogenase, catalyzes reversible reaction of H−H bond cleavage and its formation.[148] Based on the metal content and subunit composition of the enzymes, three classes of hydrogenases have been identified: (i) Fe−Fe hydrogenase;[149,150] (ii) Ni−Fe hydrogenase (some organisms contain Ni−Fe−Se hydrogenase);[135,151−153] and (iii) hydrogenases that use neither Fe nor Ni.[154,155] The Ni−Fe hydrogenases are primarily utilized for hydrogen oxidation.[152] Structures of several Ni−Fe hydrogenases have been identified. One class is composed of two subunits which are structurally conserved. The large subunit contains the Ni active site, and the small subunit that contains an Fe−S cluster appears to be used in electron transfer from the large subunit.[156] Other Ni−Fe hydrogenases are tetramers and are integral membrane proteins.[157] Based on structural and mutation analyses, two motifs have been suggested to be involved in the ligation of the Ni ion: the N-terminal RxCGxC and the C-terminal DPCxxC.[28,156,158] Similar motifs have been found in the Ni−Fe−Se hydrogenases where the first cysteine (Cys) in the DPCxxC motif is replaced with Sec (DPUxxC, U represents Sec).[159,160]

CODHs are the biological catalysts for reversible oxidation of CO to CO_2, with water as the source of oxygen.[161] In recent years, members of the CODH family have been purified and characterized from archaea and bacteria.[162−164] The active site of CODH is called cluster C, which is a complex Ni-, Fe-, and S-containing metal center.[165] A recent report of the crystal structure of CODH from *Carboxydothermus hydrogenoformans* in three states demonstrated the mechanism of CO oxidation and CO_2 reduction at the Ni−Fe site of cluster C.[166]

CODHs from acetogenic bacteria (anaerobes that can grow autotrophically on the greenhouse gas CO_2) and methanogenic archaea are bifunctional enzymes that perform both the reversible CO-oxidation reaction and the synthesis or degradation of acetyl-coenzyme A (CoA)[167] and are therefore designated CODH/ACS. Both catalytic sites for the individual reactions require Ni for catalysis and are located at different sites in the protein.[168−170]

MCR is an enzyme that catalyzes the final step in the biological synthesis of methane in methanogenic archaea.[171] In contrast to other Ni-dependent proteins, this enzyme contains Ni in a tetrapyrrolic structure known as coenzyme F_{430}, which is found exclusively in methanogens.[172] Recently, it has been reported that MCR homologs that bind a modified F_{430} in some uncultured methanotrophic archaea are involved in anaerobic oxidation of methane in marine sediments.[173]

Superoxide dismutases (SOD) are important antioxidant enzymes that guard against superoxide toxicity. Various SODs have been characterized that employ Fe/Mn, Cu−Zn, or Ni cofactors to carry out the disproportionation of superoxide. The Ni-containing SOD is a product of the *sodN* gene, which encodes a protein with an N-terminal extension that is removed in the mature enzyme.[174,175] SodN is a small protein with no sequence homology to other SODs. It is reported to function as a tetramer[174] but to remain monomeric without Ni.[176] The crystal structure of the active Ni-bound enzyme from *Streptomyces coelicolor* identified a novel SOD fold, assembly, and the Ni active site. A nine-residue structural motif (His-Cys-X-X-Pro-Cys-Gly-X-Tyr) provides almost all interactions critical for metal binding and catalysis, and thus may be diagnostic of other SodNs.[177]

4.3. Vitamin B_{12} Uptake and Biosynthesis

Vitamin B_{12} uptake is essential for B_{12}-utilizing organisms that lack the ability to synthesize the coenzyme *de novo*. To date, the only known transport system for B_{12} in prokaryotes is the BtuFCD system, which includes a periplasmic-binding protein BtuF and two ABC transport subunits BtuC and BtuD.[178,179] In Gram-negative bacteria, a TonB-dependent outer membrane receptor BtuB is also involved in B_{12} uptake and forms a complex with BtuFCD.[180] The BtuFCD system belongs to a large superfamily involved in the uptake of Fe, siderophores, and heme.[181] In mammals, a complex system has evolved for internalization of this vitamin from the diet. Three binding proteins (haptocorrin, intrinsic factor, and transcobalamin) and several specific cell surface receptors are involved in the process of intestinal absorption, plasma transport and cellular uptake.[182,183] However, the mechanism of B_{12} uptake in other eukaryotes, such as algae and nematodes, is unclear.

In microorganisms that synthesize vitamin B_{12}, it is produced via two alternative routes: oxygen-dependent (aerobic, or "late cobalt insertion") and oxygen-independent (anaerobic, or "early cobalt insertion") pathways that differ mainly in the early stages.[184,185] A comparison of the genes required for cobalamin biosynthesis between the two pathways is shown in Figure 6. The aerobic pathway incorporates molecular oxygen into the macrocycle as a prerequisite to ring contraction. The intermediates of the aerobic route from uroporphyrinogen III (uro'gen III) to adenosylcobalamin and more than 20 genes involved in these processes (named *cobA-cobW*) have been identified in *P. denitrificans*.[186] The anaerobic route, which was partially resolved in some organisms, such as *Salmonella typhimurium*, *Bacillus megaterium*, and *Propionibacterium shermanii*,[113,187−190] takes advantage of a chelated cobalt ion, in the absence of oxygen, to set the stage for ring contraction. It has been suggested that the anaerobic and aerobic pathways contain several pathway-specific enzymes.[114] For example, CbiD, CbiG, and CbiK appear to be specific to the anaerobic route of *S. typhimurium*, whereas CobE, CobF, CobG, CobN, CobS, CobT, and CobW are unique to the aerobic pathway of *P. denitrificans*. Recently, an adenosyltransferase that catalyzes the final step in the assimilation of vitamin B_{12} was found to directly transfer the cofactor to a B_{12}-dependent protein, methylmalonyl-CoA mutase, in *Methylobacterium extorquens*, suggesting that the strategy of using the final enzyme in an assimilation pathway for tailoring a cofactor

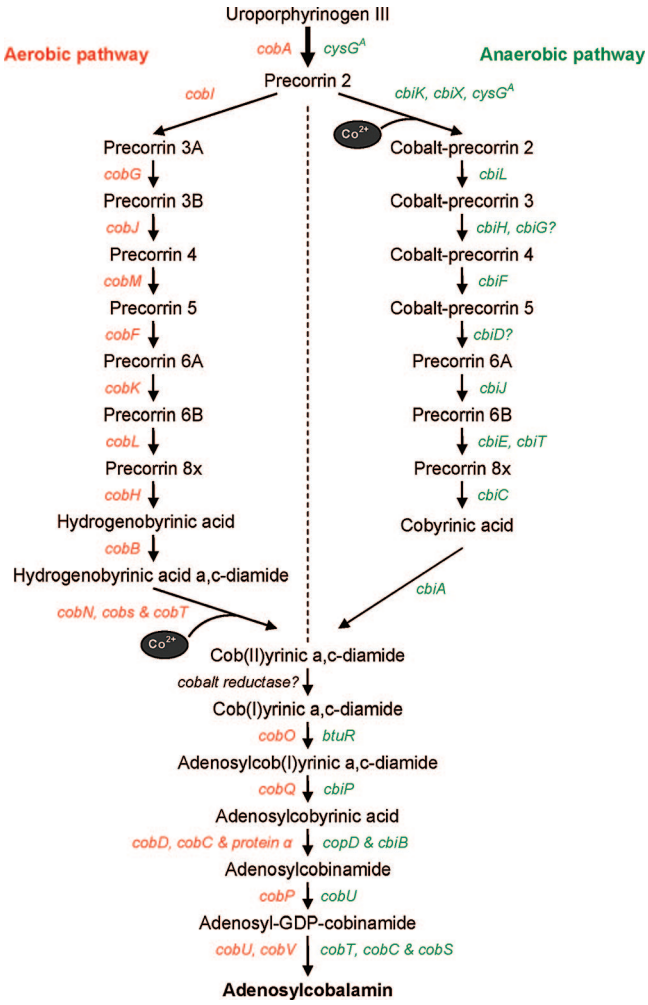

Figure 6. Biosynthetic pathways for vitamin B_{12} in bacteria. Genes involved in aerobic and anaerobic pathways are shown in red and green, respectively.

and delivering it to a dependent enzyme may also be general for cofactor trafficking.[191]

4.4. B_{12}-Dependent Enzymes

Considering that B_{12} is the major form of Co utilization and that B_{12}-containing proteins are strictly dependent on this coenzyme, coupled with the observation that non-corrin Co-containing proteins are rare and their homologs may bind other metals,[192] identification of all B_{12}-dependent enzymes is extremely important for comparative genomics of Co utilization. To date, three classes of B_{12} enzymes are known: adenosylcobalamin-dependent isomerase, methylcobalamin-dependent methyltransferase, and B_{12}-dependent reductive dehalogenase.[193,194] These classes are further divided into subclasses based on sequence similarity and reactions they catalyze (Table 2).

Adenosylcobalamin-dependent isomerases are the largest family of B_{12}-dependent enzymes and are mainly found in bacteria, where they catalyze a variety of chemically difficult 1,2-rearrangements that proceed through a mechanism involving free radical intermediates.[195,196] Subclasses of isomerases include methylmalonyl-CoA mutase (MCM),[197] isobutyryl-CoA mutase (ICM),[198] ethylmalonyl-CoA mutase (ECM or MeaA),[199] glutamate mutase (GM),[200] methyleneglutarate mutase (MGM),[201] D-lysine 5,6-aminomutase (5,6-LAM),[202] diol/glycerol dehydratase (DDH/GDH),[203] ethanolamine ammonia lyase (EAL),[204] and B_{12}-dependent ribonucleotide reductase (RNR II).[205]

MCM is the only B_{12}-dependent isomerase that is found in both microbes and mammals.[206] In many organisms, such as *S. cinnamonensis* and *P. shermanii*, it consists of two subunits, MutA and MutB, which show high sequence similarities to MCMs from other bacteria and mammals.[207,208] MCM catalyzes the isomerization of methylmalonyl-CoA to succinyl-CoA in the pathway that converts catabolites of odd-chain fatty acids, branched-chain amino acids, and cholesterol to a key intermediary metabolite.[209] In some bacteria, such as *P. shermanii*, the mutase is important in the reverse metabolic direction, linking production of propionate to succinate.[210] The crystal structure of MCM from *P. shermanii* revealed the coordination of cobalt in coenzyme B_{12} by the histidine in the conserved DXHXXG motif within the C-terminal cobalamin-binding domain.[211]

ICM and MCM are homologous proteins with different functions. ICM catalyzes the reversible rearrangement of isobutyryl-CoA to *n*-butyryl-CoA.[198,212] In *S. cinnamonensis*, this mutase was found to comprise a large subunit of 62.5 kDa (IcmA) and a small subunit of 14.3 kDa (IcmB).[198] ICM has been detected in several polyketide antibiotic-producing streptomycetes, where it appears to play a key role in valine and fatty acid catabolism as well as in the production of fatty acid-CoA thioester building blocks for polyketide antibiotic biosynthesis. The small subunit IcmB shows high sequence similarity to cobalamin-binding domains of other B_{12}-dependent enzymes (such as the cobalamin-binding domains of MutB, methyleneglutarate mutase, glutamate mutase, and B_{12}-dependent methionine synthase), including the conserved DXHXXG cobalamin-binding motif, suggesting that IcmB has taken on the role of a separate cobalamin-binding domain in ICM.

Ethylmalonyl-CoA mutase is a novel member of the family of B_{12}-dependent acyl-CoA mutases, operating in the recently discovered ethylmalonyl-CoA pathway for acetate assimilation.[199,213] It is involved in the central reaction of this novel pathway and catalyzes the transformation of ethylmalonyl-CoA to methylsuccinyl-CoA in combination with a second enzyme that was identified as promiscuous ethylmalonyl-CoA/methylmalonyl-CoA epimerase. Although this enzyme showed significant sequence similarity to MCM and ICM from the same organism,[214] sequence analysis revealed that ECM is distinct from MCM, as well as ICM, and defines a new subfamily of coenzyme B_{12}-dependent acyl-CoA mutases.[213]

Glutamate mutase is a B_{12}-dependent enzyme that catalyzes the reversible rearrangement of (2S)-glutamate to (2S,3S)-3-methylaspartate.[215] The active enzyme consists of two subunits (designated GlmE and GlmS) as an $\alpha 2\beta 2$ tetramer, whose assembly is mediated by coenzyme B_{12}. The smaller of the protein components, GlmS, is similar to the B_{12}-binding domain of MCM and has been shown to be the B_{12}-binding subunit.[216]

Coenzyme B_{12}-dependent 2-methyleneglutarate mutase from the strict anaerobe *Eubacterium barkeri* catalyzes the equilibration of 2-methyleneglutarate with (*R*)-3-methylitaconate.[201] This enzyme also contains the highly conserved motif DXHXXG(X)(41)GG, which has been shown to be critical for B_{12} binding.[217]

D-Lysine 5,6-aminomutase is an adenosylcobalamin and pyridoxal-5′-phosphate-dependent enzyme that catalyzes a 1,2 rearrangement of the terminal amino group of DL-lysine

and of L-β-lysine.[202] The X-ray structure of a substrate-free form of lysine-5,6-aminomutase from *Clostridium sticklandii* has been solved recently.[218] In the structure, a Rossmann domain covalently binds pyridoxal-5′-phosphate and positions it into the putative active site of a neighboring triosephosphate isomerase barrel domain, while simultaneously positioning the other cofactor, adenosylcobalamin, approximately 25 Å from the active site. This structure features a locking mechanism to keep the adenosylcobalamin out of the active site and prevent radical generation in the absence of substrate.

B_{12}-dependent glycerol dehydratase and diol dehydratase are highly homologous isofunctional enzymes that catalyze the elimination of water from glycerol and 1,2-propanediol (1,2-PD) to the corresponding aldehyde via a B_{12}-dependent radical mechanism.[203] The crystal structure of the substrate-free form of GDH in complex with cobalamin has been determined, whose overall fold and subunit assembly closely resemble those of DDH.[219,220] Structural analysis of the locations of conserved residues among various GDH and DDH sequences has aided in identification of residues important for substrate preference and specificity of protein−protein interactions.[220]

Ethanolamine ammonia lyase catalyzes the deamination of ethanolamine to the corresponding aldehydes.[204,221] Comparative modeling of EAL from *S. typhimurium* showed that this enzyme may have a similar TIM-barrel fold as DDH and GDH.[222,223]

Ribonucleotide reductase catalyzes the conversion of ribonucleotides to 2′-deoxyribonucleotides in all organisms, which is fundamentally important for DNA replication and repair.[224] To date, three main classes of ribonucleotide reductases have been discovered that depend on different metal cofactors for the catalytic activity: class I RNRs contain a diiron−oxygen cluster, class II contain vitamin B_{12}, and class III use an FeS cluster coupled to *S*-adenosylmethionine.[225−227] The class II enzymes are found in bacteria that can live under both aerobic and anaerobic conditions, and also in some of their phages. They utilize an adenosylcobalamin cofactor that interacts directly with an active Cys residue to form the reactive Cys radical needed for ribonucleotide reduction.

The B_{12}-dependent methyltransferases play important roles in amino acid metabolism in a variety of organisms, including mammals, as well as in carbon metabolism and CO_2 fixation in anaerobic microbes. There are two methyltransferase classes: one subclass binds simple substrates such as methanol (MtaB), methylated amines (MttB, MtbB, MtmB), methylated thiols (MtsB), methoxylated aromatics (MtvB), and methylated heavy metals, while the other, such as methionine synthase, catalyzes methyl transfer from methyltetrahydrofolate (CH_3-H_4 folate) and the methanogenic analogue methyltetrahydromethanopterin (CH_3-H_4MPT).[193,228−231]

Methionine synthase (MetH) is the most extensively studied B_{12}-dependent methyltransferase, which catalyzes the transfer of a methyl group from N_5-methyltetrahydrofolate to homocysteine, producing tetrahydrofolate and methionine.[228] This enzyme is widespread in all three domains of life. It is a modular enzyme containing separate binding domains for homocysteine, CH_3-H_4 folate, B_{12}, and adenosylmethionine (AdoMet).[229] The independently expressed modules of methionine synthase retain most of the functional properties of the native protein.[229] In addition, the B_{12} domain in its different oxidation states may interact with each of the other three domains.[193] The crystal structure of a B_{12}-containing fragment of MetH from *E. coli*, which was the first crystal structure of a protein-bound B_{12}, revealed that the histidine residue (His759) in the DXHXXG motif is the cobalt ligand and is part of a catalytic quartet, Co-His759-Asp757-Ser810, that modulates the reactivity of the B_{12} prosthetic group in MetH.[232]

Other B_{12}-dependent methyltransferases are designated as Mt*x*, where *x* denotes the methyl donor (e.g., a, methanol; v, vanillate; m, methylamine; b, dimethylamine; t, trimethylamine; and s, dimethylsulfide). These methyltransferases consist of three components (Mt_A, Mt_B, and Mt_C), which are required for the methyl transfer reaction.[193] Each component is found on a different polypeptide or domain. Mt_A methylates coenzyme M (CoM, or named mercaptoethanesulfonate), Mt_B methylates the corrinoid protein, and Mt_C is the corrinoid protein which contains B_{12}.[193,230] These methyltransferases are important in energy metabolism and in cell carbon synthesis in anaerobic microbes such as methanogenic archaea and acetogenic bacteria.[233,234] In addition, methyltetrahydromethanopterin:CoM methyltransferase (Mtr), which contains eight subunits (MtrA-H), was reported to utilize a histidine as the axial ligand to the cobalamin in MtrA.[235]

B_{12}-dependent reductive dehalogenases (CprA) play an important role in the detoxification of aromatic and aliphatic chlorinated organics in anaerobic microbes.[236,237] Most of these B_{12}-dependent reductive dehalogenases also contain Fe−S clusters. The role of B_{12} in reductive dehalogenases appears to be significantly different from those of the B_{12}-dependent isomerases and methyltransferases.[193] However, many fundamental questions regarding the reaction mechanism of dehalogenases still remain.

Only three B_{12}-dependent enzymes, MetH, MCM, and RNR II, have been reported in eukaryotes.[206,238−240] Thus, it appears that Co utilization is quite restricted in this domain of life.

4.5. Noncorrin Co-binding Proteins

Several noncorrin-Co-containing enzymes have been isolated, including methionine aminopeptidase (from *S. typhimurium*), prolidase (from *P. furiosus*), nitrile hydratase (from *Rhodococcus rhodochrous*), glucose isomerase (from *S. albus*), methylmalonyl-CoA carboxytransferase (from *P. shermanii*), aldehyde decarbonylase (from *Botryococcus braunii*), and several other proteins (for a complete review, please see ref 192). However, all of these enzymes are not strictly Co-specific and may use other metals (such as iron, zinc, and manganese) in place of Co.[192,241] It is difficult to identify the metal specificity of these enzymes by computational analysis. Among them, only nitrile hydratase (NHase) was previously suggested to have different active site motifs for Co- and Fe-binding forms.[242]

4.6. Comparative Genomics of Ni, Co, and B_{12} Utilization

Ni and Co are less frequently used in metalloenzymes than other transition metals (e.g., Mn, Fe, Zn, and Cu); however, they are essential cofactors in several enzymes. As mentioned above, Ni is used in several metalloenzymes involved in energy and nitrogen metabolism, detoxification processes, pathogenesis, enzyme inactivation, lipid peroxidation, and damage of nucleic acids, whereas Co is primarily found in

the corrin ring of coenzyme B_{12}. In recent years, several comparative genomic studies have been carried out to investigate Ni and Co utilization traits.

One study examined Ni and Co transporters in about 200 microbial genomes and demonstrated a complex and mosaic utilization of both metals in prokaryotes.[39] For functional prediction of proteins involved in Ni or Co uptake in each analyzed organism, two approaches were used: (i) analysis of the genomic locations of genes encoding candidate Ni/Co transporters and (ii) identification of regulatory signals, including NikR-dependent regulation through the NikR-binding signal, and B_{12} riboswitches that regulate many of the candidate Co transporters in bacteria.[83,243−246] This *in silico* study showed that the Ni/Co transporter genes are often colocalized with the genes for Ni-dependent and coenzyme B_{12} biosynthesis enzymes. Different families of Ni/Co transporters, including NikABCDE, Cbi/NikMNQO, NiCoT, UreH, and HupE/UreJ, showed a mosaic distribution in analyzed organisms, and the Cbi/NikMNQO system (including the NikKMLQO system) appeared to be the most widespread group of microbial transporters for Co and Ni ions.[39]

A separate comparative genomic analysis of B_{12} metabolism and regulation provided important information regarding B_{12} utilization in prokaryotes.[114] Using approaches of gene clustering, gene regulation, and phylogenetic profiling, the B_{12} biosynthesis and regulation was described in a variety of bacterial genomes. The regulatory B_{12} riboswitch was found to be widely distributed in the regions upstream of B_{12} biosynthetic/transport genes. By searching for candidate B_{12}-regulated genes, several new types of candidate Co transporters and several new proteins linked to the B_{12} biosynthesis pathway, such as chelatases and methyltransferases, were identified. In addition, the B_{12} transporters, BtuFCD, appeared to be widely distributed in bacteria and archaea and some of them were B_{12}-regulated. However, it has been difficult to selectively identify BtuFCDs among other highly similar transport systems (such as iron/heme or siderophore transporters) in sequenced organisms. It is also interesting that B_{12}-independent MetH and RNR II were also predicted to be regulated by B_{12} elements in some bacteria.[114]

Very recently, we carried out a similar but more extensive comparative genomic analysis of Ni and Co utilization in more than 700 organisms in all three domains of life.[247] Only strictly Ni-dependent metalloproteins and B_{12}-binding enzymes were used for comparative genomics of Ni and Co, respectively. Occurrence of the Ni/Co-utilization trait and Ni- or B_{12}-dependent proteins is shown in Figure 7. The distribution and dynamics of the use of Ni and Co (in the form of B_{12}) were analyzed at the level of transport systems and metalloproteomes. These analyses revealed that both metals are widely used in bacteria and archaea. Ni- or Co-dependent metalloenzymes showed a mosaic distribution. Ni-dependent urease and Ni−Fe hydrogenase, and B_{12}-dependent MetH, RNR II, and MCM families, were the most widespread users. Interestingly, deltaproteobacteria and methanosarcina generally have larger Ni- and Co-dependent metalloproteomes than other organisms. We also identified several bacteria containing Co-binding NHases based on the presence of a Co-binding motif (CTLCSCY[242]). All of them are B_{12}-utilizing organisms. Besides, Fe-containing NHases (containing a CSLCSCT sequence motif[242]) were predicted in a few organisms, and they might be newly evolved from Co-binding NHases based on phylogenetic analysis. Investigation of Ni and Co utilization in eukaryotes provided a first glimpse of the evolutionary dynamics of Ni- and Co-dependent pathways in this domain of life. The utilization of these two metals is much more restricted in eukaryotes, with regard to both the organisms that use Ni/Co and the number of Ni transporters and Ni/B_{12}-dependent protein families. Very few of these organisms utilize both metals (Figure 7). The Ni-utilizing eukaryotes are mostly fungi (except saccharomycotina) and plants, whereas most B_{12}-utilizing organisms are animals. The NiCoT transporter family is the most widespread eukaryotic Ni transporter, and urease and MetH are the most common eukaryotic Ni- and B_{12}-dependent enzymes, respectively. Finally, analysis of Ni and Co utilization based on different habitats, environments, and other factors revealed that host-associated organisms (particularly obligate intracellular parasites and endosymbionts) have a tendency for reduced Ni/Co utilization.

5. Copper

Cu is an essential trace metal in most organisms from bacteria to humans. It serves as an important cofactor for a variety of proteins and enzymes that carry out fundamental biological functions.[248] However, Cu ion is also highly toxic in the free form because of its ability to produce radicals by cycling between oxidized Cu(I) and reduced Cu(II) species.[249] Therefore, it is important for Cu-utilizing organisms to obtain sufficient levels of Cu ion to meet their needs while limiting intracellular Cu concentration, especially in the free form, to attain homeostatic balance and avoid toxicity.[250,251]

5.1. Cu Transport and Homeostasis

Cellular Cu transport processes are required by organisms for correct utilization of this element in biochemical processes and to limit the toxicity of excess Cu. Cu import into cells mainly requires the coordinate function of proteins with metal-binding domains. On the other hand, detoxification mechanisms found across species include the binding of Cu to specific proteins (e.g., metallothioneins) and its transfer into cell compartments such as periplasmic space, mitochondria and lysosomes.[252−254] Thus, understanding of Cu transport and homeostasis is important for comparative genomics of Cu utilization.

