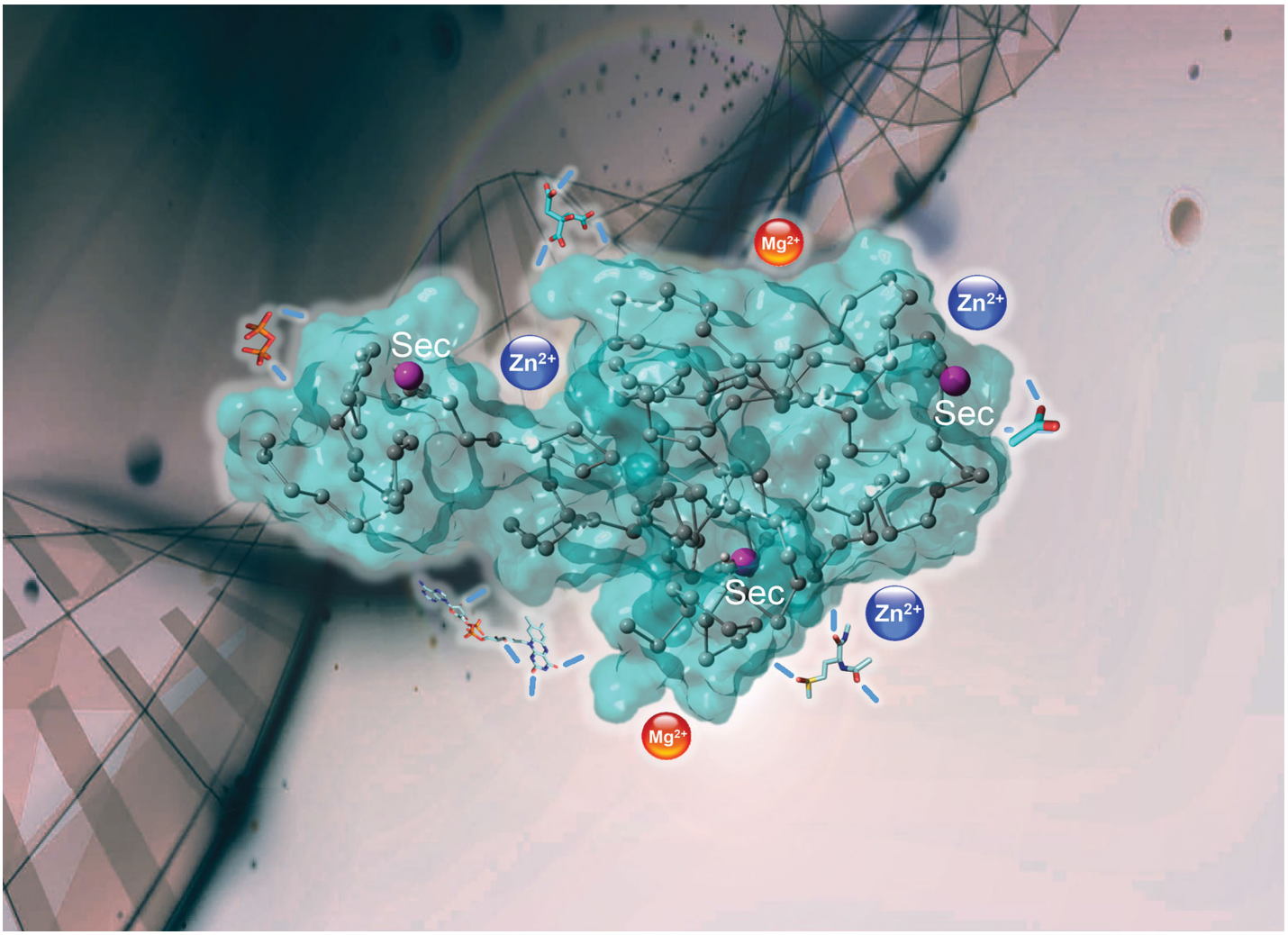

Showcasing research into avian selenoproteomes by Jin-Long Li and colleagues from the College of Veterinary Medicine, Northeast Agricultural University, Harbin, People's Republic of China

Biochemical characterization of the selenoproteome in *Gallus gallus via* bioinformatics analysis: structure–function relationships and interactions of binding molecules

This study addressed questions regarding the role of selenoproteins in chickens in relation to their biological functions. We used genomic and amino acid sequences and various other datasets for bird selenoproteins along with bioinformatics tools to comprehensively analyze the structural and biochemical functions of the selenoproteome. Our findings will draw attention to the novel biochemical functions of selenoproteins.

As featured in:

See Jin-Long Li *et al.*, *Metallomics*, 2017, **9**, 124.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/metallomics

Registered charity number: 207890

Metallomics

PAPER

View Article Online
View Journal | View Issue

Cite this: *Metallomics*, 2017, 9, 124

Biochemical characterization of the selenoproteome in *Gallus gallus via* bioinformatics analysis: structure–function relationships and interactions of binding molecules†

Shi-Yong Zhu,‡ Xue-Nan Li,‡ Xiao-Chen Sun,‡ Jia Lin, Wei Li, Cong Zhang and Jin-Long Li*

Knowledge about mammalian selenoproteins is increasing. However, the selenoproteome of birds remains considerably less understood, especially concerning its biochemical characterization, structure–function relationships and the interactions of binding molecules. In this work, the SECIS elements, subcellular localization, protein domains and interactions of binding molecules of the selenoproteome in *Gallus gallus* were analyzed using bioinformatics tools. We carried out comprehensive analyses of the structure–function relationships and interactions of the binding molecules of selenoproteins, to provide biochemical characterization of the selenoproteome in *Gallus gallus*. Our data provided a wealth of information on the biochemical functions of bird selenoproteins. Members of the selenoproteome were found to be involved in various biological processes in chickens, such as in antioxidants, maintenance of the redox balance, Se transport, and interactions with metals. Six membrane-bound selenoproteins (SelI, SelK, SelS, SelT, DIO1 and DIO3) played important roles in maintaining the membrane integrity. Chicken selenoproteins were classified according to their ligand binding sites as zinc-containing matrix metalloselenoproteins (Sep15, MsrB1, SelW and SelM), POP-containing selenoproteins (GPx1–4), FAD-interacting selenoproteins (TrxR1–3), secretory transport selenoproteins (GPx3 and SelPa) and other selenoproteins. The results of our study provided new evidence for the unknown biological functions of the selenoproteome in birds. Future research is required to confirm the novel biochemical functions of bird selenoproteins.

