Structural model of the CopA copper ATPase of *Enterococcus hirae* based on chemical cross-linking

Mathias Lübben · Reto Portmann · Gerd Kock · Raphael Stoll · Malin M. Young · Marc Solioz

Abstract The CopA copper ATPase of *Enterococcus hirae* belongs to the family of heavy metal pumping CPx-type ATPases and shares 43% sequence similarity with the human Menkes and Wilson copper ATPases. Due to a lack of suitable protein crystals, only partial three-dimensional structures have so far been obtained for this family of ion pumps. We present a structural model of CopA derived by combining topological information obtained by intramolecular cross-linking with molecular modeling. Purified CopA was cross-linked with different bivalent reagents, followed by tryptic digestion and identification of cross-linked peptides by mass spectrometry. The structural proximity of tryptic fragments provided information about the structural arrangement of the hydrophilic protein domains, which was integrated into a three-dimensional model of CopA. Comparative modeling of CopA was guided by the sequence similarity to the calcium ATPase of the sarcoplasmic reticulum, Serca1, for which detailed structures are available. In addition, known partial structures of CPx-ATPase homologous to CopA were used as modeling templates. A docking approach was used to predict the orientation of the heavy metal binding domain of CopA relative to the core structure, which was verified by distance constraints derived from cross-links. The overall structural model of CopA resembles the Serca1 structure, but reveals distinctive features of CPx-type ATPases. A prominent feature is the positioning of the heavy metal binding domain. It features an orientation of the Cu binding ligands which is appropriate for the interaction with Cu-loaded metallochaperones in solution. Moreover, a novel model of the architecture of the intramembranous Cu binding sites could be derived.

Keywords Copper ATPase · Structure model · Cross-linking · Mass spectrometry · Serca1

M. Lübben
Department of Biophysics, Ruhr-University Bochum, 44780 Bochum, Germany

R. Portmann · M. Solioz (✉)
Department of Clinical Pharmacology and Visceral Research, University of Berne, Murtenstrasse 35, 3010 Berne, Switzerland
e-mail: marc.solioz@ikp.unibe.ch

G. Kock · R. Stoll
Department of Biochemistry II, Biomolecular NMR, Ruhr-University Bochum, 44780 Bochum, Germany

M. M. Young
Sandia National Laboratory, Livermore, CA, USA

Introduction

Copper is an essential trace element for all living organisms by acting as a cofactor for redox enzymes like cytochrome *c* oxidase or superoxide dismutase (Linder and Hazegh Azam 1996). However, copper is

also toxic to cells and can damage phospholipids, enzymes and nucleic acids (Yoshida et al. 1993; Ramirez et al. 2005). Therefore uptake, intracellular distribution and export of copper have to be tightly regulated. Copper pumping ATPases fulfill a key role in copper homeostasis by pumping copper across intracellular and cytoplasmic membranes, using the energy of ATP. CopA of *Enterococcus hirae* is encoded by the *cop* operon, which is the key element of copper homeostasis in *E. hirae* (Solioz and Stoyanov 2003). The operon consists of the four genes *copY*, *copZ*, *copA* and *copB*. *CopY* encodes a copper responsive repressor which regulates all four *cop* genes, *copZ* encodes a copper metallochaperone for intracellular copper delivery, and *CopA* and *CopB* code for copper ATPases (Cobine et al. 1999, 2002; Strausak and Solioz 1997; Odermatt and Solioz 1995; Odermatt et al. 1992).

CopA is a protein of 727 amino acids and shares approximately 43% sequence similarity with the human copper pumping ATPases, the Menkes and Wilson proteins (Solioz et al. 1994; Linz and Lutsenko 2007). Like all members of the P-type ATPase family, CopA possesses the characteristic domains of these ion pumps, namely the phosphatase domain of consensus TGES, the phosphorylation domain, DKTGT, the ATP binding site GDG, and a proline residue in the membrane-embedded domain. However, CopA and related ATPases are distinct from NaK- or Ca-ATPases in several regards: (1) they possess one to six N-terminal metal binding modules with a CxxC motif, (2) they feature two additional transmembrane helices at the N-terminal side and lack four at the C terminus, (3) they contain a conserved HP motif in the second large cytoplasmic domain, and (4) they contain a conserved CPC or CPH motif, involving the conserved proline mentioned above, in membrane helix six. This motif has been proposed to form the membrane-located ion binding site (Bissig et al. 2001). Due to the CPx-feature, these ATPases are referred to as 'CPx-type' ATPases, but they have also been classified as P1B-type ATPases (Solioz and Vulpe 1996; Lutsenko and Kaplan 1995). CPx-type ATPases are thought to have evolved before the non-heavy metal ATPases to serve in host defense against toxic metals. Indeed, the spectrum of transported metal ions by CPx-type ATPases is very wide, including copper, silver, cobalt, zinc, cadmium, mercury, and lead.