The mechanisms involved in Cu transport and homeostasis in prokaryotes are only partially understood. To date, Cu trafficking in bacteria is best described in *E. coli* and in *Enterococcus hirae*.[255,256] The most relevant Cu homeostatic systems in *E. coli* are shown in Figure 8A. Several Cu-related transport and resistant proteins and systems have been characterized in different organisms, including a P-type ATPase CopA/PacS, CusCFBA, CutC, PcoABCD, and PcoE.[252,255,256] In addition, CueO, a multicopper oxidase (MCO), protects the periplasm from Cu-induced damage.[257]

In *E. coli*, the Cu(I)-translocating P-type ATPase CopA is the central component of Cu homeostasis and is required for removing excess Cu(I) from the cytoplasm.[258] CopA proteins belong to a superfamily that is involved in transport of transition or heavy metal ions across biological membranes.[259−261] These ATPases can be further divided into Cu(I)/Ag(I)-translocating ATPases (e.g., CopA) and Zn(II)/Cd(II)/Pb(II)-translocating ATPases.[261−263] *E. coli* CopA possesses two CXXC motifs in the N-terminal domain. However, they are not required for function and do not confer metal specificity, suggesting that they might have a regulatory

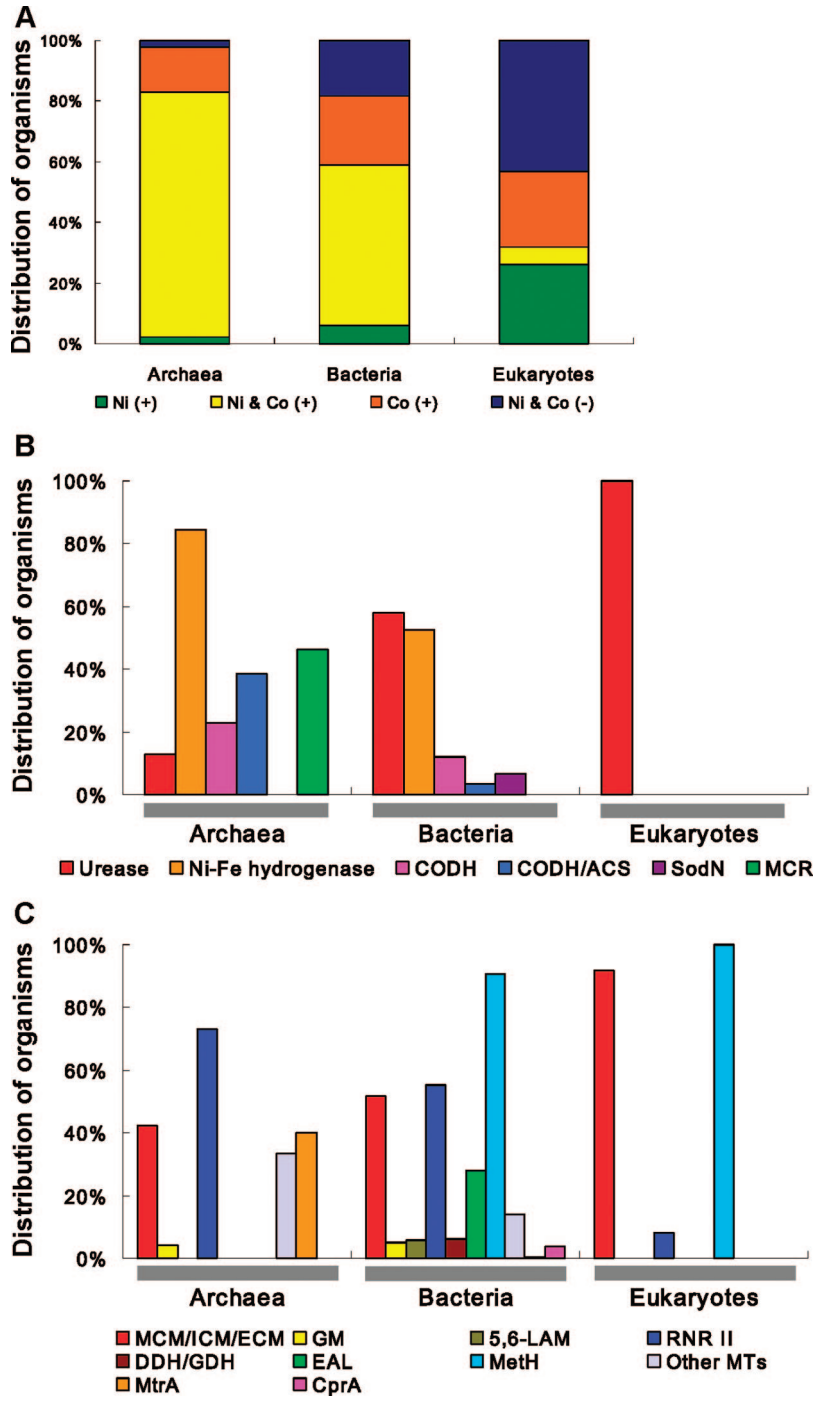

Figure 7. Occurrence of Ni and Co utilization in the three domains of life. (A) Distribution of Ni-/Co-utilizing organisms among those with completely sequenced genomes. All organisms were classified into four groups: Ni (+), i.e., containing the Ni utilization trait only; Ni & Co (+), i.e., containing Ni and Co utilization traits; Co (+), i.e., containing the Co utilization trait only; Ni & Co (−), i.e., containing neither Ni nor Co utilization traits. (B) Distribution of organisms containing different Ni-dependent proteins in Ni-utilizing organisms. CODH, carbon monoxide dehydrogenase. (C) Occurrence of B_{12}-dependent proteins in Co-utilizing organisms. CODH/ACS, acetyl-coenzyme A decarbonylase/synthase; SodN, Ni-containing superoxide dismutase; MCR, methyl-coenzyme M reductase; MCM, methylmalonyl-CoA mutase; ICM, isobutyryl-CoA mutase; ECM, ethylmalonyl-CoA mutase. The latter three subfamilies are quite similar and are combined into one group. GM, glutamate mutase; 5,6-LAM, D-lysine 5,6-aminomutase; RNR II, B_{12}-dependent ribonucleotide reductase; DDH/GDH, diol/glycerol dehydratase; EAL, ethanolamine ammonia lyase; MetH, methionine synthase; Other MTs, various B_{12}-dependent methyltransferases such as Mta, Mtm, Mtb, Mtt, Mts, and Mtv systems; MtrA, methyltetrahydromethanopterin:CoM methyltransferase subunit A; CprA, reductive dehalogenase.

effect.[264,265] In contrast, two Cys residues in a Cys-Pro-Cys (CPC) motif located in the middle of CopA are required for CopA function.[265] PacS, a CopA homolog identified in a cyanobacterium *Synechococcus sp. PCC7942*, is mainly located in the thylakoid membrane, in which the photosynthetic reactions take place.[266] Therefore, PacS in cyanobacteria may be involved in Cu homeostasis crucial to the photosynthetic thylakoid function (thylakoid import).[267] CtaA, another CopA homolog identified in cyanobacteria (*Synechococcus sp. PCC7942* and *Synechocystis PCC 6803*), was suggested to be involved in Cu import from the periplasm.[267−269] It has been shown that both CtaA and PacS are required for the use of Cu in plastocyanin and cytochrome *c* oxidase in the thylakoid, consistent with inward Cu

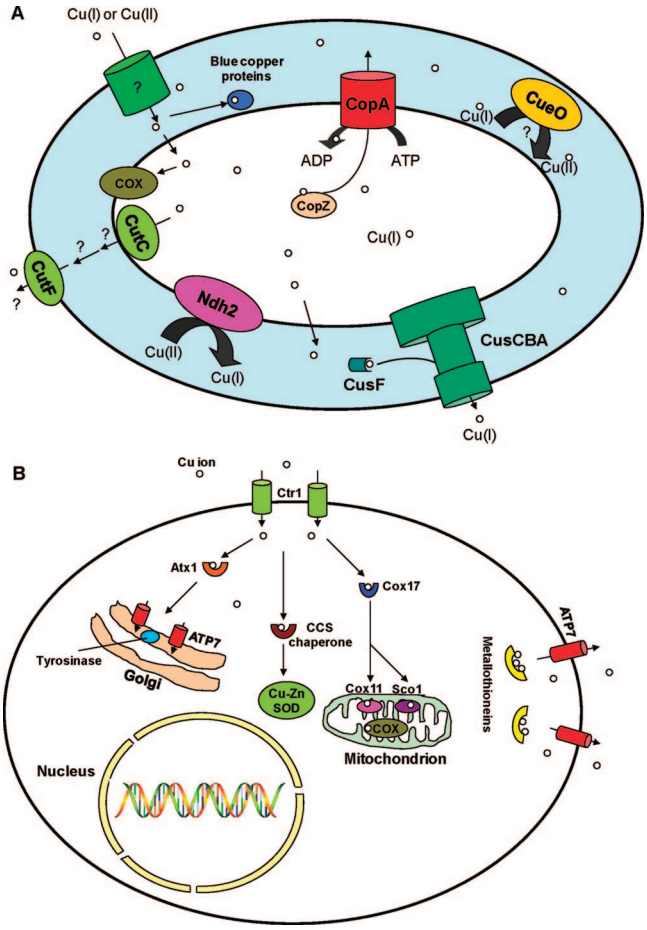

Figure 8. Schematic view of copper homeostasis. (A) Cu homeostasis in *E. coli*. CopA, the Cu(I)-translocating P-type ATPase; CusCFBA, the four-component Cu efflux system; Ndh 2, a cupric reductase; CueO, a multicopper oxidase; CutC and CutF, two proteins involved in Cu efflux and/or homeostasis; CopZ, a Cu chaperone involved in Cu export; COX, cytochrome *c* oxidase. (B) Cu homeostasis in *Drosophila melanogaster*. Atx1, CCS, and Cox17, Cu chaperones involved in various pathways; Ctr1, eukaryotic Cu importer; ATP7, eukaryotic Cu exporter (also involved in Cu transport to Golgi); COX11 and Sco1, two proteins involved in cytochrome *c* oxidase assembly; Cu−Zn SOD, Cu−Zn superoxide dismutase.

transport by them.[268] Deletion of *ctaA* impairs cellular Cu accumulation whereas deletion of *pacS* confers Cu sensitivity.[268] In *E. hirae*, two CopA homologs (named CopA and CopB) were identified.[256] The *E. hirae* CopA may be involved in Cu uptake, which is similar to the function of CtaA, whereas the *E. hirae* CopB functions as an ATPase for the export of Cu ion (like *E. coli* CopA), thereby serving in Cu resistance.[270,271] In *E. hirae* and many other organisms such as *B. subtilis* and *Staphylococcus aureus*, a Cu chaperone, CopZ, functions as part of a complex cellular machinery for Cu(I) trafficking and detoxification, in which it interacts with and delivers the metal to the Cu(I)-exporter CopA.[272] CopZ homologs are also found in eukaryotes, named Atx1, which bind Cu(I) and interact directly with the Golgi P-type ATPase Cu transporter (ATP7 family).[273] Very recently, a new Cu(I)-binding metallothionein, MymT, was identified in several pathogenic mycobacteria, which may also serve as a chaperone involved in CopA-associated Cu(I) detoxification.[274] Additional Cu-regulating proteins can be anticipated, considering that Cu-dependent enzymes require sufficient Cu, yet excessive Cu risks damage to DNA and other molecules.

Another Cu efflux system, which is exclusively present in Gram-negative bacteria, is the CusCFBA system. In *E. coli*, the CusCFBA system is encoded by one operon.[255,275] The four-part Cus complex consists of the inner membrane pump CusA, the periplasmic protein CusB, and the outer membrane protein CusC forming a channel bridging the periplasmic space. CusF is a periplasmic Cu-binding chaperone that transports Cu to the CusCBA efflux complex, thus facilitating Cu detoxification in the periplasm.[276] It has been shown that CusA and CusB are essential for Cu resistance, and CusC and CusF are required for full resistance.[277]

In *E. coli*, the *cutC* gene encodes a cytoplasmic protein CutC, which is involved in Cu tolerance. Previous studies have implicated CutC in Cu efflux, suggesting a role of CutC in intracellular trafficking of Cu(I).[278] The crystal structure of CutC from *Shigella flexneri* suggested that it is a new family of TIM barrel proteins,[279] providing a sound basis for the in-depth study of its structure−function relationship. Another outer membrane protein with potential Cu-binding sites, named CutF (or NlpE), is also required for Cu tolerance.[278] Thus, it was suggested that CutC is a cytosolic component of an efflux pathway for Cu, whereas CutF may be involved in both Cu efflux and the delivery of Cu to Cu-dependent enzymes.[278] CutC homologs have also been characterized in eukaryotes including human and *C. elegans*.,[280,281] implying that this transporter family may play a role in the protection from excess Cu in both prokaryotes and eukaryotes.

Some *E. coli* strains (e.g., APEC O1) possess additional plasmid-encoded seven-gene clusters, named *pcoABCDRSE*, that confer Cu resistance.[255] The *pcoA* gene encodes a periplasmic MCO which is the central protein of the Pco system. It is known that PcoA could functionally substitute for CueO, another MCO involved in Cu tolerance, in *E. coli*, indicating they have a similar function.[275] PcoB (or CopB) is predicted to be an outer membrane protein and is often colocalized with *pcoA* in the same operon in many other organisms, such as *Xanthomonas axonopodis* and *P. syringae*.[255,282−284] PcoC (or CopC) and PcoD (or CopD) may only be needed for full resistance because they do not co-occur with PcoAB.[255] PcoC is a periplasmic Cu carrier involved in Cu homeostasis. This protein has been shown to bind Cu atom and has two metal-binding sites distinct for Cu(II) and Cu(I).[285,286] It has been suggested that PcoC may function with PcoD (an integral membrane protein) together in Cu uptake.[287,288] Therefore, it is possible that PcoC would deliver periplasmic Cu to PcoD for uptake into the cytoplasm, probably for loading into PcoA. PcoR and PcoS are regulators required for Cu-inducible expression of Cu resistance mediated by the PcoABCD system.[289,290] Recently, a distant homolog of PcoCD, named YcnJ, was characterized in *B. subtilis*, and it may have an important role in Cu import in this organism.[291] The *pco* cluster in *E. coli* also contains another gene, *pcoE*, which does not belong to the *pcoABCD* operon. PcoE is related to Cu accumulation in the periplasm.[255,289] Overall, PcoABCDE are involved in periplasmic Cu handling. PcoB probably prevents Cu uptake, whereas PcoC and PcoD may be involved in Cu uptake across the cytoplasmic membrane.[292] It is possible that PcoE provides initial sequestration of Cu in the periplasm before the remaining genes of the Pco system are fully induced, thus minimizing the effects of Cu stress but not contributing to maximal resistance levels.

Recently, two additional genes, *cinA* encoding a putative azurin-plastocyanin-like protein and *cinQ* encoding a putative GTP cyclohydrolase/pre-Q_0 reductase, were identified to be involved in Cu resistance in several *Pseudomonas* species.[293] Although the roles for the two proteins in Cu homeostasis are undefined, the finding that CinA homologs are most often associated with multicopper oxidases suggests a possible role in electron transfer.[293]

A schematic view of Cu homeostasis in eukaryotes is shown in Figure 8B (using *Drosophila melanogaster* as a representative organism). In eukaryotes (from fungi to mammals and likely in many lower eukaryotes), Cu is acquired by the high-affinity, membrane-associated Cu importers: Cu transporter(Ctr)-family.[294−299] Members of the Ctr family contain three transmembrane domains and possess an N-terminal Met-rich domain which is important for survival under Cu starvation. These Met residues are part of the extracellular domain and are involved in the acquisition of Cu(I) ions.[294,300,301] In addition, Ctr proteins do not require ATP for Cu import.[294,302] They are probably stimulated by extracellular K^+ and the extremely low intracellular concentration of free Cu.[302,302] Different organisms may possess multiple Ctr proteins located in different biological membranes. For example, in *S. cerevisiae*, three Ctr proteins, named yCtr1−yCtr3, are present.[299,304−306] yCtr1 and yCtr3 are located in the plasma membrane, whereas yCtr2 is localized in the vacuolar membrane and imports Cu from the vacuole to the cytoplasm upon Cu depletion.[306] Humans contain two Ctr proteins, hCtr1 and hCtr2. hCtr1 is the main cellular Cu importer.[307] It is located predominantly at the plasma membrane but may also be present in intracellular vesicular perinuclear compartments.[308] hCtr2 was localized exclusively to late endosomes and lysosomes and may be involved in Cu delivery to the cytosol of human cells.[309] *D. melanogaster* also expresses three Ctr1 genes (*ctr1A*, *ctr1B*, and *ctr1C*).[298] Ctr1A resides at the plasma membrane and is the primary *Drosophila* Cu transporter during normal growth and development. Loss of Ctr1A results in developmental arrest at early larval stages.[310] Ctr1B also localizes to the plasma membrane and is not essential for development unless flies are severely Cu-deficient or are subjected to Cu toxicity.[310,311] Little is known about Ctr1C.

Studies of two human disorders of Cu transport, Menkes disease and Wilson disease, led to the identification of an important category of ATP-dependent transporters mediating cell Cu export, the ATP7 family.[312] ATP7 is homologous to bacterial CopA proteins and conserved in many other eukaryotes.[313−315] In mammals, two isoforms are expressed: ATP7A and ATP7B.[316] ATP7A, the protein nonfunctional in Menkes disease, is expressed in the intestinal epithelium as well as most other tissues other than the liver.[317] This protein is required for transport of Cu into the trans-Golgi network for biosynthesis of several secreted cuproenzymes and for basolateral efflux of Cu in the intestine and selected other cells.[318] ATP7B is predominantly expressed in the liver and is required for Cu metalation of ceruloplasmin and biliary Cu excretion, and mutations in ATP7B lead to Wilson disease.[318−320] Both ATP7A and ATP7B undergo Cu-stimulated trafficking.[321−326] *D. melanogaster* has a sole ortholog of ATP7 (named DmATP7), which is essential for *in vivo* Cu distribution by (i) delivering Cu to cuproenzymes required for pigmentation and neuronal function and (ii) removing excess cellular Cu via facilitated efflux.[315,327−329] Yeasts also have an ortholog of ATP7 protein, Ccc2.[330] *Ccc*2 is a Cu(I) ATPase embedded in the trans-Golgi membrane, which obtains Cu from the Atx1 chaperone and transfers it to secreted proteins.[324,331,332]

5.2. Cu-Dependent Proteins

The biological roles of Cu include electron transfer, oxidation of organic substrates and metal ions, dismutation of superoxide, monooxygenation, transport of dioxygen and iron, reduction of dioxygen, nitrite, and nitrous oxide, etc. Both Cu(I) and Cu(II) are utilized in biological systems. Currently, it is not possible to identify all Cu-binding proteins in one organism using computational approaches because (i) some proteins bind Cu in certain organisms but alternative metals in others and (ii) there are additional factors, such as protein folding location and the presence of certain supporting residues or other metal ions that may influence the affinity, allostery, and inherent chemical properties of individual metals and further influence the conformation or function of proteins in individual organisms.[57,333−336] Thus, this review only focuses on strictly Cu-dependent protein families which have been used for comparative genomics of Cu utilization in recent studies. Identification of all members of Cu-dependent protein families as well as their ability to bind Cu (based on conservation of Cu ligands) is essential for comparative genomic analysis of Cu utilization and evolution, especially in prokaryotes where no widespread Cu-specific importer has been detected.

Cu in proteins can be classified into three groups based on spectroscopic and magnetic properties: type 1, type 2, and type 3. Type 1 Cu (blue Cu) shows intense absorption at ~600 nm, whereas type 2 Cu (nonblue Cu) does not. Both type 1 and type 2 Cu can be detected in the electron paramagnetic resonance (EPR) spectrum. Type 3 Cu cannot be detected in the EPR studies.[337] To date, a number of Cu-containing proteins have been characterized in all three domains of life.[337−340] Some Cu-dependent proteins (such as MCOs) contain multiple Cu centers, whereas others contain only one type of Cu (for example, blue Cu proteins and type 2 Cu-containing enzymes). A list of Cu-dependent proteins is shown in Table 3.

Blue Cu proteins are a class of cuproproteins containing a single type 1 Cu center in a relatively small (10−20 kDa) protein molecule that functions in electron transfer from a donor to an acceptor protein in the respiratory and photosynthetic chains of bacteria and plants.[340−343] The major blue Cu proteins (also called cupredoxins) include plastocyanin, azurin, pseudoazurin, amicyanin, rusticyanin, auracyanin, halocyanin, plantacyanin, umecyanin, mavicyanin, and stellacyanin.[337] Among them, the most extensive studies were carried out with plastocyanin. This protein shuttles electrons from cytochrome b6/f to photosystem I, both of which are large (200−300 kDa) and membrane-embedded complexes.[343] A large number of crystal and solution structures of plastocyanin from different species have revealed that plastocyanin has an eight-stranded flattened Greek-key β-barrel fold, containing a type 1 Cu atom coordinated by two histidines, one cysteine and one methionine.[343−347] The red Cu protein, nitrosocyanin, is a variant of the blue Cu protein.[348] The red Cu site is the only presently known blue Cu-related site with an exogenous water coordinated to the Cu.[349]

The same type I Cu is also found in the larger enzymes nitrite reductase (NiR) that catalyzes the one-electron reduction of nitrite (NO_2^-) to nitric oxide (NO), and multi-Cu

Table 3. Cu-Dependent Proreins

prokaryotes	eukaryotes
plastocyanin family (including plastocyanin, amicyanin, pseudoazurin, halocyanin, etc.)	plastocyanin family
azurin family (including azurin and auracyanin)	plantacyanin family (including plantacyanin, umecyanin, mavicyanin, stellacyanin, etc.)
rusticyanin	cytochrome *c* oxidase subunit I
nitrosocyanin	cytochrome *c* oxidase subunit II
cytochrome *c* oxidase subunit I	Cu−Zn superoxide dismutase
cytochrome *c* oxidase subunit II	copper amine oxidase
nitrous oxide reductase	peptidylglycine α-hydroxylating monooxygenase
NADH dehydrogenase 2	dopamine β-monooxygenase
Cu−Zn superoxide dismutase	multicopper oxidases (including laccase, Fet3p, hephaestin, ceruloplasmin, ascorbate oxidase, etc.)
copper amine oxidase	tyrosinase (or polyphenol oxidase)
particulate methane monooxygenase	hemocyanin
multicopper oxidases (including nitrite reductase, CueO, CotA, laccase, bilirubin oxidase, phenoxazinone synthase, etc.)	Cnx1G
tyrosinase	galactose oxidase

oxidases that function in intramolecular electron transfer to Cu active sites.[337,339,350] A number of MCOs have been found in both prokaryotes and eukaryotes, such as laccase, ascorbate oxidase, CueO, PcoA, CumA, MofA, MnxG, EpoA, CotA, dihydrogeodin oxidase (sulochrin oxidase), hephaestin, ceruloplasmin, phenoxazinone synthase, bilirubin oxidase, Fet3p, etc.[337,339,350−352] Most MCOs contain four Cu centers: a type 1 Cu and a trinuclear Cu center comprised of a type 2 Cu and a pair of type 3 Cu.[337,350] The function of a type 1 Cu in MCOs is to withdraw an electron from the substrate and transfer it to the trinuclear Cu center. These MCOs catalyze the oxidation of various small molecules and cations with the concomitant four-electron reduction of oxygen to water. Some MCOs such as mammalian ceruloplasmin and yeast Fet3p are ferroxidases, oxidizing Fe(II) to Fe(III). In contrast, laccases are MCOs that derive electrons from the oxidation of phenolic compounds. MCO mechanisms have been extensively reviewed.[353] Striking similarity exists between the Cu sites of NiR and those of the MCOs, both structurally and mechanistically, suggesting that NiR and MCOs share a common ancestor.[339]

The type 2 Cu-containing enzymes include Cu−Zn superoxide dismutase (Cu−Zn SOD), Cu amine oxidase (CuAO), peptidylglycine α-hydroxylating monooxygenase (PHM), and dopamine β-monooxygenase (DBM).[338]

Many eukaryotes and prokaryotes express Cu−Zn SOD. Most of the mechanistic and crystallographic studies have focused on the enzymes from eukaryotic sources, such as yeast, bovine, and human.[354] The active site type 2 Cu of oxidized Cu−Zn SOD is coordinated by four histidine residues (e.g., His46, His48, His63, and His120 in the human protein).