Received 1st November 2016,
Accepted 3rd January 2017

DOI: 10.1039/c6mt00254d

www.rsc.org/metallomics

Significance to metallomics

Selenium (Se) participates in various physiological processes in chickens. Selenoproteins play a key role in the biochemical function of Se. However, the chicken selenoproteome remains considerably less understood. This paper provides biochemical characterization of the selenoproteome in *Gallus gallus* through comprehensive analyses of the structure–function relationship and interactions of the binding molecules of selenoproteins. We find that the chicken selenoproteome could be classified into some families according to their binding molecules, including zinc-containing matrix metalloselenoproteins, POP-containing selenoproteins, FAD-interacted selenoproteins and secreted transport selenoproteins. Our findings will draw a wide range of attention to the novel biochemical functions of selenoproteins.

College of Veterinary Medicine, Northeast Agricultural University, Harbin, 150030, People's Republic of China. E-mail: Jinlongli@neau.edu.cn; Tel: +86 451 55190407

† Electronic supplementary information (ESI) available: Table S1. The abbreviations and NCBI accession no. for the chicken selenoproteome. Table S2. The position of Sec, the ligand name and the binding residues of the chicken selenoproteome. Fig. S1. The prediction of SECIS elements in the chicken selenoproteome. Fig. S2. The comparison of amino acid sequences between the human and chicken SelN and the domain prediction in human SelN. Fig. S3. The signal peptide prediction in the sequence of chicken GPx3. The following information was supplied regarding data availability: the raw analytical data in the chicken selenoproteome have been supplied as Data S1. See DOI: 10.1039/c6mt00254d

‡ These authors contributed equally to this work.

Introduction

Selenium (Se) is an obligatory micronutrient that is critical to the normal physiology of many species, including birds.[1,2] Chickens are sensitive to the amount of dietary Se. Dietary Se can be particularly active in maintaining the production performance and affects the breeding of birds. Se deficiency is manifested as hepatic malnutrition, deformation and necrosis in chickens.[3,4] However, an excess of Se can result in reduced growth and anemia, reduced egg production and an impaired immune function. Most biological activities of Se are predominantly

mediated through selenoprotein activity. The function of Se regulation in the body is to maintain vital selenoproteins and to avoid toxicity, and these biological effects are mainly reflected in the direct incorporation of Se into selenoproteins as the 21st amino acid, selenocysteine (Sec). The synthesis of Sec and its insertion into polypeptides requires a complex machinery that recodes the UGA codon, normally a termination codon, to serve as a codon for Sec.[5,6]

Selenoproteins have been found to be involved in various biological processes in chickens, including growth performance, central nervous system function,[5,7] male fertility,[8] and muscle development.[9] The Se protein family also has various oxidoreductase functions, acting as the first line of defense against oxidants.[1] The selenoproteome is a set of selenoproteins in prokaryotes and eukaryotes; 25 selenoproteins have been identified in humans.[10] The selenoproteome in *Gallus gallus* consists of 24 selenoproteins. Although selenoproteins related to antioxidative function are homologous between humans and birds, the rest of the selenoproteins have different functions between humans and birds. Because they serve to alleviate damage caused by ROS or as scavengers, several selenoproteins have been characterized as antioxidant enzymes. These antioxidant enzymes containing Se are called selenoenzymes; the significant metabolic roles of Se in the cell are attributed to its function in the active site of these enzymes, such as glutathione peroxidase (GPx) and thioredoxin reductase (TrxR). However, the biochemical functions of most selenoproteins in birds remain uncharacterized.

Although knowledge about mammalian selenoproteins is increasing, the bird selenoproteome remains considerably less understood. Additionally, the selenoproteome is related to preventing diseases in chickens during poultry production. A supplement of Se can promote the synthesis of selenoproteins. Selenoproteins prevent poor growth and exudative diathesis in poultry production. Pancreatic atrophy is a common disease in chickens prevented by selenoproteins. A better understanding of the role of selenoproteins in *Gallus gallus* would be helpful for poultry production. The purpose of this study was to address questions regarding the role of the selenoproteome in chickens in relation to their biological functions. We used genomic and amino acid sequences and various other datasets for bird selenoproteins using bioinformatics tools to comprehensively analyze the structural and biochemical functions of the selenoproteome. These results will help characterize the biochemical function, structure–function relationships and interactions of binding molecules of the selenoproteome in birds.

Experimental

The genomic sequences and resources of the selenoproteome

All sequences used in this study were from the current Entrez Genome Project at NCBI (http://www.ncbi.nlm.nih.gov/, NCBI). Information including abbreviations and accession numbers for the chicken selenoproteome is provided in Table S1, ESI.†

Analysis of the SECIS element in the selenoproteome

The characteristic feature of the selenoprotein mRNA is UGA, a typical termination codon that can be translated into Sec with SECIS located in the untranslated regions. The SECISearch engine (http://seblastian.crg.es/, SECISearch3) was used to confirm these novel RNA sequences and identify SECIS elements and secondary structures of selenoproteins in *Gallus gallus* (Table S1, ESI†). The SECISearch3 combines the predictions from three sources, including program Infernal, Covels and original SECISearch. Then, the RNAfold package is used to refine the prediction.[11]

Analysis of the subcellular localization of the selenoproteome

To understand the functions of proteins, it is essential to know their subcellular locations. The subcellular localization of the chicken selenoproteins was predicted using Euk-mPLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/, Euk-mPLoc 2.0), PSORT II (http://psort.hgc.jp/form2.html, PSORT II) and TargetP (http://www.cbs.dtu.dk/services/TargetP/, TargetP). The former is established by gene ontology and a pseudo amino acid composition approach.[12] And PSORT II is based on the SWISS-PROT data. The program TargetP is based on signal peptides, mitochondrial targeting peptides and chloroplast transit peptides (in plants), then outputs the actual prediction.[13]

The predictions of transmembrane regions in chicken selenoproteins were achieved by TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/, TMHMM 2.0) and CCTOP (http://cctop.enzim.ttk.mta.hu/, CCTOP).

The predictions of the subcellular localization of the chicken selenoproteins from three online services and the transmembrane regions were considered to obtain the subcellular localization of 24 selenoproteins. And the known subcellular localization of the human selenoproteome and some predicted subcellular localizations of human selenoproteins were used as a positive control to improve the reliability of the predictions for subcellular localization.

Structural modeling in the chicken selenoproteome

I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/, I-TASSER) was used to model the selenoproteome structure. The I-TASSER server is based on iterative threading assembly simulations and matching structural predictions with known functional templates to achieve protein structure modelling. Then, it outputs the protein molecules with annotated biological functions.[14,15] Also, SAVES (http://services.mbi.ucla.edu/SAVES/, SAVES) was used to check the databases of I-TASSER in the PDB files. Then the modelled structures were constructed using the program PYMOL. The modeled region included the location of the Sec in 24 chicken selenoproteins.