Detailed three-dimensional structures have been published for the Ca-ATPase of the sarcoplasmic reticulum, Serca1, and the plasma membrane NaK-ATPase (Toyoshima et al. 2000, 2003; Toyoshima and Mizutani 2004; Morth et al. 2007). These structures revealed that the different domains are well separated into an actuator or A-domain, a phosphorylation or P-domain, and a nucleotide binding or N-domain. For CPx-ATPases, only rough models have been reported. They were based on electron cryomicroscopy structures, into which related, known structures were fitted (Chintalapati et al. 2008). Partial structures are available for isolated, soluble domains of CPx-type ATPases: several structures have been published for the unique N-terminal metal-binding domains of copper ATPases [heavy metal associated (HMA) domains (Achila et al. 2006; Banci et al. 2002, 2000)]. The HMA-domains appear to be modular structures which are present in one to six copies at the N terminus of CPx-type ATPases. The fourth of the six metal binding domains of the human Menkes copper ATPase and the first copper binding domain of the yeast Ccc2 copper ATPase possess essentially the same structure (Gitschier et al. 1998). Interestingly, these domains exhibit the same structure as the small chaperones which route copper intracellularly, like CopZ of *E. hirae*, Atx1 of *Saccharomyces cerevisiae*, or human Hah1 (Wimmer et al. 1999; Rosenzweig et al. 1999; Wernimont et al. 2000). The structures of the A-domain of the CopA copper ATPase of the thermophile *Archaeoglobus fulgidus* (Sazinsky et al. 2006a) as well as the combined PN-domains of *A. fulgidus* (Sazinsky et al. 2006b) and *Sulfolobus solfataricus* (Lübben et al. 2007) have also recently been solved. The lack of a complete structure of a heavy metal ATPase has prompted us to investigate the topology of the CopA copper ATPase of *E. hirae* by chemical cross-linking. Intramolecular cross-linking with bifunctional reagents, followed by mass spectrometry and molecular modeling, has been recently shown to provide useful structural information for membrane-bound proteins (Jacobsen et al. 2006). We identified intramolecular cross-links in CopA by tryptic digestion, followed by mass spectrometric identification of the cross-linked peptides. Based on the resultant proximity information, we modeled the protein fold of CopA by combining these constraints with homology modeling (Young et al. 2000).

Materials and methods

Materials

All homo-bifunctional cross-linkers were supplied by Pierce (Rockford, IL) and the sequencing grade modified trypsin by Promega (Madison, WI). All other chemicals were from Sigma Chemical Corp. (St Louis, MO) or from Merck (Darmstadt, Germany) and were of analytical grade.

Purification and cross-linking of CopA

CopA was expressed from a plasmid in *E. coli* and purified as described in Wunderli-Ye and Solioz (2001). It contained an N-terminal 25 amino acid extension with a his-tag and spacer sequence for purification (Fig. 1, *dark green*). The amino acid numbering of CopA used here does not take this modification into account and corresponds to the numbering of the native protein. Purified CopA was diluted in buffer C (100 mM HEPES pH 7.5, 1 mM EDTA, 10 mM DTT, 0.05% dodecylmaltoside) to 5 μM and was dialyzed at 4°C against buffer C for 2 h and overnight. The primary amine specific homo-bifunctional monoester cross-linkers dimethyl-adipimidate (DMA) and disuccinimidyl suberate (DSS, Fig. 2) were added from 100 mM stock solutions to a final concentration of 100 μM and the reaction mixtures incubated at room temperature for 6 h. A control sample without cross-linker was incubated in parallel. Following incubation, the reactions were separated by SDS polyacrylamide gel electrophoresis (Laemmli and Favre 1973). Gels were stained with Coomassie Blue and the bands corresponding to monomeric CopA were cut out of the gel and minced with a scalpel for subsequent processing.

Peptide modification and proteolytic digestion

Before each processing step, the gel pieces were washed with 25 mM NH$_4$HCO$_3$, followed by shrinking with 50% acetonitrile, 25 mM NH$_4$HCO$_3$. Free cysteines were modified by reduction with 10 mM DTT in 25 mM NH$_4$HCO$_3$ and alkylation with 20 mM iodoacetamide in 25 mM NH$_4$HCO$_3$. CopA was then digested with 19 ng/μl of sequencing grade modified trypsin (Promega) in 47 mM Tris–Cl, pH 9.

Peptides were extracted from the gel with 50% acetonitrile, 25 mM NH$_4$HCO$_3$. The gel was rehydrated with 25 mM NH$_4$HCO$_3$ under sonication and a second extraction with 50% acetonitrile, 25 mM NH$_4$HCO$_3$ was conducted. The pooled extracts were dried in a SpeedVac.