CuAO belongs to a larger class of amine oxidases that catalyze oxidative deamination of amines with concomitant reduction of oxygen to hydrogen peroxide. These ubiquitous enzymes are found in a large variety of organisms, from microbes (including bacteria and fungi) to plants and mammals.[355] In bacteria, CuAOs have well-established roles in providing carbon or nitrogen sources when primary amines are available. In plants, there is evidence for the role of CuAOs in defense responses and a variety of developmental processes.[356] In mammals, CuAOs are found in various locations, including placenta, blood, muscle, and endothelium; however, their functions are not well understood. Increased CuAO expression in humans is a marker of several diseases including cancer, diabetes, congestive heart failure, and liver cirrhosis.[357,358] Crystal structures for CuAO from several organisms such as *E. coli* and humans showed that the Cu atom is coordinated by three histidine residues and two water molecules.[359,360]

PHM and DBM catalyze the hydroxylation of their respective substrates. In both enzymes, two distinct Cu sites are used to split molecular oxygen, O_2, which then serves as the source of OH in the hydroxylation reaction.[338] They are found primarily in metazoa, and their functions *in vivo* are well established. PHM is one of two domains in peptidylglycine α-amidating monooxygenase (PAM), which is responsible for the activation of a variety of hormones by α-amidation, thereby improving hormone-receptor affinity.[361] DBM catalyzes a similar reaction to PHM; however, the hydroxylation of dopamine is at the β-carbon.[361] Comparison of the primary sequence of PHM and DBM indicated that DBM is homologous to PHM, suggesting that they likely evolve from a common ancestor.[362]

Two additional Cu-dependent proteins, cytochrome *c* oxidase (COX) and nitrous oxide reductase (N_2OR), have a binuclear Cu center, named Cu_A, which is a variant of type 1 Cu, and an iron−sulfur center and function to facilitate entry of electrons toward the reduction centers of dioxygen (heme a3-Cu_B) and nitrous oxide (Cu_Z), respectively.[363−365]

Cytochrome oxidase is a family of proteins which act as the terminal enzymes in respiratory chains. The two main subgroups of this family include cytochrome *c* oxidases and quinol oxidases.[366,367] Both classes have two catalytic subunits (I and II), and subunit I contains two heme centers: the first (heme a) acts as an electron input device to the second, and the second heme (heme a_3) is a part of a binuclear center containing Cu_B. However, there are significant differences between the two subgroups. In cytochrome *c* oxidases, subunit II contains the Cu center Cu_A with 2 Cu atoms, which is thought to be the immediate electron acceptor from cytochrome *c*, whereas quinol oxidase subunit II processes the quinol substrate and independently lost the Cu_A center.[368,369] Cytochrome *c* oxidases of mitochondria, of respiratory bacteria from the purple bacterial branch of the eubacteria, and one of the cytochrome *c* oxidases of *Rhodobacter sphaeroides*, are of the aa_3-type.[370,371] The ba_3-type cytochrome *c* oxidases have a heme b in place of heme a.[372] The third cytochrome *c* oxidase class is the cbb_3-type, which has two membrane-bound cytochrome *c* molecules in place of subunit II.[373] Quinol oxidases also have these subclasses (aa_3, ba_3, and cbb_3). In addition, a bo_3-type quinol oxidase has been detected, where a heme type-o replaces heme at the binuclear center.[374] Characterizing all these subtypes and

distinguishing Cu-dependent cytochrome *c* oxidase subunit II from Cu-independent quinol oxidase subunit II is essential for correct description of Cu utilization.

Nitrous oxide reductase is the terminal oxidoreductase of a respiratory electron transfer chain that transforms nitrous oxide to dinitrogen.[363] The enzyme carries six Cu atoms. Two are arranged in the mixed-valent binuclear Cu_A site similar to that of cytochrome *c* oxidase, and four make up the sulfide-bridged Cu cluster, named the Cu_Z catalytic center. The crystal structure of *P. nautica* N_2OR revealed that the catalytic Cu_Z center belongs to a new type of metal cluster in which the four Cu ions are liganded by seven histidine residues.[375]

Other Cu-dependent proteins include NADH dehydrogenase 2 (Ndh2), tyrosinase, hemocyanin, particulate methane monooxygenase (pMMO), Cnx1G, and galactose oxidase (GAO).

The Cu(II)-reductase Ndh2 from the *E. coli* respiratory chain is a membrane-bound cupric-reductase that diminishes the susceptibility of the respiratory chain to damaging effects caused by Cu and hydroperoxides and allows cells to survive in extreme Cu conditions. It contributes to antioxidant function and Cu homeostasis.[376]

Tyrosinases (or catechol/polyphenol oxidases) are Cu-containing enzymes which are nearly ubiquitously distributed in all domains of life. They are essential for pigmentation and are important factors in wound healing and primary immune response.[377] The active site is a type 3 Cu center, which is a binuclear center consisting of two Cu ions, each coordinated by three histidine residues. The Cu pair of tyrosinases binds one molecule of atmospheric oxygen to catalyze two different kinds of enzymatic reactions: (i) the *o*-hydroxylation of monophenols (cresolase activity) and (ii) the oxidation of *o*-diphenols to *o*-diquinones (catecholase activity). The best-known function of tyrosinases is the formation of melanins from L-tyrosine via L-dihydroxyphenylalanine (L-dopa). However, the complicated hydroxylation mechanism at the active site is still not completely understood.

Similar to tyrosinase, hemocyanin belongs to the type 3 Cu protein family.[378] Although they share a common active site, they exhibit different functions. While tyrosinases are enzymes, hemocyanins are oxygen carrier proteins.[379] They occur in the hemolymph of some species in the phyla arthropoda and mollusca. As extracellular oxygen carriers, hemocyanins are responsible for the precise oxygen delivery from the respiratory organs to tissues.

pMMO is a membrane-bound Cu-containing enzyme that oxidizes methane to methanol in methanotrophic bacteria.[380] Knowledge of how pMMO selectively oxidizes methane under ambient conditions could impact the development of new catalysts. The crystal structure of *Methylococcus capsulatus* (Bath) pMMO, for the first time, reveals the composition and location of three metal centers.[381] In the past several years, significant advances in the biochemical, spectroscopic, and mechanistic characterization of pMMO have provided insights into the coordination environments and oxidation states of these metal centers.[382]

Cnx1G is the G domain of Cnx1 that is involved in catalyzing the insertion of Mo into molybdopterin (see section 3.1 and Figure 3). The recent identification of Cu bound to the molybdopterin dithiolate sulfurs in Cnx1G structures, coupled with the observed Cu inhibition of Cnx1G activity, provides a molecular link between Mo and Cu metabolism.[383]

Galactose oxidase is a monomeric enzyme that contains a single Cu ion and an amino acid-derived cofactor.[384] The enzyme has been extensively studied by structural, spectroscopic, kinetic, and mutational approaches that have provided insights into the catalytic mechanism of this radical enzyme. One of the most intriguing features of the enzyme is the post-translational generation of an organic cofactor from active-site amino acid residues. Biogenesis of this cofactor involves the autocatalytic formation of a thioether bond between Cys228 and Tyr272, with the latter being one of the Cu ligands.[385]

5.3. Comparative Genomics of Cu Utilization

Several comparative genomic studies have been carried out to identify Cu-binding proteins in organisms.[386−388] One study presented an initial computational approach exploiting metal-binding patterns of metalloproteins in the PDB to search databases for new metalloproteins, and applied this method to Cu-binding proteins.[386] This approach was based on the analysis of the occurrence of conserved patterns of amino acids that are known to bind Cu, together with sequence similarity requirements. A set of Cu-binding patterns (CBPs) were generated for all Cu-binding proteins in the PDB and then used together with the primary sequences of corresponding metalloproteins to identify Cu-binding proteins by homology searches. However, this procedure also retrieved a significant number of false positive metalloproteins which are known to bind metals other than Cu.[386]

To solve this problem, additional searches were integrated with domain recognition methods, which showed better results with regard to sensitivity and selectivity.[38,389,390] Recently, using this modified approach, the occurrence of Cu-binding proteins in 57 completely sequenced genomes in prokaryotes and eukaryotes was examined.[391] It was found that the size of the Cu proteome is generally less than 1% of the total proteome of an organism. Bacterial proteomes contain between 0% and 1.3% Cu proteins, with an average of 0.3%, and eukaryotic proteomes have between 0.2% and 0.5%, with an average of 0.3%. Among analyzed prokaryotes, *Nitrosomonas europaea* and *Halobacterium sp.* contain the largest percentage of Cu-binding proteins in bacteria and archaea, respectively. The former evolved a large number of Cu-dependent oxidoreductase enzymes, whereas the latter possesses the largest archaeal repertoire of proteins belonging to the plastocyanin/azurin family. Among eukaryota, the largest proteome fraction of Cu proteins was observed in *A. thaliana* and *S. cerevisiae*. The number of putative Cu-binding proteins did not correlate with the size of the proteome, which is different from the cases for several other metals, such as Zn.[38] A small number of organisms (5 out of 57 organisms in this study), all of which were host-associated, appeared to lack Cu-binding proteins. Functional prediction of Cu-binding proteins indicated that these proteins are likely to be part of a network which may thus represent an ancient core that is crucial for Cu homeostasis.[391] It appears that the speciation of prokaryotic organisms affected only slightly this ancestral Cu proteome. On the other hand, eukaryotes may have expanded their ancestral repertoires of Cu proteins, by inventing new Cu domains and reusing old domains for new functions.

Recently, we carried out a separate comparative genomic analysis of sequenced archaea and bacteria that yielded a comprehensive view of Cu utilization in prokaryotes.[40] Using

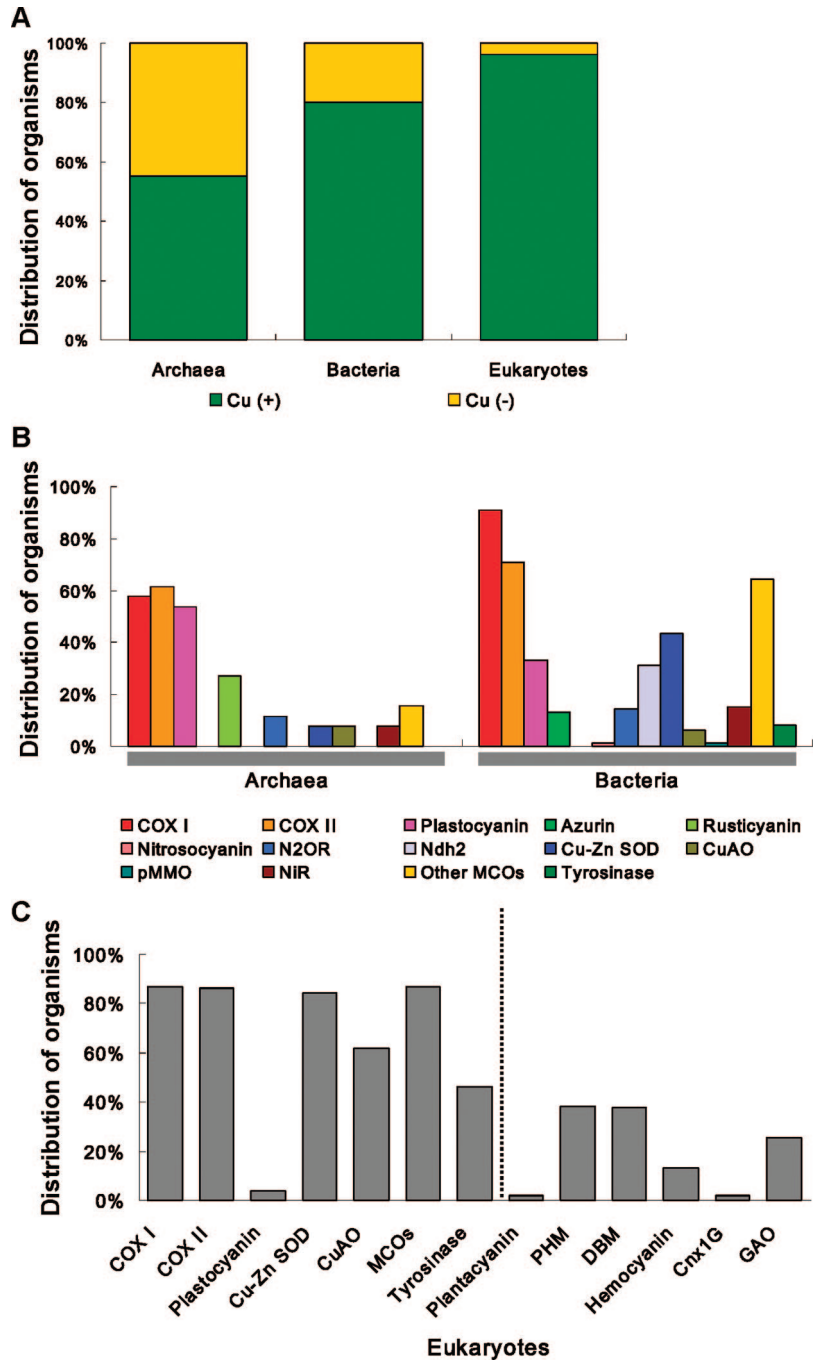

Figure 9. Occurrence of Cu utilization in the three domains of life. (A) Proportion of Cu-utilizing organisms among organisms with sequenced genomes. All organisms were classified into two groups: Cu (+), i.e., containing the Cu utilization trait; Cu (−), i.e., lacking the Cu utilization trait. (B) Occurrence of Cu-dependent proteins in Co-utilizing prokaryotes. (C) Occurrence of Cu-dependent proteins in Co-utilizing eukaryotes. Protein families on the left side of the dotted line have Cu-containing homologs in bacteria whereas others were only found in eukaryotes. COX I, cytochrome *c* oxidase subunit I; COX II, cytochrome *c* oxidase subunit II; N_2OR, nitrous oxide reductase; Ndh2, NADH dehydrogenase 2; Cu−Zn SOD, Cu−Zn superoxide dismutase; CuAO, Cu amine oxidase; pMMO, particulate methane monooxygenase; NiR, nitrite reductase; MCOs, multi-Cu oxidases; PHM, peptidylglycine α-hydroxylating monooxygenase; DBM, dopamine β-monooxygenase; GAO, galactose oxidase.

the same approach introduced in section 2, the occurrence of both Cu transporters, including CopA, CutC, and Cus-CFBA, and strictly Cu-dependent proteins (i.e., those that cannot substitute other metal ions for Cu) was examined. In addition, we also analyzed Cu utilization in sequenced eukaryotes (unpublished data), which extended the previous studies of Cu utilization in this domain. The distribution of Cu-utilizing organisms and Cu-dependent proteins is illustrated in Figure 9.

Consistent with previous observations, Cu is widely used by bacteria, with more than 70% of analyzed organisms being Cu-utilizing (i.e., having at least one Cu-dependent protein). In contrast, all or almost all organisms in some phyla (such as Thermotogae, Firmicutes/Mollicutes, Chlamydiae, and Spirochaetes) do not contain known Cu-dependent proteins. In archaea, the trend was somewhat reversed, and only half of organisms appeared to utilize Cu (Figure 9A). Analysis of Cu transporters revealed that they had different patterns

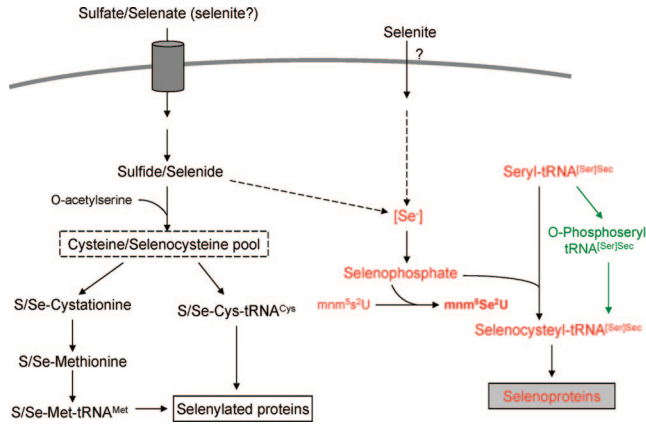

Figure 10. General scheme of Se metabolism. The specific pathway of Sec incorporation into proteins in bacteria is shown in red. A different intermediate step of Sec biosynthesis in archaea and eukaryotes is shown in green. [Se^-] represents the reactive selenium species used by the selenophosphate synthetase as substrate for the biosynthesis of selenophosphate.

of occurrence than Cu-dependent proteins. CopA was the most widespread Cu exporter in bacteria and was the only Cu transporter detected in archaea.[40] Many organisms, including those that lack Cu-dependent proteins, had Cu exporters. These data suggested that the pathways of Cu utilization and detoxification are independent and that many organisms likely protect themselves against Cu ions that inadvertently enter the cell. Thus, occurrence of Cu transporters in prokaryotes may not provide sufficient information about Cu utilization; however, it may be important for understanding of Cu homeostasis. In eukaryotes, the occurrence of Cu importer Ctr1 and exporter ATP7 was consistent with that of the Cu utilization trait (Ctr1 was detected in more than 90% of Cu-utilizing organisms and ATP7 in all Cu-utilizing organisms, unpublished data). Compared to most organisms that have 1−3 *ctr1* genes, the *Caenorhabditis* species (nematodes) possess a high number of *ctr1* genes, especially *C. elegans* that possesses 11 such genes, suggesting complex mechanisms of Cu uptake and trafficking in these organisms (unpublished data).

Among Cu-dependent proteins, COX I and COX II were the most frequently used Cu-containing proteins, which were identified in the majority of Cu-utilizing organisms in all three domains of life (Figure 9B and 9C). In addition, various MCOs were detected in most Cu-utilizing bacteria and eukaryotes. The occurrence of other Cu-containing proteins was relatively restricted. Cu−Zn SOD and CuAO appeared to be more widely used in eukaryotes than in prokaryotes. Homologs of almost half of prokaryotic Cu-dependent proteins could not be detected in eukaryotes, suggesting they either evolved in prokaryotes or were lost in eukaryotes. In addition, novel Cu-binding proteins evolved in eukaryotes, such as plantacyanin and Cnx1G in plants. Plants possess the largest Cu-dependent proteomes (cuproproteomes). Most of these proteins belong to plantacyanin, CuAO, and MCO families, suggesting important roles of these Cu-dependent proteins in plant metabolism. Only 15−16 Cu-dependent proteins were found in nematodes (unpublished data). It is unclear why nematodes need such a complex Cu uptake mechanism in the face of restricted Cu utilization. It is possible that unknown Cu-dependent proteins are present in these organisms.

An interesting finding was that organisms living in oxygen-rich environments utilize Cu, whereas the majority of anaerobic organisms do not.[40] In addition, among Cu users, cuproproteomes of aerobic organisms were generally larger than those of anaerobic organisms. These data are consistent with the idea that proteins evolved to utilize Cu following the oxygenation of the Earth.

6. Selenium

Selenium (Se) is an essential trace element in many organisms, including humans, yet it is required only in small amounts.[392,393] This element is known primarily for its functions in redox homeostasis and is recognized as one of promising cancer chemopreventive agents.[394−396] There are strong indications that it also has a role in antivirus activity, in anti-inflammatory activity, in preventing heart disease and other cardiovascular and muscle disorders, and in delaying the progression of AIDS.[397−399] In addition, Se is required for mammalian development, male reproduction, and immune function.[400−403]

6.1. Selenium Uptake and Its Major Biological Forms

Se occurs as inorganic selenite or selenate and in organic forms in organisms. It can be utilized in a specific and nonspecific manner. The nonspecific pathway is based on the chemical similarity between Se and sulfur, its neighbor element in the periodic table. It appears that Se is metabolized along the pathways of sulfur metabolism.[404,405] Se may be taken up, in the form of selenite, by the sulfate transport system and reduced to selenide via the assimilatory sulfate reduction system.[406] Recently, it was proposed that selenite uptake by plants may be mediated, at least partly, by phosphate transporters.[407]

A general scheme for metabolism and incorporation of Se into macromolecules is shown in Figure 10. The major organic forms of Se include Sec, selenomethionine (SeMet), selenosugar, and methylated low-molecular-weight Se compounds.[408] Free Sec can be converted to SeMet and is a substrate for cysteyl-tRNA synthetase, which forms selenocysteyl-$tRNA^{Cys}$ and incorporates Sec nonspecifically at Cys positions in proteins.[409] The decision whether Se is incorporated nonspecifically as either Sec or SeMet should be dependent on the relative catalytic efficiencies of cysteyl-tRNA synthetase and cystathionine synthetase for the substrate Cys and its analog Sec.[404] SeMet is the major selenocompound in plants such as grains, legumes, and soybeans and was also detected as the major selenocompound

in bacteria when growing on excessive amounts of selenite.[404] On the other hand, the specific incorporation of Sec is efficient at much lower amounts of selenite and does not produce free Sec because the biosynthesis of Sec occurs on the specific tRNA. Accordingly, free Sec, if present in cells, cannot be inserted into selenoproteins. With regard to this review, only specific Sec incorporation is relevant and will be further discussed as a subject for comparative genomic studies.

6.2. Selenocysteine: The 21st Amino Acid

Unlike metals that directly bind to proteins or are part of cofactors, Se is cotranslationally incorporated into selenoproteins and occurs in these proteins in the form of Sec, the 21st amino acid in the genetic code.[410−412] Such a form of Se in protein is widespread in all domains of life and is responsible for the majority of biological effects of Se.[410−414]

Understanding Sec biosynthesis pathways is essential for comparative genomics of Se utilization. The biosynthesis of Sec and its incorporation into nascent polypeptides requires a complex molecular machinery that recodes in-frame UGA codons, which normally function as stop signals, to serve as Sec codons.[408,415,416] The mechanisms of selenoprotein biosynthesis have been the subject of numerous studies.[410,411,413,417−423] While the fundamental mechanism of Sec insertion in the three domains of life appears to be similar, prokaryotes and eukaryotes evolved unique components that provide specific regulation of Sec biosynthesis.

6.2.1. Biosynthesis of Sec in Prokaryotes

In bacteria, the mechanism of Sec insertion in response to UGA has been most thoroughly elucidated in *E. coli* by Böck and collaborators.[409−411,424−426] The biosynthesis and specific insertion of Sec into proteins requires an in-frame UGA codon, a *cis*-acting Sec insertion sequence (SECIS) element which is a stem loop structure within the selenoprotein mRNA immediately downstream of the Sec-encoding UGA codon, and several trans-acting factors dedicated to Sec incorporation.[413,420] A model of Sec insertion into proteins in bacteria is shown in Figure 11A.