Structure model of domains in the chicken selenoproteome

A domain, which is a conserved part of the protein sequence, can evolve, function, and exist independently of the rest of the protein chain. The protein domains of chicken selenoproteins

were identified using InterProScan (http://www.ebi.ac.uk/interpro/search/sequence-search/, InterProScan) and CATH-Gene3D v4.1 (http://www.cathdb.info/, CATH/Gene3D v4.1).[16]

Analysis of the ligand binding sites in the selenoproteome

The ligand binding sites of chicken selenoproteins were analyzed using the I-TASSER server, and then the PDB files containing the ligand binding sites were constructed using PYMOL and YASARA.

Results and discussion

The selenoproteome SECIS element

Eukaryotic SECIS elements have two distinct subfamilies: type I and type II. An analysis of the 3′-untranslated regions in the chicken selenoproteome genes using the SECISearch program revealed that 16 selenoprotein SECIS elements were type II and six selenoproteins were type I (Fig. S1, ESI†). However, the chicken SelI and SelPb SECIS elements were not confirmed.

Subcellular localization of the selenoproteome

21 selenoproteins were located inside the cell, with three exceptions: SelPa, SelPb and GPx3 (Fig. 1). The endoplasmic reticulum (ER) contained DIO1–3, Sep15, SelI, SelK, SelS, SelT, SelM and SelN. The mitochondria (Mit) contained nine selenoproteins including GPx1–2, GPx4, TrxR1–3, SelM, SelO and SelU. In addition, seven selenoproteins were located in the cytoplasm including GPx1–2, GPx4, TrxR1–3 and SelW. The nucleus contained GPx4, MsrB1 and SelH. While SelT located in the Golgi apparatus, which was also predicted in mammals,[17] and SelI located in the cytomembrane.

Membrane-bound selenoproteins are inserted into cellular membranes to accomplish their biochemical and biophysical roles. Eight human selenoproteins are predicted to be membrane-bound selenoproteins.[18] Six out of the 24 bird selenoproteins were predicted to be membrane-bound proteins, including SelI, SelK, SelS, SelT, DIO1 and DIO3 (Fig. 2).

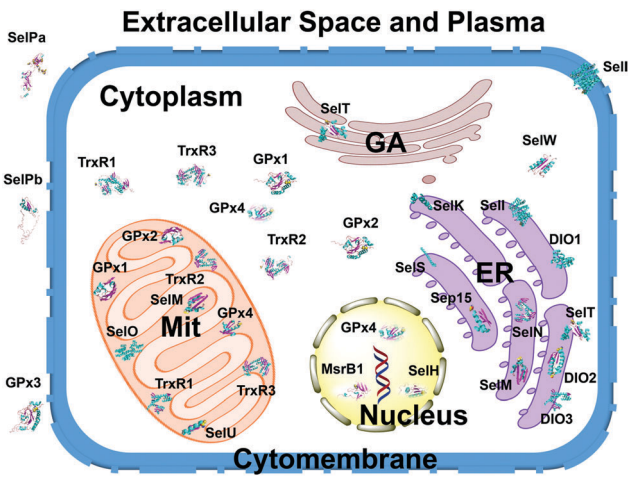

Fig. 1 Subcellular localization of the selenoproteome. The members of the selenoproteome in *Gallus gallus* are expressed at various subcellular localizations related with their functions and regulation. Chicken selenoproteins are depicted with tertiary structures. Mit: mitochondrion; ER: endoplasmic reticulum; GA: Golgi apparatus.

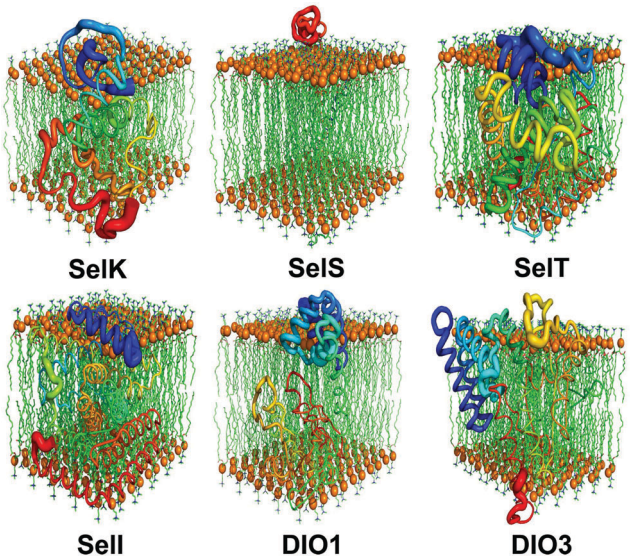

Fig. 2 Transmembrane tertiary structures of membrane-bound selenoproteins in the lipid bilayer. The orange balls and green lines represent the lipid bilayer. The tertiary structure of membrane-bound selenoproteins is depicted with a putty model, created using PYMOL. Membrane-bound selenoproteins span the TM helices through the bilayer multiple times.

Chicken SelI is located in the ER membrane (Fig. 1). This selenoprotein resides entirely in the lipid bilayer. Fig. 2 shows that SelI spans the TM helices and passes through the bilayer multiple times. Recently, structural assays have established that SelI catalyzes the formation of the lipid phosphatidylethanolamine from CDP-ethanolamine and diacylglycerol through its bilayer domain.[18] And phosphatidylethanolamine plays an important role in the organization of bio-membranes by contributing to their shapes.[19] Interestingly, the Sec of SelI is located at the C-terminus, where it is present in the cytosol rather than at the active sites in the lipid bilayer. An evolutionary study noted that SelI had no recognizable homologs in which Sec was substituted for Cys.[6] The EPT-like activity of SelI does not depend on the presence of Sec, which is necessary for selenoprotein function. Therefore, we hypothesized that the physiological role of chicken SelI Sec was related to the composition of the membrane.