Mass spectrometry

For matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF MS), the peptide pellets were taken up in 10 μl of buffer M (50% acetonitrile, 0.1% trifluoroacetic acid) and 1 μl was spotted on an AnchorChip 400/384 (Bruker Daltonics) in α-cyano-4-hydroxycinnamic acid. The spectra were recorded on a Bruker Daltonics Ultraflex. For liquid chromatography-electrospray ionization (LC-ESI)-MS/MS analysis, the pellets were dissolved in 10 μl of buffer A (2% acetonitrile, 0.1% trifluoroacetic acid), adsorbed on a C18 column and the peptides eluted over 45 min at a flow rate of 600 nl/min with a linear 5–60% gradient of buffer B (80% acetonitrile, 0.1% trifluoroacetic acid). Mass spectra were recorded on a LCQ Deca XP instrument (Thermo Finnigan).

Mass spectra assignment

A list of theoretical fragment sizes generated by the different cross-linkers was calculated using the Automated Spectrum Assignment Program (ASAP), developed at the University of California, San Francisco. The MALDI-TOF mass spectra were overlaid and the differences between the non-cross-linked sample and the DMA- and DSS-treated samples were identified. The newly appearing fragment sizes were compared with the calculated theoretical masses to identify cross-linked peptides.

Modeling of CopA

The homology modeling strategy was based on the assumed structural similarity of CopA to the Ca-ATPase, Serca1. To this end, a binary comparison of both sequences was performed with CLUSTALW (Larkin et al. 2007). N-terminal transmembrane helices TM1 and TM2 of CopA are absent in Serca1, thus TM3–TM8 of CopA were considered homologous to TM1–TM6 of Serca1. The program TM-HMM

Fig. 1 Sequence alignment of CopA and Serca1. The sequence of CopA (accession code P04191) was aligned with Serca1 (accession code P04191) as detailed under "Materials and methods". The *Tag* added for purification of CopA is indicated in *dark green*. Transmembrane helices are highlighted in *dark grey*, adjacent hydrophobic residues in *light grey*. Other functional domains are *color coded* and labeled in the figure. The following conserved sequence features are *underlined*: TGES, phosphatase domain; CPC, putative part of the copper channel through the membrane; DKTGT, phosphorylation site; HP, conserved feature of unknown function; GDG(I,V)ND, ATP binding region. Features in *dark purple* are also indicated in Fig. 4 and the amino acid numbers are indicated

```
                     Tag                            HMA-domain        17-20
CopA    MSYYHHHHHHDYDIPTTENLYFQGAMATNTKMETFVITGMTCANCSARIEKELNEQPGVM   35

CopA    SATVNLATEKASVKYTDTTTERLIKSVENIGYGAILYDEAHKQKIAEEKQTYLRKMKFDL  120
             TM1                        TM2
CopA    IFSAILTLPLMLAMIAMMLGSHGPIVSFFHLSLVQLLFALPVQFYVG-------------  142
Serca1                                              MEAAHSKSTEECL   13
                                                           A1-domain
CopA    ----------------WRFYKGAYHALKTKAPN----------MDVLVAIGTSAAFALS  175
Serca1  AYFGVSETTGLTPDQVKRHLEKYGHNELPAEEGKSLWELVIEQFEDLLVRILLLAACISF   73
           TM3              TM4                              TM1
CopA    IYNGFFPSHSHDLYFESSSMIITLILLGKYLEHTAKSKTGDAIKQMMSLQTKTAQVLRDG  235
Serca1  VLAWFEEGEETITAFVEPFVILLILIANAIVGVWQERNAENAIEALKEYEPEMGKVYRAD  133
                   TM2              A2-domain
CopA    KE--ETIAIDEVMIDDILVIRPGEQVPTDGRIIAGTS---ALDESMLTGESVPVEKK---  287
Serca1  RKSVQRIKARDIVPGDIVEVAVGDKVPADIRILSIKSTTLRVDQSILTGESVSVIKHTEP  193

CopA    ----------EKDMVFGGTINTNGLIQIQVSQIGKDTVLAQIIQMVEDAQGSKAPIQQIA  337
Serca1  VPDPRAVNQDKKNMLFSGTNIAAGKALGIVATTGVSTEIGKIRDQMAATEQDKTPLQQKL  253
                    TM5                       TM6   381-383
CopA    DKISGIFVPIVLFLALVTLLV----------TGWLTKDWQLALLHSVSVLVIACPCALGL  387
Serca1  DEFGEQLSKVISLICVAVWLINIGHFNDPVHGGSWIRGAIYYFKIAVALAVAAIPEGLPA  313
                  TM3                          TM4   309-310
                               425
CopA    ATPTAIMVGTGVGAHNGILIKGGEALEGAAHLNSIILDKTGTITQG-------------  433
Serca1  VITTCLALGTRRMAKKNAIVRSLPSVETLGCTSVICSDKTGTLTTNQMSVCKMFIIDKVD  373
                P-domain              351        N1-domain
CopA    -----------------------------------------------------------
Serca1  GDFCSLNEFSITGSTYAPEGEVLKNDKPIRSGQFDGLVELATICALCNDSSLDFNETKGV  433

CopA    ---------------------------------------------RPEVTDVIGPKEI  446
Serca1  YEKVGEATETALTTLVEKMNVFNTEVRNLSKVERANACNSVIRQLMKKEFTLEFSRDRKS  493
              485-486                                N2-domain
CopA    ISLFYSLEHASEHPLGKAIVAYGAKVGAKTQPITDFVAHPGAGISGTING-----VHYFA  501
Serca1  MSVYCSPAKSSRAAVGNKMFVKGAPEGVIDRCNYVRVGTTRVPMTGPVKEKILSVIKEWG  553

CopA    GTRKRLAEMNLSFDEFQEQALELEQAG-KTVMFLANEEQVLGMIAVADQIKEDAKQAIEQ  560
Serca1  TGRDTLRCLALATRDTPPKREEMVLDDSSRFMEYETDLTFVGVVGMLDPPRKEVMGSIQL  613
                                                       P-domain
CopA    LQQKGVDVFMVTGDNQRAAQAIGKQVGIDSDH---------------------------  592
Serca1  CRDAGIRVIMITGDNKGTAIAICRRIGIFGENEEVADRAYTGREFDDLPLAEQREACRRA  673

CopA    -IFAEVLPEEKANYVEKLQKAGKKVGMVGDGINDAPALRLADVGIAMGSGTDIAMETADV  651
Serca1  CCFARVEPSHKSKIVEYLQSYDEITAMTGDGVNDAPALKKAEIGIAMGSGTAVAKTASEM  733
                                                   TM7
CopA    TLMNSHLTSINQMISLSAATLKKIKQNLFWAFIYNT----IGIPFAAFGFLNPIIAGGAM  707
Serca1  VLADDNFSTIVAAVEEGRAIYNNMKQFIRYLISSNVGEVVCIFLTAALGLPEALIPVQLL  793
            TM8                                    TM5
CopA    AFSSISVLLNSLSLNRKTIK                                         727
Serca1  WVNLVTDGLPATALGFNPPDLDIMDRPPRSPKEPLISGWLFFRYMAIGGYVGAATVGAAA  853
             TM6                              TM7
Serca1  WWFMYAEDGPGVTYHQLTHFMQCTEDHPHFEGLDCEIFEAPEPMTMALSVLVTIEMCNAL  913
                                                     TM8
Serca1  NSLSENQSLMRMPPWVNIWLLGSICLSMSLHFLILYVDPLPMIFKLKALDLTQWLMVLKI  973
             TM9                        TM10
Serca1  SLPVIGLDEILKFIARNYLEDPEDERRK                                1001
```