Briefly, the SECIS element binds the Sec-specific elongation factor SelB and forms a quaternary complex with Sec-tRNA$^{[Ser]Sec}$, a unique tRNA whose anticodon matches the UGA codon, and GTP. tRNA$^{[Ser]Sec}$ is initially charged by seryl-tRNA synthetase with serine and serves as an adaptor for the conversion of the seryl moiety into the selenocysteyl product by Sec synthase (SelA). SelA utilizes selenophosphate provided by selenophosphate synthetase (SelD) as the selenium donor. During translation, the quaternary complex is translocated toward the ribosome, and the lower helical part of the SECIS element is melted. When the UGA arrives at the A site, SelB makes contact with the ribosome and the charged Sec-tRNA$^{[Ser]Sec}$ is released after accommodation in the A-site. After translocation of the ribosome through the melted SECIS element, the RNA can refold and serve as a target for the formation of a new quaternary complex to assist the next ribosome in decoding the UGA codon.[404]

6.2.2. Biosynthesis of Sec in Eukaryotes and Archaea

Recent studies on Sec biosynthesis in eukaryotes and archaea revealed a pathway that contains additional steps and enzymes compared to the pathway of Sec biosynthesis

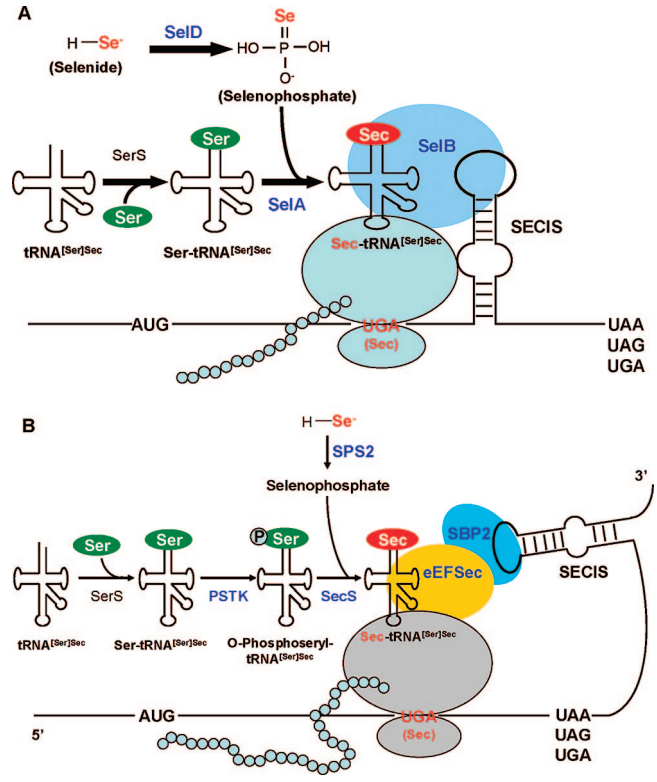

Figure 11. Models of Sec incorporation in bacteria and eukaryotes. Specific proteins involved in Sec biosynthesis and insertion into selenoproteins are highlighted in blue. (A) Sec insertion in bacteria. SelA, bacterial Sec synthase; SelB, bacterial Sec-specific elongation factor; SelD, selenophosphate synthetase; SerS, seryl-tRNA synthetase. (B) Sec insertion in eukaryotes. SPS2, selenophosphate synthetase 2; SecS, eukaryotic Sec synthase; PSTK, *O*-phosphoseryl-tRNA$^{[Ser]Sec}$ kinase; eEFSec, eukaryotic Sec-specific elongation factor; SBP2, SECIS binding protein 2.

in bacteria. In eukaryotes, the initial steps of Sec biosynthesis occur in the cytosol, whereas the maturation Sec-tRNA$^{[Ser]Sec}$ probably occurs in the nucleus.[427] The mechanism of Sec insertion in eukaryotes is shown in Figure 11B. Some eukaryotes, including all vertebrates, contain two distinct SelD homologs, SPS1 and SPS2. Only SPS2 synthesizes selenophosphate from selenide and is essential for generating the selenium donor for Sec biosynthesis.[428] By analogy with the pathway in prokaryotes, a Sec synthase is needed to convert the seryl-tRNA$^{[Ser]Sec}$ to Sec-tRNA$^{[Ser]Sec}$, but SelA is absent in eukaryotes. The recent identification and characterization of eukaryotic Sec synthase (SecS) demonstrates that Sec is synthesized on the tRNA$^{[Ser]Sec}$ from seryl-tRNA$^{[Ser]Sec}$ using seryl-tRNA synthetase, *O*-phosphoseryl-tRNA$^{[Ser]Sec}$ kinase (PSTK) and SecS.[421−423] An additional protein named Secp43 can also bind Sec-tRNA$^{[Ser]Sec}$, which might be involved in methylation of Sec-tRNA$^{[Ser]Sec}$ and promoting assembly of both Sec biosynthesis and incorporation complexes.[427,429] The eukaryotic SECIS element is located in the 3′-UTR of selenoprotein mRNAs; therefore, additional factors are required for elongation and SECIS binding function. The eukaryotic Sec-specific elongation factor eEFSec, a distant homolog of prokaryotic SelB, only binds Sec-tRNA$^{[Ser]Sec}$.[417,430,431] Two additional factors have been identified to bind SECIS element and ribosomal sites, SECIS binding protein 2 (SBP2) and ribosomal protein L30.[432,433] SBP2 is composed of three major domains: an N-terminal domain which may have a selenoprotein-specific

regulatory function, a functional domain involved in Sec incorporation, and a C-terminal RNA-binding domain required for interaction with the SECIS element.[434] The ribosomal protein L30 was identified to have a SECIS-binding activity and may compete with SBP2, suggesting it may be a component of the eukaryotic Sec recoding machinery. It has been proposed that the SECIS element acts as a molecular switch and that the dynamic exchange between L30 and SBP2 for the SECIS depends on the preformed complex.[433]

Homologs of SecS and PSTK were also detected in Sec-utilizing archaea, implying that archaea use a similar mechanism for Sec biosynthesis. However, the absence of other proteins (such as Secp43 and SBP2) also highlights differences in Sec incorpation between archaea and eukaryotes. In addition, archaeal SECIS elements are completely different from those in both bateria and eukaryotes and may be located in both 3′-UTRs and 5′-UTRs.

6.3. Selenoproteins

Historically, selenoproteins have been identified by the presence of Se in protein fractions during isolation. Selenoproteins can be metabolically labeled with ^{75}Se, which can be visualized on polyacrylamide gels with a PhosphorImager. Using this technique, several selenoproteins were identified in prokaryotes and eukaryotes.[435−437] In recent years, remarkable progress in large-scale sequencing projects provided an opportunity and resources for selenoprotein discovery. To this end, a variety of computational approaches and tools have been developed to identify selenoprotein genes in genome databases.

6.3.1. In Silico Identification of Selenoproteins in Eukaryotes and Prokaryotes

All selenoprotein genes have both Sec-encoding UGA codon and the SECIS element. The SECIS element is an essential and highly specific structure for Sec insertion and contains conserved features which can be utilized for its prediction.[413] As discussed above, in archaea and eukaryotes, SECIS elements are most often located in the 3′-UTR of selenoprotein genes. The eukaryotic SECIS element is composed of two helices separated by an internal loop; a SECIS core structure, called Quartet, located at the base of helix 2; and an apical loop (see Figure 12A). The Quartet is formed by four non-Watson−Crick base pairs and is the main functional site of the SECIS element.[438] When the apical loop is large enough, an additional ministem is formed that presumably stabilizes the SECIS element. The presence of this ministem was used to classify SECIS elements into form 1 (lacking ministem) and form 2 (having ministem) structures.[439] Two unpaired adenosines in the apical loop are also found in the majority of selenoprotein genes.[440] The archaeal SECIS elements differ from those in the eukaryote and display a motif containing a purine-only GAA sequence, an internal loop, and three consecutive C-G or G-C base pairs.[441]

As discussed above, bacterial SECIS elements differ from both eukaryotic and archaeal structures with respect to sequence and structure, and are located immediately downstream of Sec-encoding UGA codons in the coding regions of selenoprotein genes.[442] To date, the best characterized bacterial SECIS elements are in the genes encoding formate dehydrogenases H (fdhF), N (fdnG), and O (fdoG) in *E. coli*. They are composed of two domains:

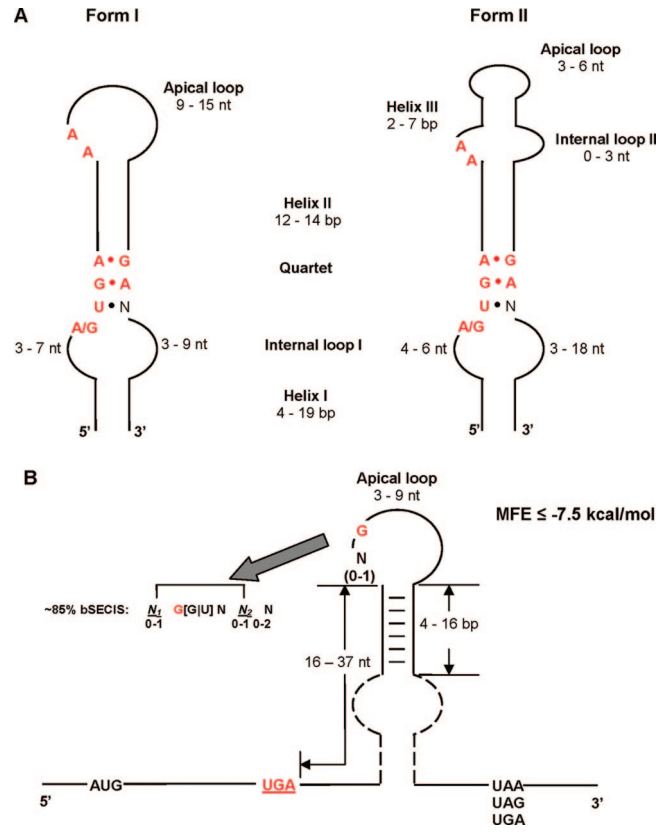

Figure 12. SECIS elements in eukaryotes and bacteria. (A) Two forms of eukaryotic SECIS elements. The allowed lengths of helices and loops are indicated. Conserved nucleotides are shown in red, including the four non-Watson−Crick base pairs (quartet or SECIS core) and two unpaired adenosines in the apical loop of form I or the internal loop of form II structures. (B) A consensus model for bacterial SECIS elements. The Sec UGA codon is underlined. The bacterial SECIS model includes the following: (i) a 3−9 nt apical loop and a 4−16 bp upper-stem; (ii) at least one guanosine (G) among the first two nucleotides (N means any nucleotide) in the apical loop, most of which contain the N_1(0−1)G[G U]NN_2(0−1)N(0−2) pattern in which N_1 and N_2 form a base pair; (iii) a spacing of 16−37 nt between the UGA codon and the apical loop; and (iv) minimum free energy (MFE) ≤ −7.5 kcal/mol.

one containing a Sec UGA codon and the other a 17-nt stem-loop separated from UGA by 11 nucleotides. An exposed GU in the apical loop and a bulged UU in the upper stem are regarded as a common core of the *E. coli* SECIS elements.[443,444] A fixed distance between the in-frame UGA codon and the apical loop is also important for SECIS function.[445] However, putative SECIS elements identified in selenoprotein mRNAs in other bacteria appeared to bear no resemblance to each other or to the *E. coli* counterparts with respect to loop sequences or lengths of the stems.[446,447] Recent studies on various bacterial SECIS elements showed that a single G, often followed with a U, is present among the first two nucleotides in a small apical loop and that the spacing between the UGA codon and the apical loop is 18−23 nucleotides for most bacterial SECIS elements.[448] A model of bacterial SECIS elements is shown in Figure 12B. These data suggested that the majority of bacterial SECIS elements can be described by a common structural model and that these structures probably occur exclusively in the downstream sequences flanking the UGA.

In recent years, several bioinformatics algorithms have been developed for the prediction of selenoprotein genes in eukaryotic, archaeal, and bacterial genomes on the basis of

Table 4. Selenoprotein Families

eukaryotes	prokaryotes
Selenoproteins in mammals	formate dehydrogenase (Fdh)
deiodinase family	formylmethanofuran dehydrogenase (FMDH)
DI, DII, and DIII	selenophosphate synthetase (SelD)
glutathione peroxidase (GPx) family	coenzyme F420-reducing hydrogenase α subunit (FrhA)
GPx1, Gpx2, Gpx3, Gpx4, and GPx6	coenzyme F420-reducing hydrogenase δ subunit (FrhD)
thioredoxin reductase (TR) family	methylviologen-reducing hydrogenase α subunit (VhuA)
TR1, TGR, and TR3	glycine reductase selenoprotein A (GrdA)
15-kDa selenoprotein (Sep15)	glycine reductase selenoprotein B (GrdB)
methionine-R-sulfoxide reductase 1 (MsrB1)	peroxiredoxin (Prx)
selenophosphate synthetase 2 (SPS2)	thioredoxin (Trx)
selenoprotein P (SelP)	glutaredoxin (Grx)
selenoprotein W (SelW)	heterodisulfide reductase α subunit (HdrA)
selenoprotein H (SelH)	HesB-like
selenoprotein I (SelI)	proline reductase PrdB
selenoprotein K (SelK)	deiodinase-like
selenoprotein M (SelM)	GPx
selenoprotein N (SelN)	SelW-like
selenoprotein O (SelO)	MsrA
selenoprotein S (SelS)	DsbG-like
selenoprotein T (SelT)	Fe−S oxidoreductase (GlpC)
selenoprotein V (SelV)	DsbA-like
	DsrE-like
Other selenoproteins	AhpD-like
methionine-S-sulfoxide reductase (MsrA)	arsenate reductase
protein disulfide isomerase (PDI)	molybdopterin biosynthesis protein MoeB-like
selenoprotein U (SelU)	glutathione S-transferase-like (GST-like)
selenoprotein J (SelJ)	OsmC-like
selenoprotein L (SelL)	NADH:ubiquinone oxidoreductase subunit E
fish 15 kDa selenoprotein-like (Fep15)	rhodanese-related protein
SAM-dependent methyltransferase	methylated-DNA−protein−cysteine methyltransferase
peroxiredoxin (Prx)-like	UGSC-containing protein
thioredoxin (Trx)-fold protein	arsenite *S*-adenosylmethyltransferase
membrane selenoprotein (MSP)	Prx-like proteins
hypothetical proteins	Trx-like proteins
	Grx-like proteins
	Trx-fold proteins
	hypothetical proteins

SECIS elements.[32,33,448−451] The general strategy is to find candidate SECIS elements and then analyze upstream regions to identify coding regions, and finally test candidate selenoproteins by ^{75}Se labeling assays. Based on these approaches, a number of novel selenoproteins have been discovered and characterized.[31,33,448,452−455]

Considering that most selenoproteins have homologs in which Sec is replaced with Cys, additional computational approaches were developed, which employ Cys-containing proteins and comprehensive protein databases (e.g., nonredundant protein database in NCBI) to search nucleotide sequence databases for selenoprotein genes.[33,456] These approaches were designed to identify TGA-containing nucleotide sequences which, when translated, are homologous to query Cys-containing proteins such that the conserved Cys residues align with translated TGA codons, i.e., these Cys/TGA-containing (or Cys/Sec) pairs should be flanked by conserved regions. Further analyses (e.g., the presence of SECIS and the occurrence of a predicted Sec-TGA codon in other organisms) were carried out with these candidate selenoproteins to identify new selenoprotein genes. Using these methods, a large number of novel selenoproteins were identified in both completely sequenced genomes and large scale environmental genome sequencing projects.[33,456−458] One deficiency of this approach is the inability to identify selenoproteins, which have no Cys-containing homologs, but luckily such situations are extremely rare.

Both SECIS-dependent and SECIS-independent algorithms described above identify similar sets of selenoprotein genes in various genomes and environmental databases, implying that both methods have excellent performance and that all or almost all selenoproteins could be identified by these bioinformatics tools.

6.3.2. Eukaryotic Selenoproteins

In the past several years, a number of studies have reported on the identification of novel selenoprotein genes in eukaryotes, most of which were discovered by the bioinformatics approaches described above (see section 6.3.1). For example, a total of 25 and 24 selenoproteins were identified in human and mouse, respectively.[31] A complete list of known eukaryotic selenoproteins is shown in Table 4. Here we only discuss the majority of these selenoproteins, focusing on those in mammals. Additional details about selenoproteins can be found in the recently published book on the molecular biology of Se.[459]

Glutathione peroxidases (GPxs), the largest group of mammalian selenoproteins, are antioxidant and redox regulatory enzymes that catalyze thiol-dependent hydroperoxide reduction. In mammals, there are eight GPx homologs and five of them are selenoproteins, including GPx1 (or cGPx, the first known animal selenoprotein),[460] GPx2 (or GI-GPx), GPx3 (or pGPx), GPx4 (or PHGPx), and GPx6. Among them, only GPx4 is known to be essential during embryogenesis in mammals.[461] In addition, it may serve a structural role in mature sperm and has been implicated in site-specific disulfide bond formation.[401,462] The functions of several Sec-containing GPxes remain unclear. Sec-containing GPx

homologs were also detected in single-cellular eukaryotes and bacteria.[455−457] They likely evolved by convergent evolution.

Deiodinases represent another group of animal selenoproteins. In mammals, there are three deiodinases (DI, DII, and DIII, all of which are selenoproteins) which activate or inactivate thyroid hormones by reductive deiodination.[463−465] Homologs of Sec-containing deiodinases were also found in bacteria and some unicellular organisms.[456,457]

Mammalian thioredoxin reductases (TRs) are selenoenzymes and are very different from the smaller Se-independent enzymes of archaea, bacteria, yeast, and plants.[466,467] Sec was found to be located at the penultimate C-terminus, which is a substrate for the N-terminal thiol-disulfide active site.[468,469] Mammalian cells contain three TRs (all are selenoproteins). TR1 (TrxR1 or TxnRd1) is a cytosolic selenoprotein, whose function is to maintain thioredoxin in the reduced state. TGR (for thioredoxin/glutathione reductase; also known as TR2 or TxnRd3) has an additional N-terminal glutaredoxin domain.[468,470] It can catalyze many thioredoxin- and glutathione-related reactions and may also be involved in the formation of disulfide bonds during sperm maturation.[471] TR3 (also known as TrxR2 or TxnRd2) is a mitochondrial selenoprotein that reduces mitochondrial trioredoxin and glutaredoxin 2.[472,473]

Selenoprotein P (SelP) is the only known mammalian selenoprotein with multiple Sec residues,[474,475] e.g., 10 Sec residues in human SelP and 17 Sec residues in zebrafish SelPa. It is known that SelP is the major plasma selenoprotein, which is synthesized in the liver and delivers selenium to other organs and tissues.[474,475] Brain appears to synthesize its own SelP pool.[476]

Methionine-R-sulfoxide reductase 1 (MsrB1, previously known as SelR or SelX) catalyzes stereospecific reduction of oxidized methionine residues in proteins.[477,478] Two other MsrBs (MsrB2 and MsrB3) are homologous to MsrB1 in which Cys residues are present in place of Sec.[478] MsrB1 is located in the cytosol and the nucleus and is the most active MsrB in mammals.[479]

The 15-kDa selenoprotein (Sep15) was identified several years ago as a protein of unknown function.[437,480,481] This protein is located in the endoplasmic reticulum (ER) and binds UDP-glucose:glycoprotein glucosyltransferase (UGT), an ER chaperone and essential regulator of the calnexin cycle.[482] The calnexin cycle is a quality control pathway localized to the ER that specifically assists in the folding of N-linked glycoproteins.[483] Structual analysis of Sep15 revealed that it has a thioredoxin-like fold and may have redox activity.[484] In addition, Sep15 may be involved in the cancer prevention effect of dietary Se.[485]

Selenoprotein W (SelW) is the smallest mammalian selenoprotein with unclear function.[486,487] The structure of this protein has a thioredoxin-like fold with the CXXU motif located in an exposed loop similarly to the redox-active site in thioredoxin.[488] SelW exhibits glutathione-dependent redox properties *in vivo* and may be a scavenger for reactive oxygen species such as hydrogen peroxide during muscle and nervous system development.[489]

SPS2 is homologous to bacterial SelD, which is used to produce selenophosphate, the selenium donor compound.[428] SPS2 is essential for Sec biosynthesis in eukaryotes (see section 6.2.2).

Other mammalian selenoproteins include selenoprotein H (SelH), selenoprotein I (SelI), selenoprotein K (SelK), selenoprotein M (SelM), selenoprotein N (SelN), selenoprotein O (SelO), selenoprotein S (SelS), selenoprotein T (SelT), and selenoprotein V (SelV). SepN is physically associated with the ryanodine receptor (RyR) and functions as a modifier of the RyR channel.[490] SelS may have an important role in influencing inflammatory response, and that role may be related with SelS as a central component of the retrotranslocation channel in ER-associated protein degradation and its antioxidant function.[491] Other selenoproteins have unknown functions.

Besides, several additional selenoproteins were identified in other eukaryotes. These selenoproteins include methionine-*S*-sulfoxide reductase (MsrA), which catalyzes a stereospecific reduction of methionine-*S*-sulfoxide, protein disulfide isomerase (PDI), which is involved in the formation of the disulfide bond in the ER, selenoprotein U (SelU), selenoprotein J (SelJ), selenoprotein L (SelL), fish 15 kDa selenoprotein (Fep15), and several newly predicted selenoproteins (e.g., SAM-dependent methyltransferase, peroxiredoxin-like, thioredoxin-fold protein, membrane selenoprotein, and a few hypothetical proteins) in protozoa, such as *Plasmodium* and *Ostreococcus*.[453,455,492]

6.3.3. Prokaryotic Selenoproteins

Computational identification of selenoprotein genes in both sequenced prokaryotic genomes and environmental genome projects revealed a much wider distribution of selenoprotein families in prokaryotes.[448,456,457] In addition, Sec-containing homologs of some eukaryotic selenoproteins, such as SPS2, deiodinase, GPx, SelW, and MsrA, were also found in some bacteria. A list of reported prokaryotic selenoprotein families is shown in Table 4.

Formate dehydrogenase (Fdh) is the most widespread prokaryotic selenoprotein family.[493,494] In this enzyme, Sec is directly coordinated with Mo and involved in the oxidation of formate to carbon dioxide.[495−497] Recent studies on the distribution of selenoproteins in prokaryotes revealed that Fdh is the only selenoprotein in many bacterial species, suggesting that it may play an important role in maintaining the Sec utilization trait.[493]

SelD is the second largest prokaryotic selenoprotein family.[493] This protein is homologous to eukaryotic SPS2 and is necessary for Sec biosynthesis in prokaryotes (see section 6.2.1).

Formylmethanofuran dehydrogenase (FMDH) is a distant homolog of Fdh and has a similar function, but uses formylmethanofuran as the substrate. Similar to Fdh, Sec is coordinated to Mo in the active site.

Some Ni-dependent hydrogenases were found to contain Sec, such as coenzyme F420-reducing hydrogenase (α and δ subunits, FrhA and FrhD) and methylviologen-reducing hydrogenase α subunit (VhuA). Sec is coordinated to Ni in these selenoproteins.[498] Two Sec residues were observed in some Sec-containing forms of FrhD.[499]

Glycine reductase selenoprotein A (GrdA) and B (GrdB) are selenoproteins belonging to a multiprotein glycine reductase complex involved in the reduction of glycine, sarcosine, betaine, and other substrates.[500,501] GrdA is known as the only selenoprotein for which no Cys homologs can be detected.

Peroxiredoxin (Prx) is a ubiquitous family of antioxidant enzymes that control peroxide levels and thereby mediate signal transduction in cells. These proteins are present in

essentially all organisms. The Sec-containing forms of Prx have been identified in some bacteria.[502]

Thioredoxin (Trx) is the major intracellular protein disulfide reductant and occurs in all organisms as an essential component. The Sec-containing forms of Trx were found in bacteria.[33,456,457]

Glutaredoxin (Grx) is a small protein disulfide oxidoreductase which uses glutathione as a cofactor. Glutaredoxins are oxidized by substrates and reduced nonenzymatically by glutathione. The Sec-containing forms of Grx were predicted in some bacterial species.[33,456,457]

Other prokaryotic selenoproteins include heterodisulfide reductase, HesB-like, proline reductase, deiodinase-like, GPx-like, SelW-like, MsrA, DsbG-like, Fe−S oxidoreductase, DsrE-like, AhpD-like, arsenate reductase, molybdopterin biosynthesis MoeB, DsbA-like, glutathione S-transferase-like, OsmC-like, rhodanese-related protein, methylated-DNA−protein−cysteine methyltransferase, UGSC-containing protein, arsenic methyltransferase, and a variety of Prx/Trx/Grx-like proteins which contain a Trx-like fold.[33,456,457,493,503] The functions of most of these selenoproteins are not known.

6.4. Selenouridine

In some prokaryotic organisms, selenophosphate is also a Se donor for the biosynthesis of a modified tRNA nucleoside, namely 5-methylaminomethyl-2-selenouridine (mnm^5Se^2U or SeU), which is located at the wobble position of the anticodons of $tRNA^{Lys}$, $tRNA^{Glu}$, and $tRNA^{Gln}$.[494,504] The proposed function of mnm^5Se^2U in these tRNAs involves codon−anticodon interactions that help base pair discrimination at the wobble position and/or translation efficiency.[494,505] A 2-selenouridine synthase (YbbB) is required *in vivo* for the specific substitution of selenium for sulfur in 2-thiouridine residues in these tRNAs.[506] In *E. coli*, a conserved Cys (the second Cys) in the CCXXG motif was demonstrated to be essential for YbbB activity *in vivo*.[506]

6.5. A Candidate Utilization Trait of Selenium

In addition to Sec and selenouridine, Se can be utilized in the form of a cofactor in certain Mo-containing enzymes.[507−510] Nicotinic acid hydroxylase and xanthine dehydrogenase are the best known representatives of this protein class. In these enzymes, Se is covalently bound to Mo in the active site, but the specific structure of the Se cofactor is not clear. Furthermore, recent identification of organisms which contain orphan SelD proteins but lack other known components of Sec and selenouridine traits suggests an additional, unknown use of Se that is also dependent on SelD.[493] Based on gene neighborhood, comparative genomics, and phylogenetic analyses, several candidate proteins, including a SirA-like protein and several other proteins with unclear function, were predicted to be involved in the unknown SelD-dependent Se utilization pathway.[511,512] Metabolic labeling of one organism containing an orphan SelD, *E. faecalis*, with ^{75}Se revealed a protein containing labile Se species that could be released by treatment with reducing agents, suggesting non-Sec utilization of Se in this organism.[512] Further studies are required to determine whether this Se-binding protein or other proteins in organisms with orphan SelDs represent the use of Se, or it is an intermediate state for further delivery to other proteins, such as Mo-dependent hydroxylases.