Protein domains of the selenoproteome

The function of a protein might be deduced from its domains. Generally, most selenoproteins, including GPx1–4, DIO1–3, TrxR3, SelT, SelH, SelW, Sep15, SelM, SelU and SelO, contain a thioredoxin-like fold (Fig. 3). Interestingly, chicken SelT contained two thioredoxin-like folds and Sec was located in the first thioredoxin-like fold (Fig. 3). The SelT helix was formed at a minimum between residues 143 and 164, and both faces in the helix were hydrophobic. One face interacted with the membrane, while the other may interact with the protein core. This result has been confirmed with human SelT.[18]

Human SelN interacts with Ca^{2+} *via* its EF-hand domain,[20] but chicken SelN did not contain any domain in this study

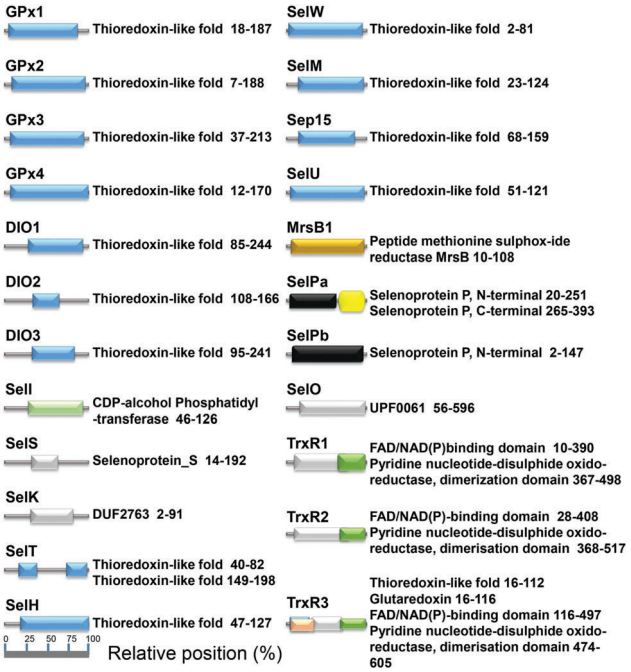

Fig. 3 Protein domains of the selenoproteome. The plotting scale functions as a relative position. The sequences are shown with gray lines. The blue diamond represents the thioredoxin-like fold; the white diamond is the FAD/NAD (P) binding domain; the pyridine nucleotide-disulphide oxidoreductase, dimerization domains are indicated by the green diamonds; the peptide methionine sulfoxide reductase MrsB domain is indicated by the orange diamond and the black diamond represents the selenoprotein P, N-terminal domain, while the selenoprotein P, C-terminal domain is shown by a yellow diamond.

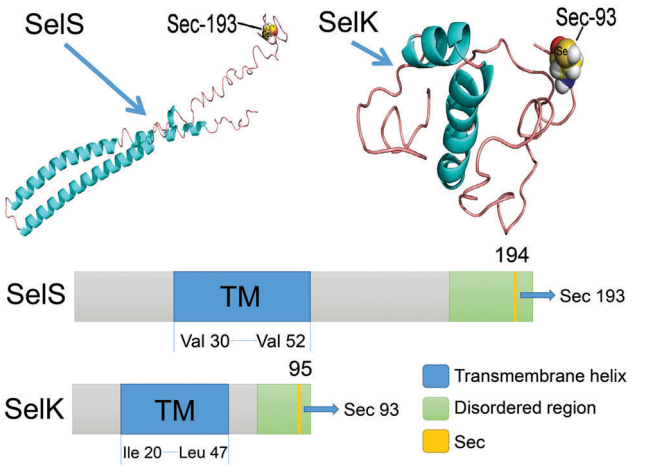

Fig. 4 Domain architecture of the SelS/SelK family in chickens. The gray diamonds represent the main chains of SelS and SelK, the TM helices are depicted by blue diamonds, the green diamonds are the disordered regions, and the Sec is shown by the orange lines. SelS and SelK share several commonalities: they contain less than 300 amino acids, each of them has a sole TM helix, and finally, each of their C-termini has a high isoelectric point, and Sec is usually positioned within the terminal five residues.

(Fig. 3 and Fig. S2, ESI†). The biochemical function of this selenoprotein could be deduced from its structure and the sequence context of the Sec. Chicken SelN was shown to contain a SCUG motif and a catalytic site similar to the GCUG motif in TrxR.[20] The statistical scores of magnitude noted that the domain predicted in SelN has been repeatedly found to be biologically similar to the classical one.[2,21] However, two other functionally important domains, the FAD- and the NADPH-binding domains in classical TrxRs, were missing from SelN. We predicted that chicken SelN interacted with unidentified partners.

Human SelK might function as a sensor of lipid peroxides.[18] Chicken SelK and SelS contained unknown domains but Sec was not present in the domain, but instead it was positioned within the C-terminal five residues (Fig. 4). Chicken SelK contained an unknown domain formed by 2–91 residues. However, Sec was in the disordered region of the C-terminus facing the cytoplasm and did not appear to be in a known redox motif. It was demonstrated that a *Drosophila melanogaster* ortholog of that SelK did not show antioxidant activity.[22] Obviously, Sec plays a role in stabilizing SelK. There was only a single-pass TM helix, consisting of 13 amino acids in SelK, which seemed insufficiently hydrophobic in the construct.[23,24] The presence of a single-pass TM helix leads to the speculation that SelK would likely be present with a protein partner or in a complex. In some selenoproteins, a vicinal Cys forms a dynamic bond with the Sec.[1,8,25] In fact, SelK was present as a homodimer with an intermolecular diselenide bond.[25] Accordingly, SelK probably complexes with itself as a putatively alternate oligomerization state *via* the diselenide bond formed by the Sec.

Analogous to SelK, it was determined that chicken SelS contained an unknown domain. The 14–193 residues that form the α-helices and other principal parts of SelS might be responsible for its redox and structural properties (Fig. 4). A multiple sequence alignment indicated that SelS was present in at least three different types, based on the quantity of the residues between Sec and its partnering Cys, which might affect the stability of the selenylsulfide bond formed by Sec and Cys.[26,27] The SelS of chickens was classified as type I, in which the Sec was separated from the Cys by 13 residues (Fig. 4). The selenylsulfide bond was reduced by the thioredoxin/TrxR system, in which it appeared that the Sec was similar to the second Cys of the CxxC motif of the active site in the thioredoxin fold. Overall, this suggested that Sec plays a role in ER stress-induced apoptosis by protecting the cell against oxidative damage.[28–30]

A multiple sequence alignment of the selenoproteome in *Gallus gallus* demonstrated a conserved CxxU motif in the sequences of SelO, SelH, SelW, SelM and SelT (Fig. 5). These five proteins shared a common β_1-α_1-β_2-β_3-β_4-α_2 thioredoxin-like domain. The Cys and Sec of chicken SelO could form CxxU, possibly indicating that this protein contained a thioredoxin-like fold. This structure suggested that chicken SelO had an oxidoreductase role. However, the Sec in chicken SelO was located at the 650th residue of the C-terminus, which contained multiple α-helices (Fig. 5 and Table S2, ESI†). Although chicken SelO contained a conserved CxxU motif, this selenoprotein could not play a role as the Trx-like protein, although it might have kinase activity.[2] Notably, Asn-314 and Asp-323 in chicken