(Krogh et al. 2001) was used to predict the putative transmembrane regions and to determine hydrophilic/hydrophobic boundaries (Fig. 1). The hydrophilic portion of CopA was dissected into the HMA-, N-, P-, and A-domains based on sequence alignments. These domains were submitted to SWISSMODEL (automatic mode). Based on the existing structural homologues 2B8E (N-, P-domains), 2HC8 (A-domain) and 1OPZ (HMA-domain), structural models of the respective domains were obtained. Structural alignments of the domains with Serca1 using DAL-ILITE (Holm and Park 2000) produced residue–
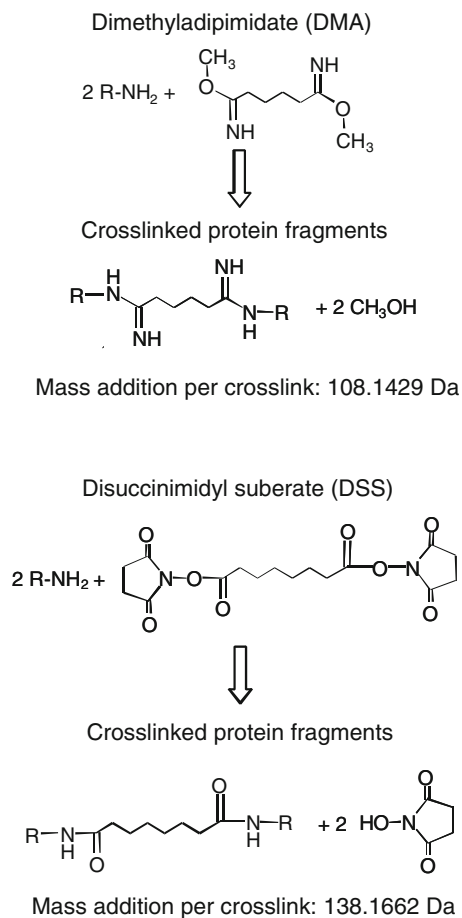
Dimethyladipimidate (DMA)