6.6. Comparative Genomics of Se Utilization

Compared to metals, the utilization of Se is easier to analyze, whether in the form of Sec or selenouridine, on the basis of specific genes involved in Sec and selenouridine biosynthesis. Both SECIS-dependent and SECIS-independent approaches discussed earlier in this review (see section 6.3.1) have been successfully used to search for selenoproteins in prokaryotic and eukaryotic genomes. For example, a computational screen of the entire *Drosophila* genome resulted in the identification of three selenoproteins: SPS2, G-rich, and BthD.[32,513] Other examples include identification of the selenoproteomes in *C. elegans*,[514,515] *C. reinhardtii*,[516] humans and mice,[31] *Plasmodium*,[453] and a variety of archaea and bacteria,[33,448] as well as in environmental genome projects.[456,457] These studies have allowed a detailed view of selenoproteomes in individual organisms and environmental samples.

Although considerable efforts in recent years have been made to elucidate molecular details of Sec decoding in different species, and the selenoproteomes of most widely used model organisms have been the subject of intensive research, fundamental issues regarding the evolution of Sec utilization remain incompletely understood. It has been reported that Sec can greatly increase the catalytic efficiency of selenoenzymes as compared with their Cys-containing homologs.[413] Despite this selective advantage and its dedicated biosynthesis and decoding machinery, Sec is a rare amino acid and is used very selectively in proteins and organisms. Understanding the wide distribution of Sec and yet its restricted use requires further studies.

One study used comparative genomic approaches to generate a map of Sec-incorporating and selenouridine-utilizing organisms, based on the analysis of about 200 completely sequenced prokaryotic genomes.[494] SelB and SelC were defined as the signature of the Sec-decoding trait, and YbbB was defined as the signature of selenouridine synthesis, with SelD defining overall Se utilization. Sec-decoding species overlap and yet are distinct from organisms that synthesize selenouridine. The two Se utilization traits can be independently maintained, although both require SelD. Phylogenetic analysis of Sec-decoding and selenouridine synthesis provided evidence for the ancient origin of these traits and demonstrated that their evolution was a dynamic process. Speciation, differential gene loss, and acquisition of entire sets of genes involved in each trait by horizontal gene transfer were observed, indicating that neither the loss nor the acquisition of the Se utilization trait is irreversible. However, the fact that many organisms that are able to decode Sec use this amino acid only in a small set of proteins or even in a single protein is puzzling. It would be interesting to determine what the factors are that restrict selenoprotein utilization.

A subsequent comprehensive study examined the dynamics of selenium utilization in all sequenced microorganisms at the level of both Se utilization traits and selenoproteomes.[493] In addition to the identification of all components involved in Sec decoding or the selenouridine pathway and all known selenoprotein families, phylogenetic analyses were performed to identify possible evolutionary histories for most proteins involved. The distribution of Sec- and selenouridine-utilizing organisms is shown in Figure 13. First, the searches revealed that less than one-fourth of sequenced bacteria utilize Sec, whose selenoproteomes have 1 to 31 selenoproteins. Selenoprotein-rich organisms (defined as containing six or more selenoproteins) were mostly Deltaproteobacteria or Firmic-

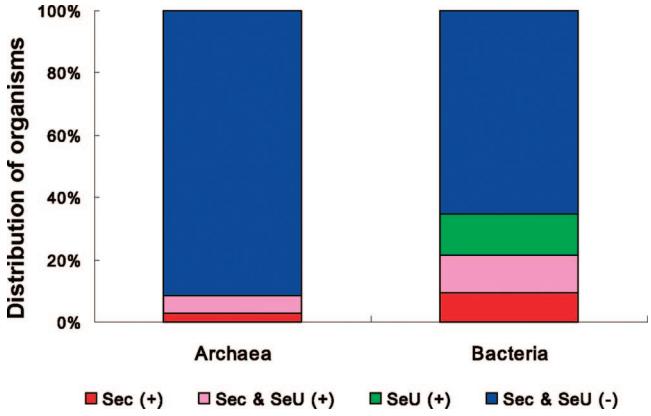

Figure 13. Occurrence of organisms with Se utilization traits in archaea and bacteria. Sec, selenocysteine; SeU, selenouridine. All organisms were classified into four groups: Sec (+), i.e., containing the Sec utilization trait only; Sec & SeU (+), i.e., containing Sec and SeU utilization traits; SeU (+), i.e., containing the SeU utilization trait only; Sec & SeU (−), i.e., containing neither Sec nor SeU utilization traits.

utes/Clostridia. Second, in most selenoprotein families, especially those containing rare selenoproteins and widespread Cys-containing homologs, selenoproteins evolved from Cys-containing ancestors, implying that the Cys-to-Sec replacement is a general trend for most selenoproteins. In contrast, only a small number of Sec-to-Cys conversions were detected, and these were mostly restricted to Fdh that is important for maintaining the Sec-decoding trait in bacteria, and SelD families. Third, specific selenoprotein gene losses were observed in a number of sister genomes of selenoprotein-rich organisms. These observations revealed a dynamic and delicate balance between Sec acquisition and selenoprotein loss. This balance is seen at three levels: loss and acquisition of the Sec-decoding trait itself, with the former as a predominant route; emergence/loss of selenoprotein families; and Cys-to-Sec or Sec-to-Cys replacements in different selenoprotein families. The Sec/Cys replacements were mostly unidirectional, and an increased utilization of Sec by existing protein families was counterbalanced by loss of selenoprotein genes or entire selenoproteomes, which may partially explain the discrepancy between the catalytic advantages offered by Sec and its limited use in nature. Fourth, lateral transfer of the Sec trait was an additional factor, and the first example of selenoprotein gene cluster (*hdrA-frhD-frhA*) transfer between Sec-decoding archaea and Deltaproteobacteria was described. Compared with other lateral gene transfers between archaea and bacteria, selenoprotein gene transfers would be much more difficult because of different mechanisms of Sec insertion into polypeptide chains.[410,413] Coherent clustering of selenoprotein genes in Sec-decoding archaea and Deltaproteobacteria, and the absence of the same operon in closely related organisms indicate that this lateral transfer might have happened recently. Finally, the oxygen requirement and optimal growth temperature appeared to influence Se utilization at the level of both Sec and selenouridine traits. Interestingly, although both of these traits utilize Se, these environmental factors affected the traits in a contrasting manner. Decreases in oxygen concentration and/or increases in optimal growth temperature appeared to preserve and even expand the use of Sec, and the former also promoted the use of Sec forms of selenoprotein families. In contrast, organisms possessing the selenouridine trait (in situations in which the Sec trait has been lost) favored aerobic environment and mesophilic conditions.[493]

A similar comparative analysis was performed to investigate the dynamics of selenoproteins in eukaryotes.[455] The complete selenoproteomes of several newly sequenced model eukaryotes were characterized, including *Ostreococcus tauri* (26 selenoproteins), *O. lucimarinus* (29 selenoproteins), *Dictyostelium discoideum* (5 selenoproteins), *D. pseudoobscura* (3 selenoproteins) and *Thalassiosira pseudonana* (16 selenoproteins). Combined with previously characterized selenoproteomes (such as mammalian selenoproteomes), the number of selenoproteins varied from zero (plants, fungi, and some protists) to 29. Significant differences in the composition of selenoproteomes could be seen even among related organisms. SelK was the most widespread selenoprotein. This protein of unknown function is present in nearly all eukaryotes that utilize Sec (but is replaced with a Cys-containing homolog in nematodes and several other organisms). An additional widespread selenoprotein was SelW, which also occurs in most (but not all) selenoprotein-containing eukaryotes. Several other selenoprotein families, such as GPx and TR, also had a wide distribution. The origin of many selenoproteins in mammals can be traced back to the ancestral, unicellular eukaryotes. Many of these selenoproteins were preserved during evolution and remain in vertebrates (including mammals), green algae, and a variety of protists, whereas many other organisms, including land plants, fungi, nematodes, insects, and some protists, manifested massive, independent selenoprotein losses. Finally, comparative analyses of selenoprotein-rich and -deficient organisms suggested an interesting correlation: large selenoproteomes tend to occur in aquatic life forms, whereas the organisms that lack selenoproteins or have small selenoproteomes are mostly terrestrial (with the notable exception of mammals, whose large bodies and intraorganismal homeostasis support an internal environment that may be less dependent on habitat). Recent chacterization of fish selenoproteomes showed that fish genomes possess a large number of selenoprotein genes (32−37 selenoprotein genes), which provides additional evidence in support of the aquatic/terrestrial hypothesis.[517]

7. Evolutionary Interactions among Trace Elements

It has been suggested that the selection of a specific metal for use in enzymatic catalysis may result from the combination of its specific physicochemical properties, such as redox potential and coordination chemistry, and its accessibility in the environment for biological systems.[518] Although comparative genomics of the utilization of Mo, Ni, Co, Cu, and Se in the three domains of life generally showed independent evolutionary histories, links or common features were observed for several of these trace elements.

7.1. Mo and Se

A correlation between Mo and Se utilization in prokaryotes was observed and appears to be mainly due to Fdh, which is a Moco-binding protein and, in many organisms, is also a selenoprotein.[493,494] Comparison of the distribution of Mo- and Sec-utilizing organisms showed that Sec-utilizing organisms were essentially a subset of Moco-dependent organisms in prokaryotes (Figure 14A). Thus, the Sec utilization trait appeared to be dependent on Mo utilization in prokaryotes,

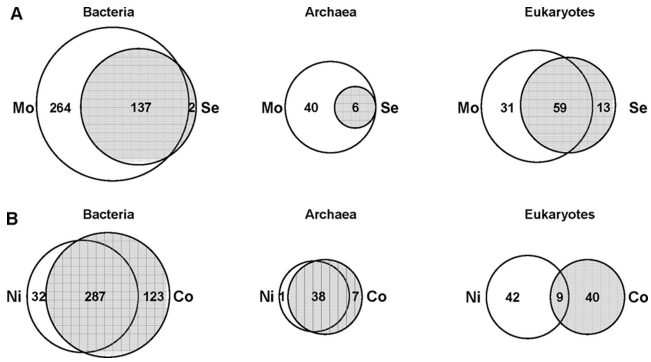

Figure 14. Interactions among trace elements. Relationships between trace element utilization traits are shown by Venn diagrams. (A) Distribution of Mo-utilizing and Sec-utilizing organisms. (B) Distribution of Ni-utilizing and Co-utilizing organisms.

most likely because of the function of Fdh, a widespread molybdoenzyme and the main user of Sec. However, this enzyme is absent in eukaryotes. Accordingly, no link between Se and Mo is seen in eukaryotes, although the majority of Mo-utilizing organisms use Sec.

7.2. Ni and Co

Ni and Co may use the same or similar transport systems in prokaryotes (see details in section 4.1), and many organisms among archaea and bacteria utilize both trace metals (section 4.6). Comparison of the distribution of Ni- and Co-utilizing organisms showed a significant overlap of the two utilization traits in prokaryotes (Figures 7A and 14B). No interaction between Ni and Co is seen in eukaryotes. Thus, the two utilization traits may have independent evolutionary histories and very few of these organisms use Ni and Co.

7.3. Other Interactions

In *A. thaliana*, the C-terminal domain of Cnx1 that catalyzes the insertion of Mo into molybdopterin was found to bind Cu.[383] Either two water molecules (original Cnx1G) or one water molecule and His618 (a Ser583Ala variant) are the Cu ligands. This feature provides a molecular link between Mo and Cu metabolism. However, when analyzing other species, it turned out that such a relationship is not conserved. A recent examination of the *in vivo* and/or *in vitro* activity of two molybdoenzymes, DMSOR and NR, in *E. coli* and *R. sphaeroides*, showed that their activities were not affected when Cu was depleted from the media.[519] Comparative analysis of Cnx1G homologs in other organisms showed that His618 is not conserved (unpublished data). It is possible that while Cu may be utilized during Moco biosynthesis in some organisms such as plants, it does not appear to be strictly required for Moco biosynthesis in many other organisms.

Similarly, a weak correlation between Ni and Se is seen because the Ni−Fe hydrogenase FrhA is a selenoprotein in several organisms. However, the Sec-containing form of Ni−Fe hydrogenase is rare and only detected in Sec-decoding archaea and several deltaproteobacteria.[33,493]

It should be noted that some of the trace elements that are subjects of this review have known interactions with more abundant metals, such as interactions between Cu and Fe, or Cu and Zn (e.g., Cu Zn SOD). But these interactions are beyond the scope of this review.

7.4. Factors That Influence Evolution of Trace Elements

Comparative genomics of trace elements has shown that certain factors may influence their utilization. A recent comparative study of Fe-, Zn-, Mn-, and Co-binding metallomes of more than 300 organisms within the three domains of life showed that the overall abundances of these metal-binding structures (or fold families) may correlate with the theorized changes in the abundances of these metals after the oxygenation of oceanic deep waters, implying that these conserved trends may be proteomic imprints of changes in trace metal bioavailability in the ancient ocean that highlight a major evolutionary shift in biological trace metal usage.[520] Among the factors analyzed, including habitat and environmental and other factors (e.g., oxygen requirement, optimal temperature, optimal pH, GC content, genome size, and gram strain), only habitat types and oxygen requirement appeared to have a significant influence (Figure 15).

Regarding Mo, Ni, and Co utilization, the majority of microbes that do not utilize any of these elements were host-associated, suggesting that the host-associated life style may result in the loss of metal utilization, perhaps due to limited space and resources. Further analysis of host-associated conditions (intra- or extracellular) and the relationship between these organisms and their hosts (i.e., symbiotic or parasitic) showed that the majority of intracellular symbionts and parasites lost the ability to utilize Mo, Ni, or Co, whereas more than 80% of extracellular symbionts utilized all of these metals (Figure 15A). It is known that many parasites and intracellular symbionts have extremely small genomes because many of their genes become dispensable.[521] Thus, it is possible that Mo, Ni, and Co are not necessary for intracellular organisms and hence have been lost due to the pressure on genome size, although these organisms may still depend on host Mo-, Ni-, or Co-dependent proteins. In contrast, their utilization mostly remained intact in extracellular symbionts, presumably because they are still essential for their survival. Similar trends were observed in eukaryotes where significant differences were found between parasites and nonparasites for Mo, Ni, and Co (Figure 15B).

However, such correlations were not observed for Cu and Se. Many parasites and symbionts still utilize Cu, suggesting that Cu is essential for their survival. On the other hand, Se utilization is quite limited and many free-living organisms also lack the ability to use Se. Interestingly, it was observed that Cu was mainly used by aerobic organisms which also have larger cuproproteomes, whereas organisms possessing the Sec-decoding trait and large selenoproteomes favor anaerobic and hyperthermic conditions (Figure 15C).[40,493] Thus, oxygen appeared to play important roles in the evolution of Cu and Se utilization, but with opposite effects. In the future, it would be interesting to identify additional factors which may affect the utilization of trace elements.

8. Conclusions

In this review, we discuss how bioinformatics and comparative genomics can be used to examine the evolution and function of trace element utilization. We describe recent research that used computational studies, especially comparative genomic analyses, to better understand the utilization of five trace elements: Mo, Ni, Co, Cu, and Se. These elements were chosen because they show scattered occurrence in organisms (i.e., some organisms use them and some

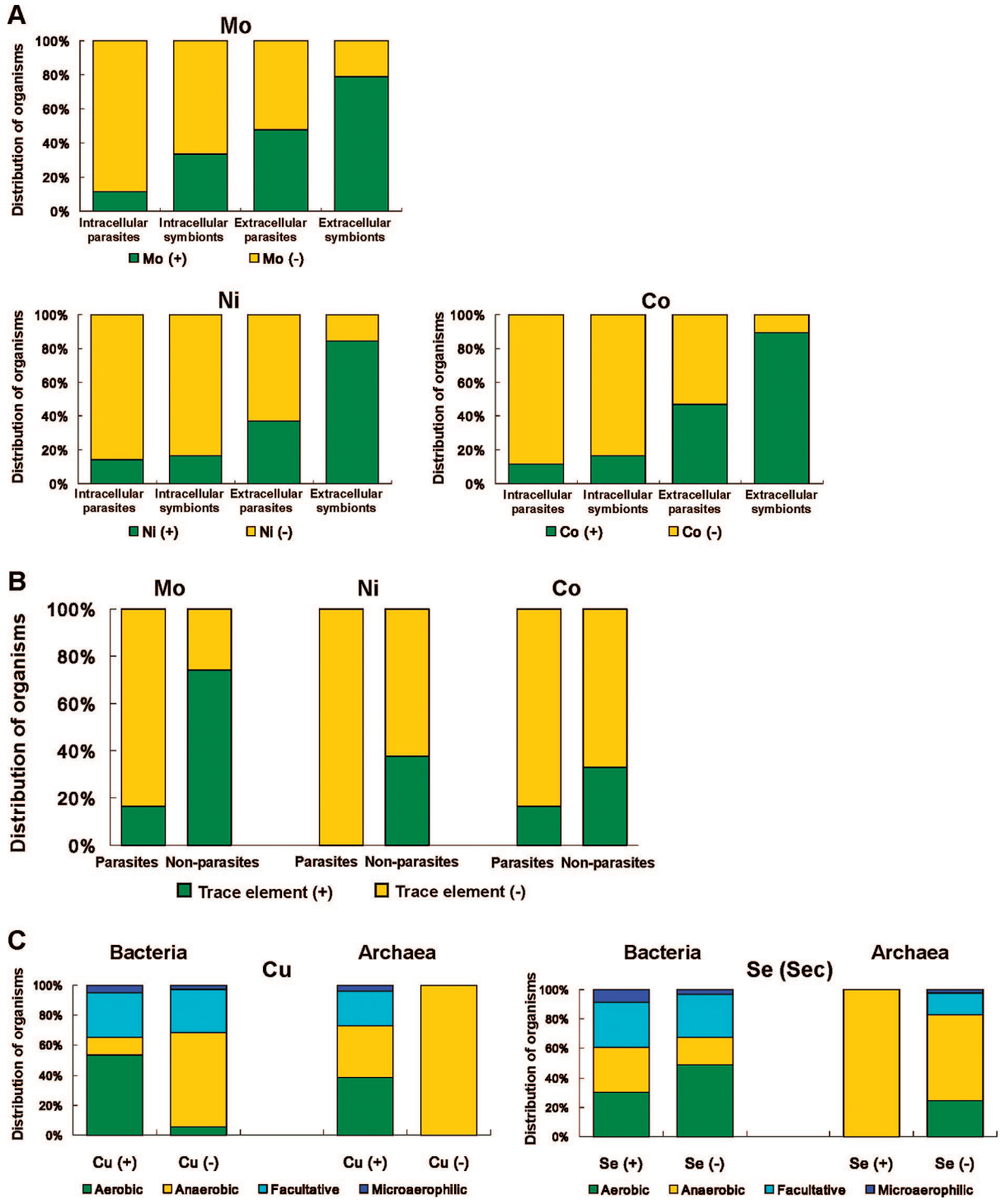

Figure 15. Relationships between various factors and trace element utilization traits. (A) Distribution of organisms with different trace element utilization traits based on various host-associated habitats in bacteria. Host-associated habitats were divided into four subgroups: intracellular symbionts, extracellular symbionts, intracellular parasites, and extracellular parasites. Archaea were not included because too few host-associated organisms in this domain have been sequenced. (B) Distribution of organisms with various trace element utilization traits based on parasitic/nonparasitic habitats in eukaryotes. (C) Distribution of organisms with different trace element utilization traits based on their requirement for oxygen. Four groups were defined: aerobic, anaerobic, facultative, and microaerophilic.

do not) yet are involved in important biological processes in the three domains of life. For these trace elements, most user proteins are well characterized and their dependence on a specific element is evolutionarily conserved. Compared to several other metals, such as Fe, Zn, and Mn, which are more widely used (presumably by all organisms) and more easily interchanged (e.g., many Mn-containing proteins also bind Zn or Fe), or other elements, such as I, Cr, and V, whose utilization is very limited, the five elements discussed in this review are especially attractive targets for comparative genomics analyses.

Much effort has previously been placed on experimental studies of metalloproteins; however, computational analyses of trace element utilization have been limited or lacking for most trace elements. Among the five elements discussed here, only Se may be completely or almost completely characterized because of the dependence of Sec utilization on specific Sec insertion machinery (which can be reliably identified)

and availability of tools for prediction of selenoprotein genes. The utilization of other metals can be understood only partially, especially if the searches rely exclusively on already known metalloproteins. Nevertheless, even incomplete information provides significant advances in our understanding of metal utilization and offers new avenues for further experimental analyses.

Several articles have been published recently that used comparative genomic approaches to analyze the utilization and evolution of trace elements. Most of them adopted similar strategies based on known metalloproteins that are strictly dependent on a specific metal, and factors involved in metal utilization (transporters, regulators, cofactor biosynthesis proteins, etc.). These efforts mark the beginning of a new approach in the area of biological trace elements. These studies may not only help unravel the general principles of utilization of trace elements across the three domains of life but they may also help explain how their utilization changed during evolution and which environmental conditions and factors played a role in these processes.

In spite of the mostly independent utilization of trace elements, several detected biological interactions provide important clues regarding the common features and correlations in their use. It may be expected that, in the next few years, additional computational studies on trace element utilization will continue to generate new insights into the biology of these elements.

9. Acknowledgments

Supported by NIH Grant GM063601 to V.N.G.