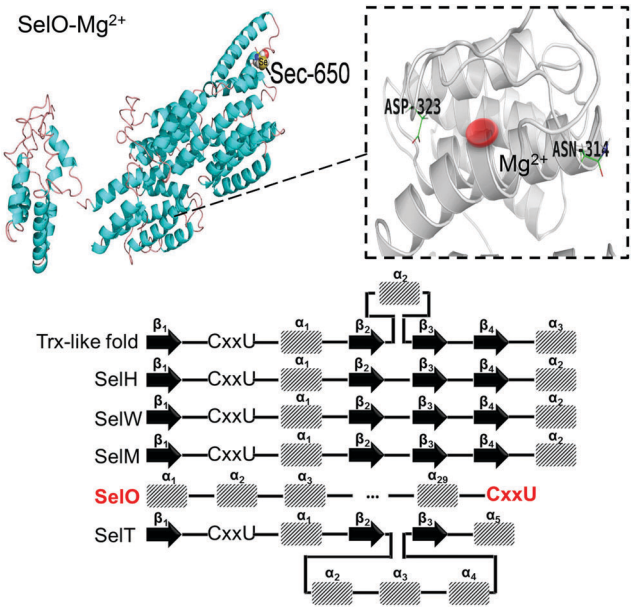

Fig. 5 Interaction between SelO and Mg^{2+} and secondary structure elements of the thioredoxin-like domain in SelO. The red ball represents Mg^{2+}. The β-strands are drawn with black arrows, and the diamond represents the α-strands.

SelO had ligand binding sites containing Mg^{2+}. These two residues were conserved in the SelO family, which may be related to the kinase activity of chicken SelO. However, a kinase role for SelO could only be demonstrated by further experiments.

Zinc-containing matrix metalloselenoproteins (Sep15, MsrB1, SelW and SelM)

The metal ions in protein complexes play a fundamental role in the organization of the secondary or tertiary structure, facilitating interactions between proteins and ligands or directly participating in catalysis.[31,32] Zn^{2+} either plays a structural role in maintaining the structural stability of the protein, or a catalytic role, by participating in chemical catalysis.[33] In this study, we determined that Zn^{2+} functioned in ligand binding in four chicken selenoproteins (Sep15, MsrB1, SelW and SelM) (Fig. 6 and Table S2, ESI†), designated as zinc-containing matrix metalloselenoproteins in birds.

Human Sep15 modulates the enzymatic activity of the glycoprotein glucosyltransferase (UGGT) by assessing the disulfide bonding or thiol/disulfide state of the UGGT substrate.[33,34] Tetrahedral coordination determines the binding between UGGT and 4 Cys residues in Sep15. Zn^{2+} functions in ligand binding with 4 Cys residues (47, 50, 65 and 68) and residues in the chicken Sep15 (Fig. 6). This structural coordination has also been found in the binding between Zn^{2+} and chicken Sep15 (Fig. 6). The combination between Sep15 and UGGT involves signal pathway activation that leads to the reduction of unfolded proteins in the ER.[34] Sep15 is an ER-selenoprotein in birds (Fig. 1). Therefore, we hypothesized that chicken Sep15 interactions with Zn^{2+} might play a role in a Zn^{2+}-regulated signaling pathway in the ER.

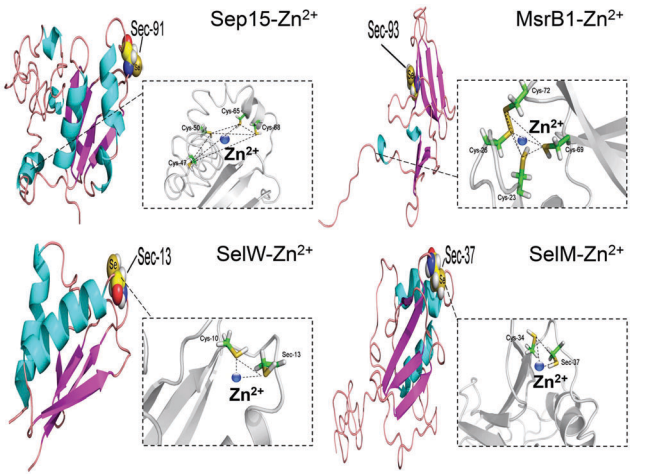

Fig. 6 Binding molecule interactions between the zinc-containing matrix metalloselenoproteins (Sep15, MsrB1, SelW and SelM) and Zn^{2+}. The model is built using the various templates suggested by searches in NCBI-PDB. The zinc sites are tetrahedrally coordinated in Sep15 and MsrB1 combined with Zn^{2+}. SelW and SelM only contain 2 zinc sites that interact with Zn^{2+}.

MsrB shares a common reductase step with the formation of a sulfenic acid intermediate. Thioredoxin acts as a reducing agent for the recycling process of MsrB1 through the formation of an intradisulfide bond.[35] Chicken MsrB1 was found to bind with Zn^{2+} through 4 Cys (23, 26, 69 and 72) residues (Fig. 6). Tetrahedral coordination was also found in the binding between Zn^{2+} and the 4 Cys residues of MsrB1. In this protein, Cys-26 and Cys-72 formed a disulfide bond, which may be the catalytic center of the domain (Fig. 6). This structure suggested that chicken MsrB1 was essential for protecting proteins against oxidative damage.

Chicken SelW contained a conserved CxxU motif, which corresponds to the CxxC redox motif of the active site of thioredoxin (Fig. 5). It was reported that SelW was involved in regulation of the cellular redox status by its thioredoxin-like fold.[36] Moreover, our data revealed that Zn^{2+} functioned in ligand binding with the Cys-10 and Sec-13 residues in chicken SelW (Fig. 6). It is known that the CxxU motif of human SelW is a 14-3-3 protein binding site. Oxidative stress conditions increased these interactions.[37] However, this study might illustrate a novel signaling function for chicken SelW *via* the regulation of redox cell signaling by interaction with Zn^{2+}.

SelM is expressed at a high level in the brain and it is involved in the onset and progression of AD.[38] Chicken SelM contained a thioredoxin-like fold in its elongated C-terminus (Fig. 5). Zn^{2+} was coordinated with the Cys-34 and Sec-37 in chicken SelM (Fig. 6), which suggested that chicken SelM was a Zn^{2+} regulator and modulates metal induced aggregation and neurotoxicity. Se and sulfur share some physicochemical similarities, but Sec has higher nucleophilicity. Chicken SelW and SelM may possess a higher affinity for Zn^{2+}. Additionally, human SelPa could bind with Zn^{2+} during the progression of AD in the brain.[7] However, these results only showed that in chickens, SelPa had 13 Secs and a His-rich domain, but does not bind with Zn^{2+} (Fig. 9).