Mass addition per crosslink: 108.1429 Da

Disuccinimidyl suberate (DSS)

Crosslinked protein fragments

Mass addition per crosslink: 138.1662 Da

Fig. 2 Cross-linking chemistry of DMA and DSS. Dimethyl adipimidate (DMA) and disuccinimidyl suberate (DSS) specifically cross-link lysine residues, resulting in the mass addition to cross-linked peptides indicated in the figure

residue correlations, which could be used to manually improve the binary sequence alignment of CopA and Serca1. The corrected data were submitted to SWISS-MODEL (alignment mode) to model CopA with Serca1 (1IWO) as a template structure. The resulting model (termed Model-1) encompassed the CopA transmembrane domain with helices TM3–TM8, connected to the A-, N- and P-domains.

To determine the binding location of the HMA domain relative to the polar domain cluster of Model-1, unbiased docking calculations with the program 3DGARDEN (Lesk and Sternberg 2008) yielded a high-scoring orientation, which placed the HMA and actuator domains next to each other. Visual inspection of the molecular surfaces in this region revealed complementary patches of opposite charges (basic

HMA, acidic actuator domain), which were assumed to form a contact interface. A detailed determination of the binding site of HMA on Model-1 was carried out with the docking program HADDOCK (Dominguez et al. 2003), for which the patching residues on HMA K26, K45, K49 and on Model-1 D313, M322 and D325 were given high priority. In order to reduce model bias, several start conditions were applied, which converged into a favorable orientation with lowest HADDOCK scores. Distance constraints from cross-links were used to select domain orientations that had the lowest number of constraint violations (exceeding the maximum Cα–Cα limit of 25 Å).

To bridge the gap between the HMA domain located at the N terminus (red colored structure in Fig. 3a) and the rest of the molecule (Model-1), the N-terminal transmembrane helices TM1 and TM2, together with their linking loops (blue colored parts in Fig. 3a), were constructed manually yielding Model-2. The conformation of the loops between HMA/TM1, TM1/TM2, and TM2/TM3 of Model-2 was energy-minimized using the program XPLOR-NIH (Schwieters et al. 2003). Graphic representations were created with the program PYMOL (DeLano 2002).

Results

Intramolecular cross-linking of purified, native CopA, was conducted with the bifunctional cross-linkers DMA and DSS. Dimethyl suberimidate was also tested but did not result in identifiable cross-links. These chemicals can react with the ε-amino groups of lysines or with protein N-termini, producing either single residue modifications or inter-residue cross-links (Fig. 2). Of course inter-molecular cross-links which lead to CopA dimers or multimers can also occur. These were removed by separating CopA monomers from dimers and multimers by SDS–PAGE after the cross-linking reaction. The band corresponding to the CopA monomer was extracted from the gel and used to identify intra-molecular cross-links as described under "Materials and methods". As a control, uncross-linked CopA was also extracted from the gel, trypsinized, and the tryptic peptides analyzed by LC-MS/MS, yielding a protein coverage of 53%. The fragments of CopA which could not be identified in the mass spectra mainly contained very hydrophobic regions, such as the transmembrane domains.
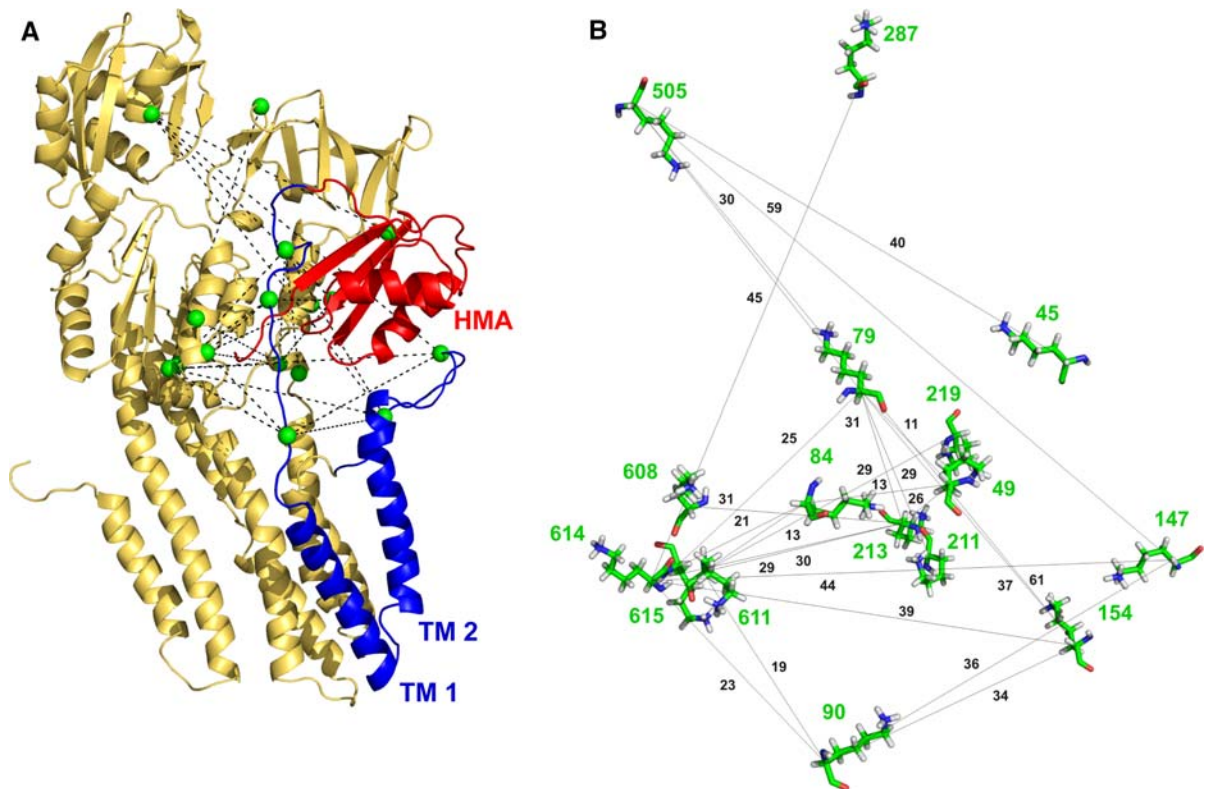
Fig. 3 Structural model (Model-2) of CopA based on chemical cross-linking and sequence similarity modeling. **a** Model of the CopA structure. TM1 and TM2, which are not present in Serca1, are shown in *blue* and the CopA HMA domain is displayed in *red*. Identified Lys-Lys cross-links are displayed as *dashed lines* between *green spheres*, which indicate the Cα-atoms of the corresponding lysine residues. **b** Spatial positions and residues numbers of cross-linked lysine residues of CopA. Lysine residues are in *green* and numbered in *green*. Identified cross-links are indicated by *dashed lines* and the distances in Å are indicated with *black* numbers