10. References

(1) Goldhaber, S. B. *Regul. Toxicol. Pharmacol.* **2003**, *38*, 232.
(2) Hambidge, M. *J. Nutr.* **2003**, *133*, 948S; Suppl 3.
(3) Van Gossum, A.; Neve, J. *Curr. Opin. Clin. Nutr. Metab. Care* **1998**, *1*, 499.
(4) Mertz, W. *Biol. Trace Elem. Res.* **1998**, *66*, 185.
(5) Galaris, D.; Pantopoulos, K. *Crit. Rev. Clin. Lab. Sci.* **2008**, *45*, 1.
(6) Frassinetti, S.; Bronzetti, G.; Caltavuturo, L.; Cini, M.; Croce, C. D. *J. Environ. Pathol. Toxicol. Oncol.* **2006**, *25*, 597.
(7) Foster, L. H.; Sumar, S. *Crit. Rev. Food Sci. Nutr.* **1997**, *37*, 211.
(8) Moghadaszadeh, B.; Beggs, A. H. *Physiology (Bethesda)* **2006**, *21*, 307.
(9) Burk, R. F. *Nutr. Clin. Care* **2002**, *5*, 75.
(10) Patzer, S. I.; Hantke, K. *Mol. Microbiol.* **1998**, *28*, 1199.
(11) Bartsevich, V. V.; Pakrasi, H. B. *J. Biol. Chem.* **1996**, *271*, 26057.
(12) Maupin-Furlow, J. A.; Rosentel, J. K.; Lee, J. H.; Deppenmeier, U.; Gunsalus, R. P.; Shanmugam, K. T. *J. Bacteriol.* **1995**, *177*, 4851.
(13) Eitinger, T.; Mandrand-Berthelot, M. A. *Arch. Microbiol.* **2000**, *173*, 1.
(14) Grass, G.; Franke, S.; Taudte, N.; Nies, D. H.; Kucharski, L. M.; Maguire, M. E.; Rensing, C. *J. Bacteriol.* **2005**, *187*, 1604.
(15) Makui, H.; Roig, E.; Cole, S. T.; Helmann, J. D.; Gros, P.; Cellier, M. F. *Mol. Microbiol.* **2000**, *35*, 1065.
(16) Dancis, A.; Yuan, D. S.; Haile, D.; Askwith, C.; Eide, D.; Moehle, C.; Kaplan, J.; Klausner, R. D. *Cell* **1994**, *76*, 393.
(17) Degen, O.; Eitinger, T. *J. Bacteriol.* **2002**, *184*, 3569.
(18) Ackland, M. L.; McArdle, H. J. *Biometals* **1996**, *9*, 29.
(19) Maguire, M. E. *Methods Mol. Biol.* **2007**, *394*, 289.
(20) Heider, J.; Böck, A. *Adv. Microb. Physiol.* **1993**, *35*, 71.
(21) Bremner, I.; Beattie, J. H. *Proc. Nutr. Soc.* **1995**, *54*, 489.
(22) Templeton, D. M. *Anal. Bioanal. Chem.* **2003**, *375*, 1062.
(23) Rajagopalan, K. V. *Biochem. Soc. Trans.* **1997**, *25*, 757.
(24) Kräutler, B. *Biochem. Soc. Trans.* **2005**, *33*, 806.
(25) Mendel, R. R.; Smith, A. G.; Marquet, A.; Warren, M. J. *Nat. Prod. Rep.* **2007**, *24*, 963.
(26) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647.
(27) Coleman, J. E. *Annu. Rev. Biochem.* **1992**, *61*, 897.
(28) Wattt, R. K.; Ludden, P. W. *Cell. Mol. Life Sci.* **1999**, *56*, 604.
(29) Böck, A.; Forchhammer, K.; Heider, J.; Leinfelder, W.; Sawers, G.; Veprek, B.; Zinoni, F. *Mol. Microbiol.* **1991**, *5*, 515.
(30) Veres, Z.; Tsai, L.; Scholz, T. D.; Politino, M.; Balaban, R. S.; Stadtman, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2975.
(31) Kryukov, G. V.; Castellano, S.; Novoselov, S. V.; Lobanov, A. V.; Zehtab, O.; Guigó, R.; Gladyshev, V. N. *Science* **2003**, *300*, 1439.
(32) Castellano, S.; Morozova, N.; Morey, M.; Berry, M. J.; Serras, F.; Corominas, M.; Guigó, R. *EMBO Rep.* **2001**, *2*, 697.
(33) Kryukov, G. V.; Gladyshev, V. N. *EMBO Rep.* **2004**, *5*, 538.
(34) Panina, E. M.; Mironov, A. A.; Gelfand, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9912.
(35) Rodionov, D. A.; Gelfand, M. S.; Todd, J. D.; Curson, A. R.; Johnston, A. W. *PLoS Comput. Biol.* **2006**, *2*, e163.
(36) Permina, E. A.; Kazakov, A. E.; Kalinina, O. V.; Gelfand, M. S. *BMC Microbiol.* **2006**, *6*, 49.
(37) Valenzuela, L.; Chi, A.; Beard, S.; Orell, A.; Guiliani, N.; Shabanowitz, J.; Hunt, D. F.; Jerez, C. A. *Biotechnol. Adv.* **2006**, *24*, 197.
(38) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. *J. Proteome Res.* **2006**, *5*, 3173.
(39) Rodionov, D. A.; Hebbeln, P.; Gelfand, M. S.; Eitinger, T. *J. Bacteriol.* **2006**, *188*, 317.
(40) Ridge, P. G.; Zhang, Y.; Gladyshev, V. N. *PLoS ONE* **2008**, *3*, e1378.
(41) Zhang, Y.; Gladyshev, V. N. *J. Mol. Biol.* **2008**, *379*, 881.
(42) Hardison, R. C. *PLoS Biol.* **2003**, *1*, e58.
(43) Filipski, A.; Kumar, S. *The Evolution of the Genome*; Elsevier: San Diego, 2005; p 521.
(44) Gregory, T. R.; DeSalle, R. *The Evolution of the Genome*; Elsevier: San Diego, 2005; p 585.
(45) Huynen, M. A.; Snel, B. *Adv. Protein Chem.* **2000**, *54*, 345.
(46) Margulies, E. H.; Birney, E. *Nat. Rev. Genet.* **2008**, *9*, 303.
(47) Gladyshev, V. N.; Kryukov, G. V.; Fomenko, D. E.; Hatfield, D. L. *Annu. Rev. Nutr.* **2004**, *24*, 579.
(48) Shi, W.; Zhan, C.; Ignatov, A.; Manjasetty, B. A.; Marinkovic, N.; Sullivan, M.; Huang, R.; Chance, M. R. *Structure* **2005**, *13*, 1473.
(49) Bateman, A.; Birney, E.; Cerruti, L.; Durbin, R.; Etwiller, L.; Eddy, S. R.; Griffiths-Jones, S.; Howe, K. L.; Marshall, M.; Sonnhammer, E. L. *Nucleic Acids Res.* **2002**, *30*, 276.
(50) Hulo, N.; Bairoch, A.; Bulliard, V.; Cerutti, L.; Cuche, B. A.; de Castro, E.; Lachaize, C.; Langendijk-Genevaux, P. S.; Sigrist, C. J. *Nucleic Acids Res.* **2008**, *36*, D245.
(51) Attwood, T. K.; Bradley, P.; Flower, D. R.; Gaulton, A.; Maudling, N.; Mitchell, A. L.; Moulton, G.; Nordle, A.; Paine, K.; Taylor, P.; Uddin, A.; Zygouri, C. *Nucleic Acids Res.* **2003**, *31*, 400.
(52) Marchler-Bauer, A.; Anderson, J. B.; Derbyshire, M. K.; DeWeese-Scott, C.; Gonzales, N. R.; Gwadz, M.; Hao, L.; He, S.; Hurwitz, D. I.; Jackson, J. D.; Ke, Z.; Krylov, D.; Lanczycki, C. J.; Liebert, C. A.; Liu, C.; Lu, F.; Lu, S.; Marchler, G. H.; Mullokandov, M.; Song, J. S.; Thanki, N.; Yamashita, R. A.; Yin, J. J.; Zhang, D.; Bryant, S. H. *Nucleic Acids Res.* **2007**, *35*, D237.
(53) Henikoff, J. G.; Greene, E. A.; Pietrokovski, S.; Henikoff, S. *Nucleic Acids Res.* **2000**, *28*, 228.
(54) Bru, C.; Courcelle, E.; Carrère, S.; Beausse, Y.; Dalmar, S.; Kahn, D. *Nucleic Acids Res.* **2005**, *33*, D212.
(55) Reichard, K.; Kaufmann, M. *Bioinformatics* **2003**, *19*, 784.
(56) Castagnetto, J. M.; Hennessy, S. W.; Roberts, V. A.; Getzoff, E. D.; Tainer, J. A.; Pique, M. E. *Nucleic Acids Res.* **2002**, *30*, 379.
(57) Waldron, K. J.; Robinson, N. J. *Nat. Rev. Microbiol.* **2009**, *7*, 25.
(58) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.
(59) Rajagopalan, K. V.; Johnson, J. L. *J. Biol. Chem.* **1992**, *267*, 10199.
(60) Mendel, R. R.; Bittner, F. *Biochim. Biophys. Acta* **2006**, *1763*, 621.
(61) Allen, R. M.; Chatterjee, R.; Madden, M. S.; Ludden, P. W.; Shah, V. K. *Crit. Rev. Biotechnol.* **1994**, *14*, 225.
(62) Chan, M. K.; Mukund, S.; Kletzin, A.; Adams, M. W.; Rees, D. C. *Science* **1995**, *267*, 1463.
(63) Hagen, W. R.; Arendsen, A. F. *Struct. Bonding (Berlin)* **1998**, *90*, 161.
(64) L'vov, N. P.; Nosikov, A. N.; Antipov, A. N. *Biochemistry (Moscow)* **2002**, *67*, 196.
(65) Stewart, V. *Microbiol. Rev.* **1988**, *52*, 190.
(66) Hinton, S. M.; Dean, D. *Crit. Rev. Microbiol.* **1990**, *17*, 169.
(67) Grunden, A. M.; Shanmugam, K. T. *Arch. Microbiol.* **1997**, *168*, 345.
(68) Pau, R. N.; Lawson, D. M. *Met. Ions Biol. Syst.* **2002**, *39*, 31.
(69) Bevers, L. E.; Hagedoorn, P. L.; Krijger, G. C.; Hagen, W. R. *J. Bacteriol.* **2006**, *188*, 6498.
(70) Makdessi, K.; Andreesen, J. R.; Pich, A. *J. Biol. Chem.* **2001**, *276*, 24557.
(71) Hu, Y.; Rech, S.; Gunsalus, R. P.; Rees, D. C. *Nat. Struct. Biol.* **1997**, *4*, 703.
(72) Hollenstein, K.; Frei, D. C.; Locher, K. P. *Nature* **2007**, *446*, 213.
(73) Andreesen, J. R.; Makdessi, K. *Ann. N.Y. Acad. Sci.* **2008**, *1125*, 215.
(74) Grunden, A. M.; Ray, R. M.; Rosentel, J. K.; Healy, F. G.; Shanmugam, K. T. *J. Bacteriol.* **1996**, *178*, 735.

(75) Hall, D. R.; Gourley, D. G.; Leonard, G. A.; Duke, E. M.; Anderson, L. A.; Boxer, D. H.; Hunter, W. N. *EMBO J.* **1999**, *18*, 1435.
(76) Tao, H.; Hasona, A.; Do, P. M.; Ingram, L. O.; Shanmugam, K. T. *Arch. Microbiol.* **2005**, *184*, 225.
(77) Anderson, L. A.; Palmer, T.; Price, N. C.; Bornemann, S.; Boxer, D. H.; Pau, R. N. *Eur. J. Biochem.* **1997**, *246*, 119.
(78) McNicholas, P. M.; Mazzotta, M. M.; Rech, S. A.; Gunsalus, R. P. *J. Bacteriol.* **1998**, *180*, 4638.
(79) Wiethaus, J.; Wirsing, A.; Narberhaus, F.; Masepohl, B. *J. Bacteriol.* **2006**, *188*, 8441.
(80) Zahalak, M.; Pratte, B.; Werth, K. J.; Thiel, T. *Mol. Microbiol.* **2004**, *51*, 539.
(81) Delgado, M. J.; Tresierra-Ayala, A.; Talbi, C.; Bedmar, E. J. *Microbiology* **2006**, *152*, 199.
(82) Studholme, D. J.; Pau, R. N. *BMC Microbiol.* **2003**, *3*, 24.
(83) Rodionov, D. A.; Dubchak, I.; Arkin, A.; Alm, E.; Gelfand, M. S. *Genome Biol.* **2004**, *5*, R90.
(84) Tomatsu, H.; Takano, J.; Takahashi, H.; Watanabe-Takahashi, A.; Shibagaki, N.; Fujiwara, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18807.
(85) Tejada-Jiménez, M.; Llamas, A.; Sanz-Luque, E.; Galván, A.; Fernández, E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20126.
(86) Baxter, I.; Muthukumar, B.; Park, H. C.; Buchner, P.; Lahner, B.; Danku, J.; Zhao, K.; Lee, J.; Hawkesford, M. J.; Guerinot, M. L.; Salt, D. E. *PLoS Genet.* **2008**, *4*, e1000004.
(87) Enemark, J. H.; Young, C. G. *Adv. Inorg. Chem.* **1993**, *40*, 1.
(88) Hille, R. *Chem. Rev.* **1996**, *96*, 2757.
(89) Schwarz, G.; Mendel, R. R. *Annu. Rev. Plant Biol.* **2006**, *57*, 623.
(90) Webb, M. E.; Marquet, A.; Mendel, R. R.; Rébeillé, F.; Smith, A. G. *Nat. Prod. Rep.* **2007**, *24*, 988.
(91) Mendel, R. R. *J. Exp. Bot.* **2007**, *58*, 2289.
(92) Shanmugam, K. T.; Stewart, V.; Gunsalus, R. P.; Boxer, D. H.; Cole, J. A.; Chippaux, M.; DeMoss, J. A.; Giordano, G.; Lin, E. C.; Rajagopalan, K. V. *Mol. Microbiol.* **1992**, *6*, 3452.
(93) Mendel, R. R.; Schwarz, G. *Met. Ions Biol. Syst.* **2002**, *39*, 317.
(94) Millar, L. J.; Heck, I. S.; Sloan, J.; Kana'n, G. J.; Kinghorn, J. R.; Unkles, S. E. *Mol. Genet. Genomics* **2001**, *266*, 445.
(95) Reiss, J.; Cohen, N.; Dorche, C.; Mandel, H.; Mendel, R. R.; Stallmeyer, B.; Zabot, M. T.; Dierks, T. *Nat. Genet.* **1998**, *20*, 51.
(96) Stallmeyer, B.; Drugeon, G.; Reiss, J.; Haenni, A. L.; Mendel, R. R. *Am. J. Hum. Genet.* **1999**, *64*, 698.
(97) Stallmeyer, B.; Schwarz, G.; Schulze, J.; Nerlich, A.; Reiss, J.; Kirsch, J.; Mendel, R. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1333.
(98) Kisker, C.; Schindelin, H.; Rees, D. C. *Annu. Rev. Biochem.* **1997**, *66*, 233.
(99) McNaughton, R. L.; Lim, B. S.; Knottenbelt, S. Z.; Holm, R. H.; Kirk, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 4628.
(100) Duval, S.; Ducluzeau, A. L.; Nitschke, W.; Schoepp-Cothenet, B. *BMC Evol. Biol.* **2008**, *8*, 206.
(101) Ma, K.; Hutchins, A.; Sung, S. J.; Adams, M. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9608.
(102) Mulrooney, S. B.; Hausinger, R. P. *FEMS Microbiol. Rev.* **2003**, *27*, 239.
(103) Watt, R. K.; Ludden, P. W. *Cell. Mol. Life Sci.* **1999**, *56*, 604.
(104) Barceloux, D. G. *J. Toxicol. Clin. Toxicol.* **1999**, *37*, 201.
(105) Eitinger, T.; Mandrand-Berthelot, M. A. *Arch. Microbiol.* **2000**, *173*, 1.
(106) Eitinger, T.; Suhr, J.; Moore, L.; Smith, J. A. *BioMetals* **2005**, *18*, 399.
(107) Navarro, C.; Wu, L. F.; Mandrand-Berthelot, M. A. *Mol. Microbiol.* **1993**, *9*, 1181.
(108) de Pina, K.; Navarro, C.; McWalter, L.; Boxer, D. H.; Price, N. C.; Kelly, S. M.; Mandrand-Berthelot, M. A.; Wu, L. F. *Eur. J. Biochem.* **1995**, *227*, 857.
(109) Shepherd, M.; Heath, M. D.; Poole, R. K. *Biochemistry* **2007**, *46*, 5030.
(110) Cherrier, M. V.; Cavazza, C.; Bochot, C.; Lemaire, D.; Fontecilla-Camps, J. C. *Biochemistry* **2008**, *47*, 9937.
(111) Addy, C.; Ohara, M.; Kawai, F.; Kidera, A.; Ikeguchi, M.; Fuchigami, S.; Osawa, M.; Shimada, I.; Park, S. Y.; Tame, J. R.; Heddle, J. G. *Acta. Crystallogr., D: Biol. Crystallogr.* **2007**, *63*, 221.
(112) Sebbane, F.; Mandrand-Berthelot, M. A.; Simonet, M. *J. Bacteriol.* **2002**, *184*, 5706.
(113) Roth, J. R.; Lawrence, J. G.; Rubenfield, M.; Kieffer-Higgins, S.; Church, G. M. *J. Bacteriol.* **1993**, *175*, 3303.
(114) Rodionov, D. A.; Vitreschak, A. G.; Mironov, A. A.; Gelfand, M. S. *J. Biol. Chem.* **2003**, *278*, 41148.
(115) Chen, Y. Y.; Burne, R. A. *J. Bacteriol.* **2003**, *185*, 6773.
(116) Bossé, J. T.; Gilmour, H. D.; MacInnes, J. I. *J. Bacteriol.* **2001**, *183*, 1242.
(117) Degen, O.; Kobayashi, M.; Shimizu, S.; Eitinger, T. *Arch. Microbiol.* **1999**, *171*, 139.
(118) Hebbeln, P.; Eitinger, T. *FEMS Microbiol. Lett.* **2004**, *230*, 129.
(119) Hidalgo, E.; Palacios, J. M.; Murillo, J.; Ruiz-Argüeso, T. *J. Bacteriol.* **1992**, *174*, 4130.
(120) Eitinger, T.; Degen, O.; Bohnke, U.; Muller, M. *J. Biol. Chem.* **2000**, *275*, 18029.
(121) Baginsky, C.; Palacios, J. M.; Imperial, J.; Ruiz-Argüeso, T.; Brito, B. *FEMS Microbiol. Lett.* **2004**, *237*, 399.
(122) Niegowski, D.; Eshaghi, S. *Cell. Mol. Life Sci.* **2007**, *64*, 2564.
(123) Persans, M. W.; Nieman, K.; Salt, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9995.
(124) Kim, D.; Gustin, J. L.; Lahner, B.; Persans, M. W.; Baek, D.; Yun, D. J.; Salt, D. E. *Plant J.* **2004**, *39*, 237.
(125) Conklin, D. S.; McMaster, J. A.; Culbertson, M. R.; Kung, C. *Mol. Cell. Biol.* **1992**, *12*, 3678.
(126) De Pina, K.; Desjardin, V.; Mandrand-Berthelot, M. A.; Giordano, G.; Wu, L. F. *J. Bacteriol.* **1999**, *181*, 670.
(127) Dosanjh, N. S.; Michel, S. L. *Curr. Opin. Chem. Biol.* **2006**, *10*, 123.
(128) Busenlehner, L. S.; Pennella, M. A.; Giedroc, D. P. *FEMS Microbiol. Rev.* **2003**, *27*, 131.
(129) Pennella, M. A.; Giedroc, D. P. *Biometals* **2005**, *18*, 413.
(130) Pennella, M. A.; Shokes, J. E.; Cosper, N. J.; Scott, R. A.; Giedroc, D. P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3713.
(131) Ragsdale, S. W. *Curr. Opin. Chem. Biol.* **1998**, *2*, 208.
(132) Ragsdale, S. W.; Kumar, M. *Chem. Rev.* **1996**, *96*, 2515.
(133) Maroney, M. J. *Curr. Opin. Chem. Biol.* **1999**, *3*, 188.
(134) Ermler, U.; Grabarse, W.; Shima, S.; Goubeaud, M.; Thauer, R. K. *Curr. Opin. Struct. Biol.* **1998**, *8*, 749.
(135) Fontecilla-Camps, J. C.; Volbeda, A.; Cavazza, C.; Nicolet, Y. *Chem. Rev.* **2007**, *107*, 4273.
(136) Clugston, S. L.; Barnard, J. F.; Kinach, R.; Miedema, D.; Ruman, R.; Daub, E.; Honek, J. F. *Biochemistry* **1998**, *37*, 8754.
(137) Sukdeo, N.; Clugston, S. L.; Daub, E.; Honek, J. F. *Biochem. J.* **2004**, *384*, 111.
(138) Greig, N.; Wyllie, S.; Vickers, T. J.; Fairlamb, A. H. *Biochem. J.* **2006**, *400*, 217.
(139) Saint-Jean, A. P.; Phillips, K. R.; Creighton, D. J.; Stone, M. J. *Biochemistry* **1998**, *37*, 10345.
(140) Baird, M. L.; Garber, E. D. *Biochem. Genet.* **1981**, *19*, 1101.
(141) Zonia, L. E.; Stebbins, N. E.; Polacco, J. C. *Plant Physiol.* **1995**, *107*, 1097.
(142) Denkhaus, E.; Salnikow, K. *Crit. Rev. Oncol. Hematol.* **2002**, *42*, 35.
(143) Bencko, V. *J. Hyg. Epidemiol. Microbiol. Immunol.* **1983**, *27*, 237.
(144) Takishima, K.; Suga, T.; Mamiya, G. *Eur. J. Biochem.* **1988**, *175*, 151.
(145) Mobley, H. L.; Hausinger, R. P. *Microbiol. Rev.* **1989**, *53*, 85.
(146) Benini, S.; Ciurli, S.; Nolting, H. F.; Mangani, S. *Eur. J. Biochem.* **1996**, *239*, 61.
(147) Jabri, E.; Carr, M. B.; Hausinger, R. P.; Karplus, P. A. *Science* **1995**, *268*, 998.
(148) Vignais, P. M. *Results Probl. Cell Differ.* **2008**, *45*, 223.
(149) Adams, M. W. *Biochim. Biophys. Acta* **1990**, *1020*, 115.
(150) Horner, D. S.; Heil, B.; Happe, T.; Embley, T. M. *Trends Biochem. Sci.* **2002**, *27*, 148.
(151) Li, C.; Peck, H. D., Jr.; LeGall, J.; Przybyla, A. E. *DNA* **1987**, *6*, 539.
(152) Siegbahn, P. E.; Tye, J. W.; Hall, M. B. *Chem. Rev.* **2007**, *107*, 4414.
(153) Voordouw, G.; Menon, N. K.; LeGall, J.; Choi, E. S.; Peck, H. D., Jr.; Przybyla, A. E. *J. Bacteriol.* **1989**, *171*, 2894.
(154) Vignais, P. M.; Colbeau, A. *Curr. Issues Mol. Biol.* **2004**, *6*, 159.
(155) Lyon, E. J.; Shima, S.; Buurman, G.; Chowdhuri, S.; Batschauer, A.; Steinbach, K.; Thauer, R. K. *Eur. J. Biochem.* **2004**, *271*, 195.
(156) Przybyla, A. E.; Robbins, J.; Menon, N.; Peck, H. D., Jr. *FEMS Microbiol. Rev.* **1992**, *8*, 109.
(157) Ballantine, S. P.; Boxer, D. H. *Eur. J. Biochem.* **1986**, *156*, 277.
(158) Black, L. K.; Fu, C.; Maier, R. J. *J. Bacteriol.* **1994**, *176*, 7102.
(159) Eidsness, M. K.; Scott, R. A.; Prickril, B. C.; DerVartanian, D. V.; Legall, J.; Moura, I.; Moura, J. J.; Peck, H. D., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 147.
(160) Volbeda, A.; Charon, M. H.; Piras, C.; Hatchikian, E. C.; Frey, M.; Fontecilla-Camps, J. C. *Nature* **1995**, *373*, 580.
(161) Ragsdale, S. W. *J. Inorg. Biochem.* **2007**, *101*, 1657.
(162) Bonam, D.; Ludden, P. W. *J. Biol. Chem.* **1987**, *262*, 2980.
(163) Dobbek, H.; Svetlitchnyi, V.; Gremer, L.; Huber, R.; Meyer, O. *Science* **2001**, *293*, 1281.
(164) Gong, W.; Hao, B.; Wei, Z.; Ferguson, D. J., Jr.; Tallant, T.; Krzycki, J. A.; Chan, M. K. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9558.
(165) Lindahl, P. A. *Biochemistry* **2002**, *41*, 2097.
(166) Jeoung, J. H.; Dobbek, H. *Science* **2007**, *318*, 1461.
(167) Maupin-Furlow, J. A.; Ferry, J. G. *J. Bacteriol.* **1996**, *178*, 6849.
(168) Lindahl, P. A. *J. Biol. Inorg. Chem.* **2004**, *9*, 516.
(169) Kumar, M.; Lu, W. P.; Ragsdale, S. W. *Biochemistry* **1994**, *33*, 9769.
(170) Ragsdale, S. W. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 165.