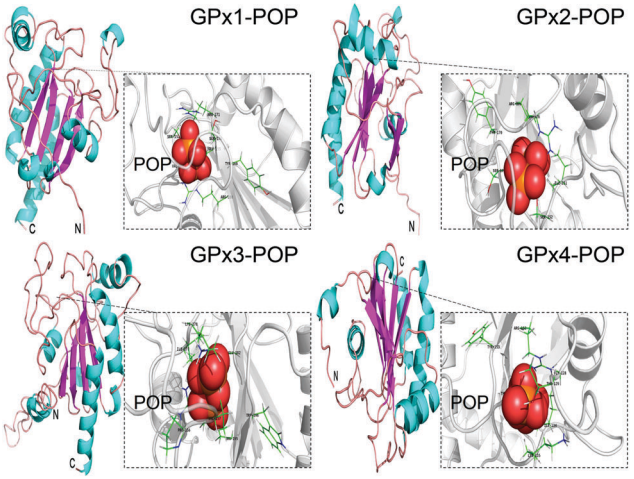

Fig. 7 Binding molecule interactions between POP-containing selenoproteins (GPx1–4) and POP. The α-helix is drawn in blue, the β-strand is indicated by the purple color, and the coil is depicted by pink lines in the tertiary structure of GPx1–4. POP represents the pyrophosphate ion drawn as red balls.

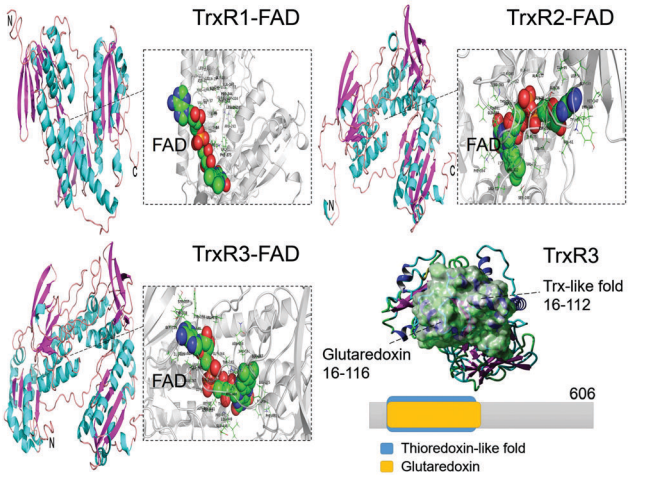

Fig. 8 Binding molecule interaction between TrxR1–3 and FAD and the structure of TrxR3. The helices are in light blue, and the strands are in pink. The thioredoxin-like fold is shown in blue and glutaredoxin in orange. FAD: flavin-adenine dinucleotide.

POP-containing selenoproteins (GPx1–4)

GPx1, a member of the GPx family, was the first selenoprotein identified[39] and fully characterized.[6,40–42] The GPx family (GPx1–4) is also the largest selenoprotein family in chickens. Chicken GPx1–2 had a common subcellular localization, but GPx3 was extracellular and GPx4 was located in the nucleus (Fig. 1). In this study, POP was the binding agent of the GPx family (GPx1–4) in *Gallus gallus* (Fig. 7 and Table S2, ESI†), which suggested that POP had a novel antioxidant defense function in the chicken GPx family.

FAD-interacted selenoproteins (TrxR1–3)

TrxRs contain FAD/NAD (P) binding domains and catalyze the formation of thioredoxin from thioredoxin disulfide.[43,44] Numerous studies using ^{75}Se have determined that TrxR3 was more highly expressed than TrxR1–2 in most cases.[45–47] In this study, FAD was the binding ligand in the TrxR family (TrxR1–3) in *Gallus gallus* (Fig. 8 and Table S2, ESI†). Remarkably, chicken TrxR3 contained a thioredoxin-like fold and a glutaredoxin domain. This structure was not evident in chicken Trx1–2, which suggested that TrxR3 might bind FAD and NAD and accelerate the formation of GSSG to GSH in some cases. Therefore, TrxR3 appeared to be more important in the GSH system than TrxR1–2 in birds.

Secreted transport selenoproteins (GPx3 and SelPa)

GPx3 is a unique selenoprotein in the GPx family. It contains approximately 20% of the Se found in plasma,[48] although this ratio may vary according to the Se status of an individual.[49] We recently reported that the main source of chicken GPx3 in the plasma is the kidney.[50] In mammals, the GPx3 gene encodes an N-terminal secretion signal, and is highly expressed in rodent kidneys.[51,52] The chicken GPx3 gene also encodes a very homologous signal peptide sequence (Fig. S3, ESI†). SelPa is

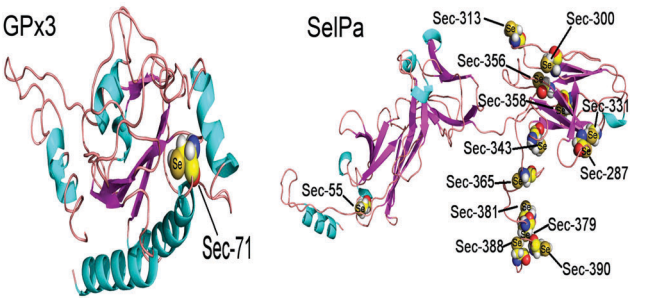

Fig. 9 The tertiary structures of GPx3 and SelPa. GPx3 has 1 Sec in its chain. Chicken SelPa contains 13 Secs. The helices are shown in light blue, and the strands in pink. For each protein, the Secs are shown as spacefills.

extracellular and comprises 40–50% of the Se in plasma.[53,54] SelPa contains 10 Sec residues in humans and mice. Notably, 13 Sec residues were found in chicken SelPa (Fig. 9). The SelPa gene contained two SECIS elements in *Gallus gallus* (Fig. S1, ESI†). These results suggested that GPx3 and SelPa are secreted transport selenoproteins and more important for Se transport in birds.

Only DIO2 contains 2 Secs in the DIO family

Chicken DIO2 contains 2 Secs (Sec-132 and Sec-265) (Fig. 10). The DIO2 mRNA contains a second UGA codon. Recent studies have reported that, in DIO2 in a cell culture system, the second UGA would insert a Sec if the first UGA codon was mutated.[55] However, the thioredoxin-like fold only contained the first Sec. Cloning studies have shown that DIO2 in *Rana catesbeiana* contains only one Sec.[55] The sequence RPLVVNFGSATUPPFT in DIO2, which appears to form the catalytic core, contains the first Sec. This amino acid sequence was more than 80% identical to the active sequence RPLILNFGSCTUPPFM in DIO1 and DIO3. Thus, it could be deduced that the first Sec, present in the putative active center of DIO2, plays a central role in the