Due to the low mass resolution of the LC-MS/MS and the difficulty of interpreting MS/MS spectra of cross-linked peptides, MALDI-TOF MS was used to identify cross-linked peptides. The masses of uncross-linked and cross-linked CopA peptides were assessed (1) by recording multiple spectra of different individual cross-linking experiments and (2) by comparison of the measured fragment masses between all control and cross-linked samples. Peptide masses were compared with a list of theoretical fragment masses calculated for each cross-linker using ASAP.

Since trypsin cleaves proteins at lysine (and arginine) residues, it was taken into consideration that modified lysine residues inhibit trypsin cleavage. This type of analysis resulted in the identification of 18 intramolecular cross-links in six regions of the ATPase, namely the C-terminal part of the copper binding domain, the N-terminal part of the transmembrane helix TM1, the loop between TM2 and TM3, the C-terminal part of TM4, the region close to the HP motif, and the ATP binding domain (Table 1).

Using CLUSTALW and the BLOSUM80 scoring matrix, partial amino acid sequences of CopA (amino acids 168-752) were aligned with the Serca1 sequence (Fig. 1). The alignment showed that the region between transmembrane helices TM3 and TM8 of CopA exhibits extensive sequence similarity to the region between TM1 and TM6 of Serca1, except for two major regions which are absent in CopA. Based on this, the A-domain of Serca1 can be divided into an A1-domain and an A2-domain, of which CopA only possesses the A2-domain. Similarly, CopA only possesses a homologue of the N2-domain of Serca1. Not surprisingly, the "accessory" A1- and N1-domains of Serca1 have surface locations on the respective domains and do not make contact to other domains.

Table 1 Identified cross-links and intramolecular Cα–Cα distances between lysine residues