(171) Ermler, U.; Grabarse, W.; Shima, S.; Goubeaud, M.; Thauer, R. K. *Science* **1997**, *278*, 1457.
(172) Friedmann, H. C.; Klein, A.; Thauer, R. K. *FEMS Microbiol. Rev.* **1990**, *7*, 339.
(173) Shima, S.; Thauer, R. K. *Curr. Opin. Microbiol.* **2005**, *8*, 643.
(174) Youn, H. D.; Kim, E. J.; Roe, J. H.; Hah, Y. C.; Kang, S. O. *Biochem. J.* **1996**, *318*, 889.
(175) Dupont, C. L.; Neupane, K.; Shearer, J.; Palenik, B. *Environ. Microbiol.* **2008**, *10*, 1831.
(176) Choudhury, S. B.; Lee, J. W.; Davidson, G.; Yim, Y. I.; Bose, K.; Sharma, M. L.; Kang, S. O.; Cabelli, D. E.; Maroney, M. J. *Biochemistry* **1999**, *38*, 3744.
(177) Barondeau, D. P.; Kassmann, C. J.; Bruns, C. K.; Tainer, J. A.; Getzoff, E. D. *Biochemistry* **2004**, *43*, 8038.
(178) Borths, E. L.; Locher, K. P.; Lee, A. T.; Rees, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16642.
(179) Hvorup, R. N.; Goetz, B. A.; Niederer, M.; Hollenstein, K.; Perozo, E.; Locher, K. P. *Science* **2007**, *317*, 1387.
(180) Cadieux, N.; Phan, P. G.; Cafiso, D. S.; Kadner, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10688.
(181) Köster, W. *Res. Microbiol.* **2001**, *152*, 291.
(182) Seetharam, B.; Bose, S.; Li, N. *J. Nutr.* **1999**, *129*, 1761.
(183) Cuadros, E. V.; Nakayama, Y.; Sequeiro, J. M. *Biochem. Biophys. Res. Commun.* **2005**, *327*, 1006.
(184) Scott, A. I.; Roessner, C. A. *Biochem. Soc. Trans.* **2002**, *30*, 613.
(185) Warren, M. J.; Raux, E.; Schubert, H. L.; Escalante-Semerena, J. C. *Nat. Prod. Rep.* **2002**, *19*, 390.
(186) Debussche, L.; Thibaut, D.; Cameron, B.; Crouzet, J.; Blanche, F. *J. Bacteriol.* **1993**, *175*, 7430.
(187) Roessner, C. A.; Scott, A. I. *J. Bacteriol.* **2006**, *188*, 7331.
(188) Raux, E.; Lanois, A.; Warren, M. J.; Rambach, A.; Thermes, C. *Biochem. J.* **1998**, *335*, 159.
(189) Raux, E.; Lanois, A.; Rambach, A.; Warren, M. J.; Thermes, C. *Biochem. J.* **1998**, *335*, 167.
(190) Roessner, C. A.; Huang, K. X.; Warren, M. J.; Raux, E.; Scott, A. I. *Microbiology* **2002**, *148*, 1845.
(191) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. *Nat. Chem. Biol.* **2008**, *4*, 194.
(192) Kobayashi, M.; Shimizu, S. *Eur. J. Biochem.* **1999**, *261*, 1.
(193) Banerjee, R.; Ragsdale, S. W. *Annu. Rev. Biochem.* **2003**, *72*, 209.
(194) Marsh, E. N. *Essays Biochem.* **1999**, *34*, 139.
(195) Banerjee, R. *Biochemistry* **2001**, *40*, 6191.
(196) Marsh, E. N.; Drennan, C. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 499.
(197) Mancia, F.; Keep, N. H.; Nakagawa, A.; Leadlay, P. F.; McSweeney, S.; Rasmussen, B.; Bösecke, P.; Diat, O.; Evans, P. R. *Structure* **1996**, *4*, 339.
(198) Zerbe-Burkhardt, K.; Ratnatilleke, A.; Philippon, N.; Birch, A.; Leiser, A.; Vrijbloed, J. W.; Hess, D.; Hunziker, P.; Robinson, J. A. *J. Biol. Chem.* **1998**, *273*, 6508.
(199) Erb, T. J.; Berg, I. A.; Brecht, V.; Müller, M.; Fuchs, G.; Alber, B. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10631.
(200) Gruber, K.; Kratky, C. *Curr. Opin. Chem. Biol.* **2002**, *6*, 598.
(201) Beatrix, B.; Zelder, O.; Linder, D.; Buckel, W. *Eur. J. Biochem.* **1994**, *221*, 101.
(202) Chang, C. H.; Frey, P. A. *J. Biol. Chem.* **2000**, *275*, 106.
(203) Daniel, R.; Bobik, T. A.; Gottschalk, G. *FEMS Microbiol. Rev.* **1998**, *22*, 553.
(204) Schrauzer, G. N.; Stadlbauer, E. A. *Bioinorg. Chem.* **1975**, *4*, 185.
(205) Cowles, J. R.; Evans, H. J.; Russell, S. A. *J. Bacteriol.* **1969**, *97*, 1460.
(206) Banerjee, R. *Chem. Biol.* **1997**, *4*, 175.
(207) Francalanci, F.; Davis, N. K.; Fuller, J. Q.; Murfitt, D.; Leadlay, P. F. *Biochem. J.* **1986**, *236*, 489.
(208) Birch, A.; Leiser, A.; Robinson, J. A. *J. Bacteriol.* **1993**, *175*, 3511.
(209) Banerjee, R.; Vlasie, M. *Biochem. Soc. Trans.* **2002**, *30*, 621.
(210) Allen, S. H. G.; Kellermeyer, R. W.; Stjernholm, R.; Wood, H. G. *J. Bacteriol.* **1964**, *87*, 171.
(211) McCarthy, A. A.; Baker, H. M.; Shewry, S. C.; Patchett, M. L.; Baker, E. N. *Structure* **2001**, *9*, 637.
(212) Ratnatilleke, A.; Vrijbloed, J. W.; Robinson, J. A. *J. Biol. Chem.* **1999**, *274*, 31679.
(213) Erb, T. J.; Retey, J.; Fuchs, G.; Alber, B. E. *J. Biol. Chem.* **2008**, *283*, 32283.
(214) Zhang, W.; Reynolds, K. A. *J. Bacteriol.* **2001**, *183*, 2071.
(215) Marsh, E. N. *Bioorg. Chem.* **2000**, *28*, 176.
(216) Hoffmann, B.; Konrat, R.; Bothe, H.; Buckel, W.; Kräutler, B. *Eur. J. Biochem.* **1999**, *263*, 178.
(217) Pierik, A. J.; Ciceri, D.; Lopez, R. F.; Kroll, F.; Bröker, G.; Beatrix, B.; Buckel, W.; Golding, B. T. *Biochemistry* **2005**, *44*, 10541.
(218) Berkovitch, F.; Behshad, E.; Tang, K. H.; Enns, E. A.; Frey, P. A.; Drennan, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15870.
(219) Toraya, T. *Chem. Rec.* **2002**, *2*, 352.
(220) Liao, D. I.; Dotson, G.; Turner, I., Jr.; Reiss, L.; Emptage, M. *J. Inorg. Biochem.* **2003**, *93*, 84.
(221) Bradbeer, C. *J. Biol. Chem.* **1965**, *240*, 4675.
(222) Sun, L.; Warncke, K. *Proteins* **2006**, *64*, 308.
(223) Shibata, N.; Masuda, J.; Tobimatsu, T.; Toraya, T.; Suto, K.; Morimoto, Y.; Yasuoka, N. *Structure* **1999**, *7*, 997.
(224) Jordan, A.; Reichard, P. *Annu. Rev. Biochem.* **1998**, *67*, 71.
(225) Kolberg, M.; Strand, K. R.; Graff, P.; Andersson, K. K. *Biochim. Biophys. Acta* **2004**, *1699*, 1.
(226) Sintchak, M. D.; Arjara, G.; Kellogg, B. A.; Stubbe, J.; Drennan, C. L. *Nat. Struct. Biol.* **2002**, *9*, 293.
(227) Logan, D. T.; Andersson, J.; Sjöberg, B. M.; Nordlund, P. *Science* **1999**, *283*, 1499.
(228) Banerjee, R. V.; Matthews, R. G. *FASEB J.* **1990**, *4*, 1450.
(229) Goulding, C. W.; Postigo, D.; Matthews, R. G. *Biochemistry* **1997**, *36*, 8082.
(230) Sauer, K.; Thauer, R. K. *Eur. J. Biochem.* **1999**, *261*, 674.
(231) Gottschalk, G.; Thauer, R. K. *Biochim. Biophys. Acta* **2001**, *1505*, 28.
(232) Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Lidwig, M. L. *Science* **1994**, *266*, 1669.
(233) Sauer, K.; Thauer, R. K. *Eur. J. Biochem.* **2000**, *267*, 2498.
(234) Hagemeier, C. H.; Krer, M.; Thauer, R. K.; Warkentin, E.; Ermler, U. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18917.
(235) Sauer, K.; Thauer, R. K. *FEBS Lett.* **1998**, *436*, 401.
(236) Krasotkina, J.; Walters, T.; Maruya, K. A.; Ragsdale, S. W. *J. Biol. Chem.* **2001**, *276*, 40991.
(237) Copley, S. D. *Curr. Opin. Chem. Biol.* **1998**, *2*, 613.
(238) Banerjee, R. *ACS Chem. Biol.* **2006**, *1*, 149.
(239) Torrents, E.; Trevisiol, C.; Rotte, C.; Hellman, U.; Martin, W.; Reichard, P. *J. Biol. Chem.* **2006**, *281*, 5604.
(240) Toohey, J. I. *Biofactors* **2006**, *26*, 45.
(241) Roderick, S. L.; Matthews, B. W. *Biochemistry* **1993**, *32*, 3907.
(242) Banerjee, A.; Sharma, R.; Banerjee, U. C. *Appl. Microbiol. Biotechnol.* **2002**, *60*, 33.
(243) Chivers, P. T.; Sauer, R. T. *J. Biol. Chem.* **2000**, *275*, 19735.
(244) Leitch, S.; Bradley, M. J.; Rowe, J. L.; Chivers, P. T.; Maroney, M. J. *J. Am. Chem. Soc.* **2007**, *129*, 5085.
(245) Vitreschak, A. G.; Rodionov, D. A.; Mironov, A. A.; Gelfand, M. S. *RNA* **2003**, *9*, 1084.
(246) Nahvi, A.; Barrick, J. E.; Breaker, R. R. *Nucleic Acids Res.* **2004**, *32*, 143.
(247) Zhang, Y.; Rodionov, D. A.; Gelfand, M. S.; Gladyshev, V. N. *BMC Genomics* **2009**, *10*, 78.
(248) Peña, M. M.; Lee, J.; Thiele, D. J. *J. Nutr.* **1999**, *129*, 1251.
(249) Gaetke, L. M.; Chow, C. K. *Toxicology* **2003**, *189*, 147.
(250) O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **2000**, *275*, 25057.
(251) Silver, S. *Gene* **1996**, *179*, 9.
(252) Vulpe, C. D.; Packman, S. *Annu. Rev. Nutr.* **1995**, *15*, 293.
(253) Finney, L. A.; O'Halloran, T. V. *Science* **2003**, *300*, 931.
(254) Rosen, B. P. *Comp. Biochem. Physiol. A: Mol. Integr. Physiol* **2002**, *133*, 689.
(255) Rensing, C.; Grass, G. *FEMS Microbiol. Rev.* **2003**, *27*, 197.
(256) Solioz, M.; Stoyanov, J. V. *FEMS Microbiol. Rev.* **2003**, *27*, 183.
(257) Roberts, S. A.; Weichsel, A.; Grass, G.; Thakali, K.; Hazzard, J. T.; Tollin, G.; Rensing, C.; Montfort, W. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2766.
(258) Petersen, C.; Møller, L. B. *Gene* **2000**, *261*, 289.
(259) Axelsen, K. B.; Palmgren, M. G. *J. Mol. Evol.* **1998**, *46*, 84.
(260) Solioz, M.; Vulpe, C. *Trends Biochem. Sci.* **1996**, *21*, 237.
(261) Rensing, C.; Ghosh, M.; Rosen, B. P. *J. Bacteriol.* **1999**, *181*, 5891.
(262) Beard, S. J.; Hashim, R.; Membrillo-Hernandez, J.; Hughes, M. N.; Poole, R. K. *Mol. Microbiol.* **1997**, *25*, 883.
(263) Rensing, C.; Mitra, B.; Rosen, B. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14326.
(264) Fan, B.; Grass, G.; Rensing, C.; Rosen, B. P. *Biochem. Biophys. Res. Commun.* **2001**, *286*, 414.
(265) Fan, B.; Rosen, B. P. *J. Biol. Chem.* **2002**, *277*, 46987.
(266) Kanamaru, K.; Kashiwagi, S.; Mizuno, T. *Mol. Microbiol.* **1994**, *13*, 369.
(267) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Kandias, N. G.; Robinson, N. J.; Spyroulias, G. A.; Su, X. C.; Tottey, S.; Vanarotti, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8320.
(268) Tottey, S.; Rich, P. R.; Rondet, S. A.; Robinson, N. J. *J. Biol. Chem.* **2001**, *276*, 19999.
(269) Tottey, S.; Rondet, S. A.; Borrelly, G. P.; Robinson, P. J.; Rich, P. R.; Robinson, N. J. *J. Biol. Chem.* **2002**, *277*, 5490.
(270) Wunderli-Ye, H.; Solioz, M. *Biochem. Biophys. Res. Commun.* **2001**, *280*, 713.
(271) Solioz, M.; Odermatt, A. *J. Biol. Chem.* **1995**, *270*, 9217.
(272) Zhou, L.; Singleton, C.; Le Brun, N. E. *Biochem. J.* **2008**, *413*, 459.
(273) Rosenzweig, A. C.; O'Halloran, T. V. *Curr. Opin. Chem. Biol.* **2000**, *4*, 140.

(274) Gold, B.; Deng, H.; Bryk, R.; Vargas, D.; Eliezer, D.; Roberts, J.; Jiang, X.; Nathan, C. *Nat. Chem. Biol.* **2008**, *4*, 609.
(275) Grass, G.; Rensing, C. *J. Bacteriol.* **2001**, *183*, 2145.
(276) Bagai, I.; Rensing, C.; Blackburn, N. J.; McEvoy, M. M. *Biochemistry* **2008**, *47*, 11408.
(277) Franke, S.; Grass, G.; Rensing, C.; Nies, D. H. *J. Bacteriol.* **2003**, *185*, 3804.
(278) Gupta, S. D.; Lee, B. T.; Camakaris, J.; Wu, H. C. *J. Bacteriol.* **1995**, *177*, 4207.
(279) Zhu, D. Y.; Zhu, Y. Q.; Huang, R. H.; Xiang, Y.; Yang, N.; Lu, H. X.; Li, G. P.; Jin, Q.; Wang, D. C. *Proteins* **2005**, *58*, 764.
(280) Li, J.; Ji, C.; Chen, J.; Yang, Z.; Wang, Y.; Fei, X.; Zheng, M.; Gu, X.; Wen, G.; Xie, Y.; Mao, Y. *Biochem. Biophys. Res. Commun.* **2005**, *337*, 179.
(281) Calafato, S.; Swain, S.; Hughes, S.; Kille, P.; Stürzenbaum, S. R. *Toxicol. Sci.* **2008**, *106*, 384.
(282) Cha, J. S.; Cooksey, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 8915.
(283) Mills, S. D.; Jasalavich, C. A.; Cooksey, D. A. *J. Bacteriol.* **1993**, *175*, 1656.
(284) Teixeira, E. C.; Franco de Oliveira, J. C.; Marques Novo, M. T.; Bertolini, M. C. *Microbiology* **2008**, *154*, 402.
(285) Hussain, F.; Sedlak, E.; Wittung-Stafshede, P. *Arch. Biochem. Biophys.* **2007**, *467*, 58.
(286) Arnesano, F.; Banci, L.; Bertini, I.; Thompsett, A. R. *Structure* **2002**, *10*, 1337.
(287) Cha, J.; Cooksey, D. A. *Appl. Environ. Microbiol.* **1993**, *59*, 1671.
(288) Koay, M.; Zhang, L.; Yang, B.; Maher, M. J.; Xiao, Z.; Wedd, A. G. *Inorg. Chem.* **2005**, *44*, 5203.
(289) Rouch, D. A.; Brown, N. L. *Microbiology* **1997**, *143*, 1191.
(290) Brown, N. L.; Barrett, S. R.; Camakaris, J.; Lee, B. T.; Rouch, D. A. *Mol. Microbiol.* **1995**, *17*, 1153.
(291) Chillappagari, S.; Miethke, M.; Trip, H.; Kuipers, O. P.; Marahiel, M. A. *J. Bacteriol.* **2009**, *191*, 2362.
(292) Lee, S. M.; Grass, G.; Rensing, C.; Barrett, S. R.; Yates, C. J.; Stoyanov, J. V.; Brown, N. L. *Biochem. Biophys. Res. Commun.* **2002**, *295*, 616.
(293) Quaranta, D.; McCarty, R.; Bandarian, V.; Rensing, C. *J. Bacteriol.* **2007**, *189*, 5361.
(294) Petris, M. J. *Pflugers Arch.* **2004**, *447*, 752.
(295) Dancis, A.; Haile, D.; Yuan, D. S.; Klausner, R. D. *J. Biol. Chem.* **1994**, *269*, 25660.
(296) Sharp, P. A. *Int. J. Biochem. Cell Biol.* **2003**, *35*, 288.
(297) Valentine, J. S.; Gralla, E. B. *Science* **1997**, *278*, 817.
(298) Zhou, H.; Cadigan, K. M.; Thiele, D. J. *J. Biol. Chem.* **2003**, *278*, 48210.
(299) Rees, E. M.; Thiele, D. J. *Curr. Opin. Microbiol.* **2004**, *7*, 175.
(300) Guo, Y.; Smith, K.; Lee, J.; Thiele, D. J.; Petris, M. J. *J. Biol. Chem.* **2004**, *279*, 17428.
(301) Jiang, J.; Nadas, I. A.; Kim, M. A.; Franz, K. J. *Inorg. Chem.* **2005**, *44*, 9787.
(302) Lee, J.; Peña, M. M.; Nose, Y.; Thiele, D. J. *J. Biol. Chem.* **2002**, *277*, 4380.
(303) Rae, T. D.; Schmidt, P. J.; Pufahl, R. A.; Culotta, V. C.; O'Halloran, T. V. *Science* **1999**, *284*, 805.
(304) Van Ho, A.; Ward, D. M.; Kaplan, J. *Annu. Rev. Microbiol.* **2002**, *56*, 237.
(305) Pena, M. M.; Puig, S.; Thiele, D. J. *J. Biol. Chem.* **2000**, *275*, 33244.
(306) Rees, E. M.; Lee, J.; Thiele, D. J. *J. Biol. Chem.* **2004**, *279*, 54221.
(307) Zhou, B.; Gitschier, J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7481.
(308) Klomp, A. E.; Tops, B. B.; Van Denberg, I. E.; Berger, R.; Klomp, L. W. *Biochem. J.* **2002**, *364*, 497.
(309) van den Berghe, P. V.; Folmer, D. E.; Malingré, H. E.; van Beurden, E.; Klomp, A. E.; van de Sluis, B.; Merkx, M.; Berger, R.; Klomp, L. W. *Biochem. J.* **2007**, *407*, 49.
(310) Turski, M. L.; Thiele, D. J. *J. Biol. Chem.* **2007**, *282*, 24017.
(311) Selvaraj, A.; Balamurugan, K.; Yepiskoposyan, H.; Zhou, H.; Egli, D.; Georgiev, O.; Thiele, D. J.; Schaffner, W. *Genes Dev.* **2005**, *19*, 891.
(312) Bull, P. C.; Thomas, G. R.; Rommens, J. M.; Forbes, J. R.; Cox, D. W. *Nat. Genet.* **1993**, *5*, 327.
(313) Camakaris, J.; Voskoboinik, I.; Mercer, J. F. *Biochem. Biophys. Res. Commun.* **1999**, *261*, 225.
(314) Prohaska, J. R.; Gybina, A. A. *J. Nutr.* **2004**, *134*, 1003.
(315) Balamurugan, K.; Egli, D.; Hua, H.; Rajaram, R.; Seisenbacher, G.; Georgiev, O.; Schaffner, W. *EMBO J.* **2007**, *26*, 1035.
(316) Linz, R.; Lutsenko, S. *J. Bioenerg. Biomembr.* **2007**, *39*, 403.
(317) Lönnerdal, B. *Am. J. Clin. Nutr.* **2008**, *88*, 846S.
(318) Prohaska, J. R. *Am. J. Clin. Nutr.* **2008**, *88*, 826S.
(319) La Fontaine, S.; Mercer, J. F. *Arch. Biochem. Biophys.* **2007**, *463*, 149.
(320) Lutsenko, S.; Barnes, N. L.; Bartee, M. Y.; Dmitriev, O. Y. *Physiol. Rev.* **2007**, *87*, 1011.
(321) Vulpe, C.; Levinson, B.; Whitney, S.; Packman, S.; Gitschier, J. *Nat. Genet.* **1993**, *3*, 7.
(322) Voskoboinik, I.; Camakaris, J. *J. Bioenerg. Biomembr.* **2002**, *34*, 363.
(323) Voskoboinik, I.; Mar, J.; Strausak, D.; Camakaris, J. *J. Biol. Chem.* **2001**, *276*, 28620.
(324) Balamurugan, K.; Schaffner, W. *Biochim. Biophys. Acta* **2006**, *1763*, 737.
(325) Lutsenko, S.; LeShane, E. S.; Shinde, U. *Arch. Biochem. Biophys.* **2007**, *463*, 134.
(326) Setty, S. R.; Tenza, D.; Sviderskaya, E. V.; Bennett, D. C.; Raposo, G.; Marks, M. S. *Nature* **2008**, *454*, 1142.
(327) Norgate, M.; Lee, E.; Southon, A.; Farlow, A.; Batterham, P.; Camakaris, J.; Burke, R. *Mol. Biol. Cell* **2006**, *17*, 475.
(328) Burke, R.; Commons, E.; Camakaris, J. *Int. J. Biochem. Cell Biol.* **2008**, *40*, 1850.
(329) Southon, A.; Burke, R.; Norgate, M.; Batterham, P.; Camakaris, J. *Biochem. J.* **2004**, *383*, 303.
(330) Fu, D.; Beeler, T. J.; Dunn, T. M. *Yeast* **1995**, *11*, 283.
(331) Singleton, C.; Le Brun, N. E. *Biometals* **2007**, *20*, 275.
(332) Miras, R.; Morin, I.; Jacquin, O.; Cuillel, M.; Guillain, F.; Mintz, E. *J. Biol. Inorg. Chem.* **2008**, *13*, 195.
(333) Tottey, S.; Waldron, K. J.; Firbank, S. J.; Reale, B.; Bessant, C.; Sato, K.; Cheek, T. R.; Gray, J.; Banfield, M. J.; Dennison, C.; Robinson, N. J. *Nature* **2008**, *455*, 1138.
(334) Viles, J. H.; Klewpatinond, M.; Nadal, R. C. *Biochem. Soc. Trans.* **2008**, *36*, 1288.
(335) Kung, C. C.; Huang, W. N.; Huang, Y. C.; Yeh, K. C. *Proteomics* **2006**, *6*, 2746.
(336) Barondeau, D. P.; Getzoff, E. D. *Curr. Opin. Struct. Biol.* **2004**, *14*, 765.
(337) Sakurai, T.; Kataoka, K. *Cell. Mol. Life Sci.* **2007**, *64*, 2642.
(338) MacPherson, I. S.; Murphy, M. E. T. *Cell. Mol. Life Sci.* **2007**, *64*, 2887.
(339) Nakamura, K.; Go, N. *Cell. Mol. Life Sci.* **2005**, *62*, 2050.
(340) Nersissian, A. M.; Shipp, E. L. *Adv. Protein Chem.* **2002**, *60*, 271.
(341) Solomon, E. I.; Szilagyi, R. K.; DeBeer George, S.; Basumallick, L. *Chem. Rev.* **2004**, *104*, 419.
(342) Dennison, C. *Nat. Prod. Rep.* **2008**, *25*, 15.
(343) De Rienzo, F.; Gabdoulline, R. R.; Wade, R. C.; Sola, M.; Menziani, M. C. *Cell. Mol. Life Sci.* **2004**, *61*, 1123.
(344) Adman, E. T. *Adv. Protein Chem.* **1991**, *42*, 145.
(345) Gray, H. B.; Malmström, B. G.; Williams, R. J. *J. Biol. Inorg. Chem.* **2000**, *5*, 551.
(346) Redinbo, M. R.; Yeates, T. O.; Merchant, S. *J. Bioenerg. Biomembr.* **1994**, *26*, 49.
(347) Marcaida, M. J.; Schlarb-Ridley, B. G.; Worrall, J. A.; Wastl, J.; Evans, T. J.; Bendall, D. S.; Luisi, B. F.; Howe, C. J. *J. Mol. Biol.* **2006**, *360*, 968.
(348) Arciero, D. M.; Pierce, B. S.; Hendrich, M. P.; Hooper, A. B. *Biochemistry* **2002**, *41*, 1703.
(349) Basumallick, L.; Sarangi, R.; DeBeer George, S.; Elmore, B.; Hooper, A. B.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **2005**, *127*, 3531.
(350) Sakurai, T.; Kataoka, K. *Chem. Rec.* **2007**, *7*, 220.
(351) Claus, H. *Micron.* **2004**, *35*, 93.
(352) Hellman, N. E.; Gitlin, J. D. *Annu. Rev. Nutr.* **2002**, *22*, 439.
(353) Bento, I.; Carrondo, M. A.; Lindley, P. F. *J. Biol. Inorg. Chem.* **2006**, *11*, 539.
(354) Pelmenschikov, V.; Siegbahn, P. E. *Inorg. Chem.* **2005**, *44*, 3311.
(355) Brazeau, B. J.; Johnson, B. J.; Wilmot, C. M. *Arch. Biochem. Biophys.* **2004**, *428*, 22.
(356) Rea, G.; Metoui, O.; Infantino, A.; Federico, R.; Angelini, R. *Plant Physiol.* **2002**, *128*, 865.
(357) Pietrangeli, P.; Mondovi, B. *Neurotoxicology* **2004**, *25*, 317.
(358) Boomsma, F.; Hut, H.; Bagghoe, U.; van der Houwen, A.; van den Meiracker, A. *Med. Sci. Monit.* **2005**, *11*, 122.
(359) Parsons, M. R.; Convery, M. A.; Wilmot, C. M.; Yadav, K. D.; Blakeley, V.; Corner, A. S.; Phillips, S. E.; McPherson, M. J.; Knowles, P. F. *Structure* **1995**, *3*, 1171.
(360) Airenne, T. T.; Nymalm, Y.; Kidron, H.; Smith, D. J.; Pihlavisto, M.; Salmi, M.; Jalkanen, S.; Johnson, M. S.; Salminen, T. A. *Protein Sci.* **2005**, *14*, 1964.
(361) Prigge, S. T.; Mains, R. E.; Eipper, B. A.; Amzel, L. M. *Cell. Mol. Life Sci.* **2000**, *57*, 1236.
(362) Klinman, J. P. *J. Biol. Chem.* **2006**, *281*, 3013.
(363) Zumft, W. G. *J. Mol. Microbiol. Biotechnol.* **2005**, *10*, 154.
(364) Moser, C. C.; Page, C. C.; Dutton, P. L. *Philos. Trans. R. Soc. London B: Biol. Sci.* **2006**, *361*, 1295.
(365) Beinert, H. *Eur. J. Biochem.* **1997**, *245*, 521.
(366) Saraste, M. *Antonie Van Leeuwenhoek* **1994**, *65*, 285.
(367) Musser, S. M.; Chan, S. I. *J. Mol. Evol.* **1998**, *46*, 508.
(368) Qin, L.; Hiser, C.; Mulichak, A.; Garavito, R. M.; Ferguson-Miller, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16117.