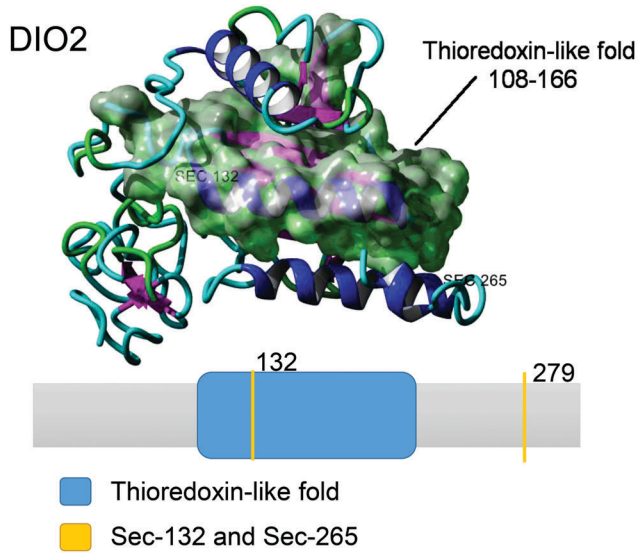

Fig. 10 The tertiary and secondary structures of DIO2. The tertiary structure: the thioredoxin-like fold is drawn to show its surface and the Sec is in the 132nd amino-acid residue. The secondary structure: the gray line represents the main chain of DIO2, and the thioredoxin-like fold is depicted by a blue line.

deiodination process. A TGA codon followed by a pyrimidine is more likely to be translated as a Sec, whereas termination is favored if it is followed by a purine.[56] The second TGA triplet in chicken DIO2 was followed by a purine, which suggested that this second TGA codon might be a termination signal. Therefore, the second Sec could not be connected with any known biological process in birds.

Conclusions

Selenoproteins have long been studied in a wide range of species. In this study, a comprehensive bioinformatics analysis of the structure–function relationships and interactions of binding molecules of selenoproteins was carried out to biochemically characterize the selenoproteome in *Gallus gallus*. Our data provided a wealth of information on the biochemical functions of bird selenoproteins. Members of the selenoproteome were found to participate in various biological processes in chickens, including as antioxidants, maintaining the redox balance, Se transport, and interactions with metals. Six membrane-bound selenoproteins (SelI, SelK, SelS, SelT, DIO1 and DIO3) play an important role in membrane integrity. The chicken selenoproteins were classified according to their ligand binding sites as zinc-containing matrix metalloselenoproteins (Sep15, MsrB1, SelW and SelM), POP-containing selenoproteins (GPx1–4), FAD-interacting selenoproteins (TrxR1–3), secreted transport selenoproteins (GPx3 and SelPa) and other selenoproteins. The results of our study provide new evidence for the previously unknown biological functions of the selenoproteome in birds. Future research is necessary to confirm the novel biochemical functions of bird selenoproteins.

Competing interests

The authors declare that there are no competing interests.

Acknowledgements

This work was supported by China New Century Excellent Talents in University (No. NECT-1207-02), Program for New Century Excellent Talents In Heilongjiang Provincial University (No. 1252-NCET-009), the National Natural Science Foundation of China (No. 31572586) and the Academic Backbone Project of Northeast Agricultural University (No. 15XG16). We also acknowledge the valuable help provided by Prof. Shi-Wen Xu of the Northeast Agricultural University and all of the workers involved.

References

1 F. Li, P. B. Lutz, Y. Pepelyayeva, E. S. Arner, C. A. Bayse and S. Rozovsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 6976–6981.
2 M. Dudkiewicz, T. Szczepinska, M. Grynberg and K. Pawlowski, *PLoS One*, 2012, **7**, e32138.
3 N.-P. Becker, J. Martitz, K. Renko, M. Stoedter, S. Hybsier, T. Cramer and L. Schomburg, *Metallomics*, 2014, **6**, 1079–1086.
4 J. X. Xu, C. Zhang, C. Y. Cao, S. Y. Zhu, H. Li, Y. C. Sun and J. L. Li, *Biol. Trace Elem. Res.*, 2016, **170**, 438–448.
5 B. R. Cardoso, B. R. Roberts, A. I. Bush and D. J. Hare, *Metallomics*, 2015, **7**, 1213–1228.
6 M. Mariotti, P. G. Ridge, Y. Zhang, A. V. Lobanov, T. H. Pringle, R. Guigo, D. L. Hatfield and V. N. Gladyshev, *PLoS One*, 2012, **7**, e33066.
7 X. Du, H. Li, Z. Wang, S. Qiu, Q. Liu and J. Ni, *Metallomics*, 2013, **5**, 861–870.
8 J. Lu and A. Holmgren, *J. Biol. Chem.*, 2009, **284**, 723–727.
9 H. Yao, W. Liu, W. Zhao, R. Fan, X. Zhao, P. A. Khoso, Z. Zhang and S. Xu, *RSC Adv.*, 2014, **4**, 64032–64042.
10 G. V. Kryukov, S. Castellano, S. V. Novoselov, A. V. Lobanov, O. Zehtab, R. Guigo and V. N. Gladyshev, *Science*, 2003, **300**, 1439–1443.
11 M. Mariotti, A. V. Lobanov, R. Guigo and V. N. Gladyshev, *Nucleic Acids Res.*, 2013, **41**, e149.
12 K. C. Chou and H. B. Shen, *PLoS One*, 2010, **5**, e9931.
13 C. S. Yu, C. J. Lin and J. K. Hwang, *Protein Sci.*, 2004, **13**, 1402–1406.
14 J. Y. Yang and Y. Zhang, *Nucleic Acids Res.*, 2015, **43**, W174–W181.
15 J. Y. Yang, R. X. Yan, A. Roy, D. Xu, J. Poisson and Y. Zhang, *Nat. Methods*, 2015, **12**, 7–8.
16 I. Sillitoe, T. E. Lewis, A. Cuff, S. Das, P. Ashford, N. L. Dawson, N. Furnham, R. A. Laskowski, D. Lee, J. G. Lees, S. Lehtinen, R. A. Studer, J. Thornton and C. A. Orengo, *Nucleic Acids Res.*, 2015, **43**, D376–D381.
17 M. E. Moustafa and H. A. Antar, *Biochem. Genet.*, 2012, **50**, 736–747.