Observed M + H$^+$	Cross-linked peptides	Lys-Lys cross-links	Distance Cα–Cα (Å)	Domain contact	Cross-linker	Error (ppm)
943.582	612-615, 287-289	614-287	44.5	P-A2	DSS	91
1,254.712	505-506, 78-84	505-79	30.2	N2-HMA/TM1[a]	DMA	82
1,576.965	612-615, 85-92	614-90	23.4	P-HMA/TM1	DMA	21
1,836.024	505-506, 145-156	505-(147 or 154)	58.9, 61.5	N-TM2/3[b]	DMA	56
1,888.027	612-615, 80-90	614-84	20.4	P-HMA/TM1	DMA	72
1,940.285	78-84, 148-156	79-154	36.3	HMA/TM1-TM2/3	DMA	65
1,964.130	612-615, 145-156	614-(147 or 154)	43.3, 38.7	P-TM2/3	DMA	52
1,969.365	90-92, 145-156	90-(147 or 154)	36.0, 34.4	HMA/TM1-TM2/3	DSS	92
2,297.306	212-219, 56-57	213-49	26.3	A2-HMA	DMA	20
2,353.081	78-84, 46-57	79-49	10.3	HMA/TM1-HMA	DSS	93
2,391.439	609-614, 615-630	611-615	9.2	P-P	DMA	>1
2,418.523	609-614, 80-92	611-(84 or 90)	13.4, 18.8	P-HMA/TM1	DSS	24
2,430.334	609-614, 239-252	611-219	29.0	P-A2	DMA	35
2,430.422	212-219, 603-615	213-(608, 611 or 614)	30.5, 29.7, 28.7	A2-P	DMA	29
2,564.269	78-84, 590-605	104-640	24.3	HMA/TM1-P	DMA	70
2,613.502	78-84, 205-219	104-(236 or 238)	27.4, 30.4	HMA/TM1-A2	DMA	6
2,728.434	80-89, 46-57	84-49	13.3	HMA/TM1-HMA	DMA	16
2,825.572	505-506, 27-49	505-45	40.4	N2-HMA	DMA	3

[a] Loop between HMA and TM1

[b] Loop between TM2 and TM3

The number and arrangement of transmembrane helices of CPx-type ATPases deviates considerably from that of other P-type ATPases. TM1 and TM2 of CopA have no counterpart in Serca1 and TM3–TM8 of CopA correspond to TM1–TM6 of Serca1 (cf. Fig. 1). The last four transmembrane helices of Serca1, TM7–TM10 are absent in CopA. All the major hydrophilic domains of CopA exhibit sequence similarity to proteins or domains of known structures: The HMA-domain resembles similar domains of yeast and human copper ATPases, the A2-domain resembles the A-domain of *A. fulgidus* CopA (Sazinsky et al. 2006a), and the N/P-domains resemble the respective domains of *A. fulgidus* and *S. solfataricus* copper ATPases. Given these similarities, it became feasible to confidently model the structure of CopA, using cross-linking information for the overall localization of protein domains (Table 1). This was most important for structural assignment of the HMA domain relative to the CPx-ATPase core, because no other experimental data or computational hints by overall structural homology were available. The direct contacts of cross-links between the lysine pairs 213-49 and 505-45 connecting the HMA domain with actuator and nucleotide binding domains were pivotal to unequivocally assign the correct placement of the HMA out of a selection of alternative positions provided by the programs 3DGARDEN and HADDOCK.

The derived structural model of CopA has the typical domain architecture of P-type ATPases, consisting of transmembrane helical elements and cytoplasmic domains that form a large cluster (Fig. 3a). For modeling, strong emphasis was given to the preservation of the relative orientations of the compact A-, N-, and P-domains, which were assumed to behave like rigid bodies. Almost all of the lysines found to be chemically cross-linked are located in the hydrophilic parts of the molecule. The network formed by these cross-links is schematically displayed in Fig. 3b. In the final model, some of the Lys-Lys distances are considerably larger than the
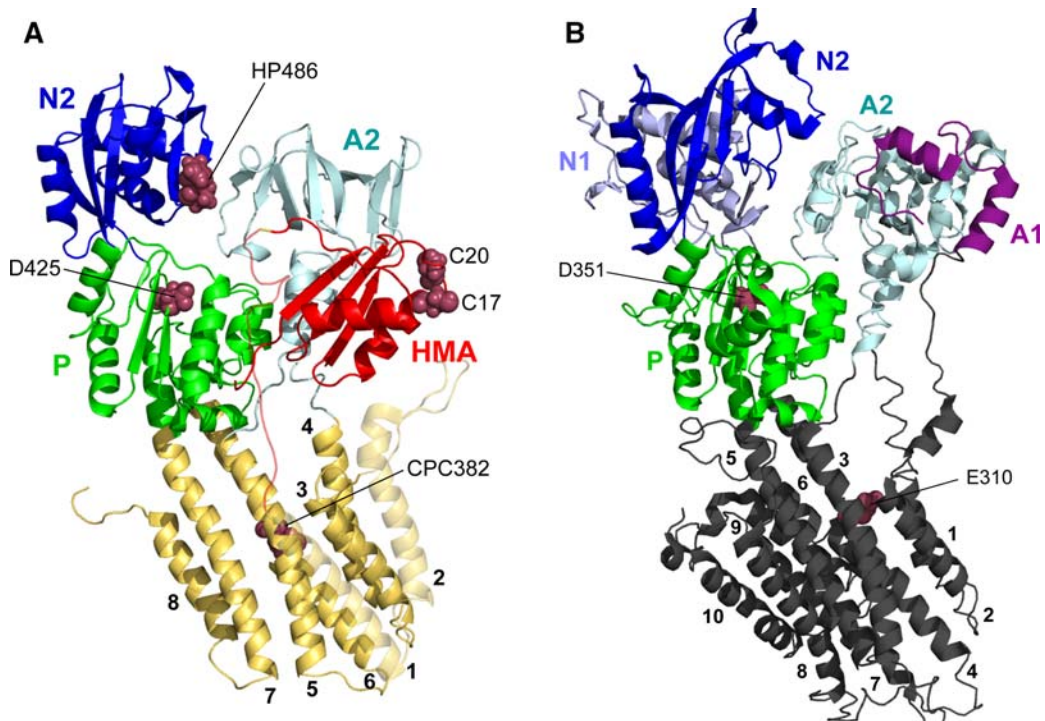
Fig. 4 Comparison of the CopA model (Model-2) with the structure of Serca1. **a** CopA model with the HMA-domain colored in *red*, the A2-domain in *grayish blue*, the N2-domain in *dark blue*, the P-domain in *green*, and the transmembranous domain in *ochre*. TM1 and TM2 and their connecting loops are displayed in a lighter shade to indicate the spatial uncertainty. **b** Structure of Serca1 (PDB entry 1IWO). The A1-domain is colored in *dark purple*, the A2-domain in *grayish blue*, the N1-domain in *light blue*, the N2-domain in *dark blue*, the P-domain in *green*. The membrane domain is displayed in *dark grey* and the membrane helices are numbered. Characteristic residues discussed in the text and highlighted in Fig. 1 are shown in *brown* space-filling representation and are labeled