(369) Ma, J.; Katsonouri, A.; Gennis, R. B. *Biochemistry* **1997**, *36*, 11298.
(370) Shapleigh, J. P.; Gennis, R. B. *Mol. Microbiol.* **1992**, *6*, 635.
(371) Michel, H. *Biochemistry* **1999**, *38*, 15129.
(372) Farver, O.; Chen, Y.; Fee, J. A.; Pecht, I. *FEBS Lett.* **2006**, *580*, 3417.
(373) Pitcher, R. S.; Watmough, N. J. *Biochim. Biophys. Acta* **2004**, *1655*, 388.
(374) Contreras-Zentella, M.; Mendoza, G.; Membrillo-Hernández, J.; Escamilla, J. E. *J. Biol. Chem.* **2003**, *278*, 31473.
(375) Brown, K.; Tegoni, M.; Prudêncio, M.; Pereira, A. S.; Besson, S.; Moura, J. J.; Moura, I.; Cambillau, C. *Nat. Struct. Biol.* **2000**, *7*, 191.
(376) Rodríguez-Montelongo, L.; Volentini, S. I.; Farías, R. N.; Massa, E. M.; Rapisarda, V. A. *Arch. Biochem. Biophys.* **2006**, *451*, 1.
(377) Claus, H.; Decaer, H. *Syst. Appl. Microbiol.* **2006**, *29*, 3.
(378) Decaer, H.; Schweikardt, T.; Nillius, D.; Salzbrunn, U.; Jaenicke, E.; Tuczek, F. *Gene* **2007**, *398*, 183.
(379) Jaenicke, E.; Decker, H. *ChemBioChem* **2004**, *5*, 163.
(380) Chan, S. I.; Chen, K. H.; Yu, S. S.; Chen, C. L.; Kuo, S. S. *Biochemistry* **2004**, *43*, 4421.
(381) Lieberman, R. L.; Rosenzweig, A. C. *Nature* **2005**, *434*, 177.
(382) Balasubramanian, R.; Rosenzweig, A. C. *Acc. Chem. Res.* **2007**, *40*, 573.
(383) Kuper, J.; Llamas, A.; Hecht, H. J.; Mendel, R. R.; Schwarz, G. *Nature* **2004**, *430*, 803.
(384) Whittaker, M. M.; Whittaker, J. W. *J. Biol. Chem.* **2003**, *278*, 22090.
(385) Firbank, S.; Rogers, M.; Guerrero, R. H.; Dooley, D. M.; Halcrow, M. A.; Phillips, S. E.; Knowles, P. F.; McPherson, M. J. *Biochem. Soc. Symp.* **2004**, *71*, 15.
(386) Andreini, C.; Bertini, I.; Rosato, A. *Bioinformatics* **2004**, *20*, 1373.
(387) Hemp, J.; Han, H.; Roh, J. H.; Kaplan, S.; Martinez, T. J.; Gennis, R. B. *Biochemistry* **2007**, *46*, 9963.
(388) Ellis, M. J.; Grossmann, J. G.; Eady, R. R.; Hasnain, S. S. *J. Biol. Inorg. Chem.* **2007**, *12*, 1119.
(389) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. *J. Proteome Res.* **2006**, *5*, 196.
(390) Andreini, C.; Banci, L.; Bertini, I.; Elmi, S.; Rosato, A. *Proteins* **2007**, *67*, 317.
(391) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. *J. Proteome Res.* **2008**, *7*, 209.
(392) Oldfield, J. E. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 1.
(393) Thomson, C. D. *Eur. J. Clin. Nutr.* **2004**, *58*, 391.
(394) Rayman, M. P. *Lancet* **2000**, *356*, 233.
(395) Rayman, M. P. *Proc. Nutr. Soc.* **2005**, *64*, 527.
(396) Boosalis, M. G. *Nutr. Clin. Pract.* **2008**, *23*, 152.
(397) Beck, M. A.; Matthews, C. C. *Proc. Nutr. Soc.* **2000**, *59*, 581.
(398) Alissa, E. M.; Bahijri, S. M.; Ferns, G. A. *Med. Sci. Monit.* **2003**, *9*, RA9.
(399) Ryan-Harshman, M.; Aldoori, W. *Can. J. Diet. Pract. Res.* **2005**, *66*, 98.
(400) Hatfield, D. L.; Carlson, B. A.; Xu, X. M.; Mix, H.; Gladyshev, V. N. *Prog. Nucleic Acid Res. Mol. Biol.* **2006**, *81*, 97.
(401) Flohé, L. *Biol. Chem.* **2007**, *388*, 987.
(402) Hoffmann, P. R. *Arch. Immunol. Ther. Exp. (Warsz).* **2007**, *55*, 289.
(403) Espinoza, S. E.; Guo, H.; Fedarko, N.; DeZern, A.; Fried, L. P.; Xue, Q. L.; Leng, S.; Beamer, B.; Walston, J. D. *J. Gerontol. A: Biol. Sci. Med. Sci.* **2008**, *63*, 505.
(404) Böck, A.; Rother, M. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 10.
(405) Springer, S. E.; Huber, R. E. *Arch. Biochem. Biophys.* **1973**, *156*, 595.
(406) Lindblow-Kull, C.; Kull, F. J.; Shrift, A. *J. Bacteriol.* **1985**, *163*, 1267.
(407) Li, H. F.; McGrath, S. P.; Zhao, F. J. *New Phytol.* **2008**, *178*, 92.
(408) Papp, L. V.; Lu, J.; Holmgren, A.; Khanna, K. K. *Antioxid. Redox Signal.* **2007**, *9*, 775.
(409) Müller, S.; Heider, J.; Böck, A. *Arch. Microbiol.* **1997**, *168*, 421.
(410) Böck, A. *Biofactors* **2000**, *11*, 77.
(411) Böck, A.; Forchhammer, K.; Heider, J.; Leinfelder, W.; Sawers, G.; Veprek, B.; Zinoni, F. *Mol. Microbiol.* **1991**, *5*, 515.
(412) Stadtman, T. C. *Annu. Rev. Biochem.* **1996**, *65*, 83.
(413) Hatfield, D. L.; Gladyshev, V. N. *Mol. Cell. Biol.* **2002**, *22*, 3565.
(414) Copeland, P. R. *Genome Biol.* **2005**, *6*, 221.
(415) Tate, W. P.; Mansell, J. B.; Mannering, S. A.; Irvine, J. H.; Major, L. L.; Wilson, D. N. *Biochemistry (Moscow)* **1999**, *64*, 1342.
(416) Allmang, C.; Krol, A. *Biochimie* **2006**, *88*, 1561.
(417) Squires, J. E.; Berry, M. J. *IUBMB Life* **2008**, *60*, 232.
(418) Rother, M.; Resch, A.; Wilting, R.; Böck, A. *Biofactors* **2001**, *14*, 75.
(419) Driscoll, D. M.; Copeland, P. R. *Annu. Rev. Nutr.* **2003**, *23*, 17.
(420) Low, S.; Berry, M. J. *Trends Biochem. Sci.* **1996**, *21*, 203.
(421) Yuan, J.; Palioura, S.; Salazar, J. C.; Su, D.; O'Donoghue, P.; Hohn, M. J.; Cardoso, A. M.; Whitman, W. B.; Söll, D. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18923.
(422) Xu, X. M.; Carlson, B. A.; Mix, H.; Zhang, Y.; Saira, K.; Glass, R. S.; Berry, M. J.; Gladyshev, V. N.; Hatfield, D. L. *PloS Biol.* **2007**, *5*, e4.
(423) Hendrickson, T. L. *Nat. Struct. Mol. Biol.* **2007**, *14*, 100.
(424) Ehrenreich, A.; Forchhammer, K.; Tormay, P.; Veprek, B.; Böck, A. *Eur. J. Biochem.* **1992**, *206*, 767.
(425) Thanbichler, M.; Böck, A. *Methods Enzymol.* **2002**, *347*, 3.
(426) Zinoni, F.; Heider, J.; Böck, A. *Proc. Natl. Acad. Sci.* **1990**, *87*, 4660.
(427) Small-Howard, A. L.; Berry, M. J. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 83.
(428) Xu, X. M.; Carlson, B. A.; Irons, R.; Mix, H.; Zhong, N.; Gladyshev, V. N.; Hatfield, D. L. *Biochem. J.* **2007**, *404*, 115.
(429) Small-Howard, A.; Morozova, N.; Stoytcheva, Z.; Forry, E. P.; Mansell, J. B.; Harney, J. W.; Carlson, B. A.; Xu, X. M.; Hatfield, D. L.; Berry, M. J. *Mol. Cell. Biol.* **2006**, *26*, 2337.
(430) Berry, M. J.; Tujebajeva, R. M.; Copeland, P. R.; Xu, X. M.; Carlson, B. A.; Martin, G. W., 3rd.; Low, S. C.; Mansell, J. B.; Grundner-Culemann, E.; Harney, J. W.; Driscoll, D. M.; Hatfield, D. L. *Biofactors* **2001**, *14*, 17.
(431) Hoffmann, P. R.; Berry, M. J. *Thyroid* **2005**, *15*, 769.
(432) Copeland, P. R.; Driscoll, D. M. *J. Biol. Chem.* **1999**, *274*, 25447.
(433) Chavatte, L.; Brown, B. A.; Driscoll, D. M. *Nat. Struct. Mol. Biol.* **2005**, *12*, 408.
(434) Driscoll, D. M. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 63.
(435) Behne, D.; Kyriakopoulos, A.; Meinhold, H.; Köhrle, J. *Biochem. Biophys. Res. Commun.* **1990**, *173*, 1143.
(436) Tamura, T.; Stadtman, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 1006.
(437) Gladyshev, V. N.; Jeang, K. T.; Wootton, J. C.; Hatfield, D. L. *J. Biol. Chem.* **1998**, *273*, 8910.
(438) Walczak, R.; Westhof, E.; Carbon, P.; Krol, A. *RNA* **1996**, *2*, 367.
(439) Grundner-Culemann, E.; Martin, G. W., 3rd.; Harney, J. W.; Berry, M. J. *RNA* **1999**, *5*, 625.
(440) Korotkov, K. V.; Novoselov, S. V.; Hatfield, D. L.; Gladyshev, V. N. *Mol. Cell. Biol.* **2002**, *22*, 1402.
(441) Wilting, R.; Schorling, S.; Persson, B. C.; Böck, A. *J. Mol. Biol.* **1997**, *266*, 637.
(442) Berg, B. L.; Baron, C.; Stewart, V. *J. Biol. Chem.* **1991**, *266*, 22386.
(443) Heider, J.; Baron, C.; Böck, A. *EMBO J.* **1992**, *11*, 3759.
(444) Hüttenhofer, A.; Westhof, E.; Böck, A. *RNA* **1996**, *2*, 354.
(445) Chen, G. F.; Fang, L.; Inouye, M. *J. Biol. Chem.* **1993**, *268*, 23128.
(446) Heider, J.; Forchhammer, K.; Sawers, G.; Böck, A. *Arch. Microbiol.* **1991**, *155*, 221.
(447) Gursinsky, T.; Jäger, J.; Andreesen, J. R.; Söhling, B. *Arch. Microbiol.* **2000**, *174*, 200.
(448) Zhang, Y.; Gladyshev, V. N. *Bioinformatics* **2005**, *21*, 2580.
(449) Kryukov, G. V.; Kryukov, V. M.; Gladyshev, V. N. *J. Biol. Chem.* **1999**, *274*, 33888.
(450) Lescure, A.; Gautheret, D.; Carbon, P.; Krol, A. *J. Biol. Chem.* **1999**, *274*, 38147.
(451) Backofen, R.; Narayanaswamy, N. S.; Swidan, F. *In Silico Biol.* **2002**, *2*, 275.
(452) Lobanov, A. V.; Gromer, S.; Salinas, G.; Gladyshev, V. N. *Nucleic Acids Res.* **2006**, *34*, 4012.
(453) Lobanov, A. V.; Delgado, C.; Rahlfs, S.; Novoselov, S. V.; Kryukov, G. V.; Gromer, S.; Hatfield, D. L.; Becker, K.; Gladyshev, V. N. *Nucleic Acids Res.* **2006**, *34*, 496.
(454) Castellano, S.; Lobanov, A. V.; Chapple, C.; Novoselov, S. V.; Albrecht, M.; Hua, D.; Lescure, A.; Lengauer, T.; Krol, A.; Gladyshev, V. N.; Guigó, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16188.
(455) Lobanov, A. V.; Fomenko, D. E.; Zhang, Y.; Sengupta, A.; Hatfield, D. L.; Gladyshev, V. N. *Genome Biol.* **2007**, *8*, R198.
(456) Zhang, Y.; Fomenko, D. E.; Gladyshev, V. N. *Genome Biol.* **2005**, *6*, R37.
(457) Zhang, Y.; Gladyshev, V. N. *PLoS Genet.* **2008**, *4*, e1000095.
(458) Fomenko, D. E.; Xing, W.; Adair, B. M.; Thomas, D. J.; Gladyshev, V. N. *Science* **2007**, *315*, 387.
(459) Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; *Selenium: Its molecular biology and role in human health*, 2nd edition; Springer: New York, 2006.
(460) Rotruck, J. T.; Pope, A. L.; Ganther, H. E.; Swanson, A. B.; Hafeman, D. G.; Hoekstra, W. G. *Science* **1973**, *179*, 588.
(461) Borchert, A.; Wang, C. C.; Ufer, C.; Schiebel, H.; Savaskan, N. E.; Kuhn, H. *J. Biol. Chem.* **2006**, *281*, 19655.

(462) Ursini, F.; Heim, S.; Kiess, M.; Maiorino, M.; Roveri, A.; Wissing, J.; Flohé, L. *Science* **1999**, *285*, 1393.
(463) Köhrle, J. *Thyroid* **2005**, *15*, 841.
(464) Bianco, A. C.; Larsen, P. R. *Thyroid* **2005**, *15*, 777.
(465) Bianco, A. C.; Larsen, P. R. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 207.
(466) Holmgren, A. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 183.
(467) Conrad, M.; Bornkamm, G. W. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 195.
(468) Sun, Q. A.; Kirnarsky, L.; Sherman, S.; Gladyshev, V. N. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3673.
(469) Sandalova, T.; Zhong, L.; Lindqvist, Y.; Holmgren, A.; Schneider, G. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9533.
(470) Sun, Q. A.; Su, D.; Novoselov, S. V.; Carlson, B. A.; Hatfield, D. L.; Gladyshev, V. N. *Biochemistry* **2005**, *44*, 14528.
(471) Su, D.; Novoselov, S. V.; Sun, Q. A.; Moustafa, M. E.; Zhou, Y.; Oko, R.; Hatfield, D. L.; Gladyshev, V. N. *J. Biol. Chem.* **2005**, *280*, 26491.
(472) Biterova, E. I.; Turanov, A. A.; Gladyshev, V. N.; Barycki, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15018.
(473) Turanov, A. A.; Su, D.; Gladyshev, V. N. *J. Biol. Chem.* **2006**, *281*, 22953.
(474) Burk, R. F.; Hill, K. E. *Annu. Rev. Nutr.* **2005**, *25*, 215.
(475) Burk, R. F.; Olson, G. E.; Hill, K. E. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 111.
(476) Schweizer, U. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 233.
(477) Kryukov, G. V.; Kumar, R. A.; Koc, A.; Sun, Z.; Gladyshev, V. N. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4245.
(478) Kim, H. Y.; Gladyshev, V. N. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 123.
(479) Kim, H. Y.; Gladyshev, V. N. *PLoS Biol.* **2005**, *3*, e375.
(480) Labunskyy, V. M.; Hatfield, D. L.; Gladyshev, V. N. *IUBMB Life* **2007**, *59*, 1.
(481) Labunskyy, V. M.; Gladyshev, V. N. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 141.
(482) Korotkov, K. V.; Kumaraswamy, E.; Zhou, Y.; Hatfield, D. L.; Gladyshev, V. N. *J. Biol. Chem.* **2001**, *276*, 15330.
(483) Wu, J. C.; Liang, Z. Q.; Qin, Z. H. *Acta. Biochim. Biophys. Sin (Shanghai)* **2006**, *38*, 219.
(484) Ferguson, A. D.; Labunskyy, V. M.; Fomenko, D. E.; Araç, D.; Chelliah, Y.; Amezcua, C. A.; Rizo, J.; Gladyshev, V. N.; Deisenhofer, J. *J. Biol. Chem.* **2006**, *281*, 3536.
(485) Diwadkar-Navsariwala, V.; Diamond, A. M. *J. Nutr.* **2004**, *134*, 2899.
(486) Vendeland, S. C.; Beilstein, M. A.; Chen, C. L.; Jensen, O. N.; Barofsky, E.; Whanger, P. D. *J. Biol. Chem.* **1993**, *268*, 17103.
(487) Whanger, P. D. *Cell. Mol. Life Sci.* **2000**, *57*, 1846.
(488) Aachmann, F. L.; Fomenko, D. E.; Soragni, A.; Gladyshev, V. N.; Dikiy, A. *J. Biol. Chem.* **2007**, *282*, 37036.
(489) Kioussi, C. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 135.
(490) Jurynec, M. J.; Xia, R.; Mackrill, J. J.; Gunther, D.; Crawford, T.; Flanigan, K. M.; Abramson, J. J.; Howard, M. T.; Grunwald, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 12485.
(491) Zeng, J.; Du, S.; Zhou, J.; Huang, K. *Arch. Biochem. Biophys.* **2008**, *478*, 1.
(492) Gladyshev, V. N. In *Selenium: Its molecular biology and role in human health*, 2nd ed.; Hatfield, D. L., Berry, M. J., Gladyshev, V. N., Eds.; Springer: New York, 2006; p 99.
(493) Zhang, Y.; Romero, H.; Salinas, G.; Gladyshev, V. N. *Genome Biol.* **2006**, *7*, R94.
(494) Romero, H.; Zhang, Y.; Gladyshev, V. N.; Salinas, G. *Genome Biol.* **2005**, *6*, R66.
(495) Gladyshev, V. N.; Khangulov, S. V.; Axley, M. J.; Stadtman, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7708.
(496) Boyington, J. C.; Gladyshev, V. N.; Khangulov, S. V.; Stadtman, T. C.; Sun, P. D. *Science* **1997**, *275*, 1305.
(497) Khangulov, S. V.; Gladyshev, V. N.; Dismukes, G. C.; Stadtman, T. C. *Biochemistry* **1998**, *37*, 3518.
(498) Garcin, E.; Vernede, X.; Hatchikian, E. C.; Volbeda, A.; Frey, M.; Fontecilla-Camps, J. C. *Structure* **1999**, *7*, 557.
(499) Wilting, R.; Schorling, S.; Persson, B. C.; Böck, A. *J. Mol. Biol.* **1997**, *266*, 637.
(500) Sliwkowski, M. X.; Stadtman, T. C. *J. Biol. Chem.* **1987**, *262*, 4899.
(501) Kreimer, S.; Andreesen, J. R. *Eur. J. Biochem.* **1995**, *234*, 192.
(502) Söhling, B.; Parther, T.; Rücknagel, K. P.; Wagner, M. A.; Andreesen, J. R. *Biol. Chem.* **2001**, *382*, 979.
(503) Zhang, Y.; Gladyshev, V. N. *Nucleic Acids Res.* **2007**, *35*, 4952.
(504) Veres, Z.; Tsai, L.; Scholz, T. D.; Politino, M.; Balaban, R. S.; Stadtman, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2975.
(505) Kramer, G. F.; Ames, B. N. *J. Bacteriol.* **1988**, *170*, 736.
(506) Wolfe, M. D.; Ahmed, F.; Lacourciere, G. M.; Lauhon, C. T.; Stadtman, T. C.; Larson, T. J. *J. Biol. Chem.* **2004**, *279*, 1801.
(507) Schräder, T.; Rienhöfer, A.; Andreesen, J. R. *Eur. J. Biochem.* **1999**, *264*, 862.
(508) Gladyshev, V. N.; Khangulov, S. V.; Stadtman, T. C. *Biochemistry* **1996**, *35*, 212.
(509) Self, W. T.; Wolfe, M. D.; Stadtman, T. C. *Biochemistry* **2003**, *42*, 11382.
(510) Alhapel, A.; Darley, D. J.; Wagener, N.; Eckel, E.; Elsner, N.; Pierik, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12341.
(511) Haft, D. H.; Self, W. T. *Biol. Direct* **2008**, *3*, 4.
(512) Zhang, Y.; Turanov, A. A.; Hatfield, D. L.; Gladyshev, V. N. *BMC Genomics* **2008**, *9*, 251.
(513) Martin-Romero, F. J.; Kryukov, G. V.; Lobanov, A. V.; Carlson, B. A.; Lee, B. J.; Gladyshev, V. N.; Hatfield, D. L. *J. Biol. Chem.* **2001**, *276*, 29798.
(514) Gladyshev, V. N.; Krause, M.; Xu, X. M.; Korotkov, K. V.; Kryukov, G. V.; Sun, Q. A.; Lee, B. J.; Wootton, J. C.; Hatfield, D. L. *Biochem. Biophys. Res. Commun.* **1999**, *259*, 244.
(515) Taskov, K.; Chapple, C.; Kryukov, G. V.; Castellano, S.; Lobanov, A. V.; Korotkov, K. V.; Guigó, R.; Gladyshev, V. N. *Nucleic Acids Res.* **2005**, *33*, 2227.
(516) Novoselov, S. V.; Rao, M.; Onoshko, N. V.; Zhi, H.; Kryukov, G. V.; Xiang, Y.; Weeks, D. P.; Hatfield, D. L.; Gladyshev, V. N. *EMBO J.* **2002**, *21*, 3681.
(517) Lobanov, A. V.; Hatfield, D. L.; Gladyshev, V. N. *Genome Biol.* **2008**, *9*, R62.
(518) Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. J. *Biol. Inorg. Chem.* **2008**, *13*, 1205.
(519) Morrison, M. S.; Cobine, P. A.; Hegg, E. L. *J. Biol. Inorg. Chem.* **2007**, *12*, 1129.
(520) Dupont, C. L.; Yang, S.; Palenik, B.; Bourne, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17822.
(521) Koonin, E. V.; Wolf, Y. I. *Nucleic Acids Res.* **2008**, *36*, 6688.

CR800557S