18 J. Liu and S. Rozovsky, *Antioxid. Redox Signaling*, 2015, **23**, 795–813.
19 P. A. Janmey and P. K. J. Kinnunen, *Trends Cell Biol.*, 2006, **16**, 538–546.
20 A. Lescure, M. Rederstorff, A. Krol, P. Guicheney and V. Allamand, *Biochim. Biophys. Acta*, 2009, **1790**, 1569–1574.
21 K. Pawlowski, A. Muszewska, A. Lenart, T. Szczepinska, A. Godzik and M. Grynberg, *BMC Genomics*, 2010, **11**, 590.
22 C. L. Chen, M. S. Shim, J. Chung, H. S. Yoo, J. M. Ha, J. Y. Kim, J. Choi, S. L. Zang, X. Hou, B. A. Carlson, D. L. Hatfield and B. J. Lee, *Biochem. Biophys. Res. Commun.*, 2006, **348**, 1296–1301.
23 V. A. Shchedrina, R. A. Everley, Y. Zhang, S. P. Gygi, D. L. Hatfield and V. N. Gladyshev, *J. Biol. Chem.*, 2011, **286**, 42937–42948.
24 A. Polo, G. Colonna, S. Guariniello, G. Ciliberto and S. Costantini, *Mol. BioSyst.*, 2016, **12**, 758–772.
25 D. L. Hatfield, P. A. Tsuji, B. A. Carlson and V. N. Gladyshev, *Trends Biochem. Sci.*, 2014, **39**, 112–120.
26 L. C. Christensen, N. W. Jensen, A. Vala, J. Kamarauskaite, L. Johansson, J. R. Winther, K. Hofmann, K. Teilum and L. Ellgaard, *J. Biol. Chem.*, 2012, **287**, 26388–26399.
27 A. Roseler, J. H. Prieto, R. Iozef, B. Hecker, R. H. Schirmer, S. Kulzer, J. Przyborski, S. Rahlfs and K. Becker, *Antioxid. Redox Signaling*, 2012, **17**, 534–543.
28 Y. Gao, H. C. Feng, K. Walder, K. Bolton, T. Sunderland, N. Bishara, M. Quick, L. Kantham and G. R. Collier, *FEBS Lett.*, 2004, **563**, 185–190.
29 Y. Gao, J. Pagnon, H. C. Feng, N. Konstantopolous, J. B. Jowett, K. Walder and G. R. Collier, *Biochem. Biophys. Res. Commun.*, 2007, **356**, 636–641.
30 J. E. Curran, J. B. Jowett, K. S. Elliott, Y. Gao, K. Gluschenko, J. Wang, D. M. Abel Azim, G. Cai, M. C. Mahaney, A. G. Comuzzie, T. D. Dyer, K. R. Walder, P. Zimmet, J. W. MacCluer, G. R. Collier, A. H. Kissebah and J. Blangero, *Nat. Genet.*, 2005, **37**, 1234–1241.
31 J. A. Tainer, V. A. Roberts and E. D. Getzoff, *Curr. Opin. Biotechnol.*, 1991, **2**, 582–591.
32 J. Zhang, W. Yang, J.-P. Piquemal and P. Ren, *J. Chem. Theory Comput.*, 2012, **8**, 1314–1324.
33 Y. M. Lee and C. Lim, *J. Mol. Biol.*, 2008, **379**, 545–553.
34 V. M. Labunskyy, A. D. Ferguson, D. E. Fomenko, Y. Chelliah, D. L. Hatfield and V. N. Gladyshev, *J. Biol. Chem.*, 2005, **280**, 37839–37845.
35 F. M. Ranaivoson, F. Neiers, B. Kauffmann, S. Boschi-Muller, G. Branlant and F. Favier, *J. Mol. Biol.*, 2009, **394**, 83–93.
36 M. A. Beilstein, S. C. Vendeland, E. Barofsky, O. N. Jensen and P. D. Whanger, *J. Inorg. Biochem.*, 1996, **61**, 117–124.
37 Y. H. Jeon, K. Y. Ko, J. H. Lee, K. J. Park, J. K. Jang and I. Y. Kim, *Biochim. Biophys. Acta*, 2016, **1863**, 10–18.
38 K. V. Korotkov, S. V. Novoselov, D. L. Hatfield and V. N. Gladyshev, *Mol. Cell. Biol.*, 2002, **22**, 1402–1411.
39 M. A. Reeves and P. R. Hoffmann, *Cell. Mol. Life Sci.*, 2009, **66**, 2457–2478.
40 L. Flohe, W. A. Gunzler and H. H. Schock, *FEBS Lett.*, 1973, **32**, 132–134.
41 J. W. Forstrom, J. J. Zakowski and A. L. Tappel, *Biochemistry*, 1978, **17**, 2639–2644.
42 I. Chambers, J. Frampton, P. Goldfarb, N. Affara, W. McBain and P. R. Harrison, *EMBO J.*, 1986, **5**, 1221–1227.
43 J. Lu and A. Holmgren, *Free Radical Biol. Med.*, 2014, **66**, 75–87.
44 E. S. Arner and A. Holmgren, *Eur. J. Biochem.*, 2000, **267**, 6102–6109.
45 A. K. Rundlof, M. Carlsten and E. S. Arner, *J. Biol. Chem.*, 2001, **276**, 30542–30551.
46 L. K. Crosley, C. Meplan, F. Nicol, A. K. Rundlof, E. S. Arner, J. E. Hesketh and J. R. Arthur, *Arch. Biochem. Biophys.*, 2007, **459**, 178–188.
47 B. A. Carlson, M. E. Moustafa, A. Sengupta, U. Schweizer, R. Shrimali, M. Rao, N. Zhong, S. Wang, L. Feigenbaum, B. J. Lee, V. N. Gladyshev and D. L. Hatfield, *J. Biol. Chem.*, 2007, **282**, 32591–32602.
48 H. Koyama, K. Omura, A. Ejima, Y. Kasanuma, C. Watanabe and H. Satoh, *Anal. Biochem.*, 1999, **267**, 84–91.
49 R. F. Burk and K. E. Hill, *Annu. Rev. Nutr.*, 2005, **25**, 215–235.
50 J.-L. Li and R. A. Sunde, *PLoS One*, 2016, **11**, e0152392.
51 S. Yoshimura, K. Watanabe, H. Suemizu, T. Onozawa, J. Mizoguchi, K. Tsuda, H. Hatta and T. Moriuchi, *J. Biochem.*, 1991, **109**, 918–923.
52 R. F. Burk, G. E. Olson, V. P. Winfrey, K. E. Hill and D. Yin, *Am. J. Physiol.: Gastrointest. Liver Physiol.*, 2011, **301**, G32–G38.
53 K. E. Hill, J. Zhou, W. J. McMahan, A. K. Motley, J. F. Atkins, R. F. Gesteland and R. F. Burk, *J. Biol. Chem.*, 2003, **278**, 13640–13646.
54 L. Schomburg, U. Schweizer, B. Holtmann, L. Flohe, M. Sendtner and J. Kohrle, *Biochem. J.*, 2003, **370**, 397–402.
55 D. Salvatore, J. W. Harney and P. R. Larsen, *Biochimie*, 1999, **81**, 535–538.
56 K. K. McCaughan, C. M. Brown, M. E. Dalphin, M. J. Berry and W. P. Tate, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 5431–5435.