theoretical 25 Å, especially those involving loops between domains or transmembrane segments. Some of the longest distances bridged by the cross-linkers were found to involve K505, which is located on the N-domain (Table 1). This domain is flexibly attached to the P-domain via hinge peptides and has been shown to undergo considerable movements during catalytic activity in the Ca-ATPase.

Figure 4 shows the similarity of the orientations between the A-, N- and P-domains of CopA and Serca1. The most prominent residue in the P-domain is the invariant phosphorylated aspartic acid (D425 in CopA, D351 in Ca-ATPase). Together with the highly conserved H485–P486 pair in the N-domain, these residues delineate the contour of the ATP binding site. The A-domain is in direct contact with the N- and P-domains in both, CopA and Serca1. The HMA-domain is a typical feature of CPx-type ATPases and is not present in Serca1 and other non-heavy metal ATPases. In our model, this domain

is located next to the A-domain and is attached to TM1 by an extended loop. It has to be pointed out that the structural ambiguity of this loop segment is fairly high, because of the lack of sequence similarity to Serca1. The invariant proline at position 382 in TM6 of CopA comprises part of the CPC motif. This motif is believed to be part of the ion channel across the membrane, similar to the PE310 motif of Serca1. As expected, the CPC motif is embedded in the integral membrane portion of the protein in the CopA model. C381 and C383 on TM6, together with Y685 and N686 on TM7 and M707 and S711 on TM8 form the ligands responsible for the transmembrane high affinity copper binding sites I and II as identified by Argüello and coworkers (Gonzalez-Guerrero et al. 2008). Our Model-2 (Fig. 5) provides a feasible structure of these two copper binding sites. Furthermore, the cysteine residues of the CxxC motif of the HMA-domain, C17 and C20, are facing to the outside of the molecule. In this position, the metal binding
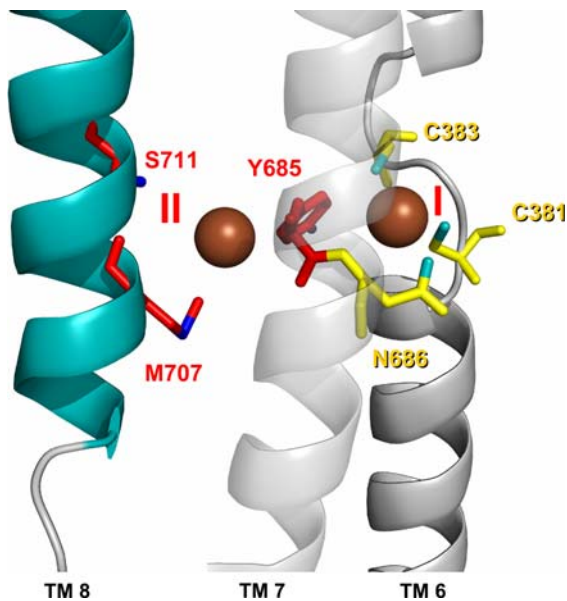
Fig. 5 Enlarged view of the Cu binding sites I and II located within the transmembrane helices TM6 (*dark grey*), TM7 (*light grey*) and TM8 (*blue*). Ligands placed as in our Model-2 using the following assignment: site I (*yellow residues*): C381 and C383 on TM6, and N686 on TM7; site II (*red residues*): Y685 on TM7, and M707 and S711 on TM8

site would be in favorable position to capture copper from copper-loaded chaperones, which presumably donate copper to the ATPase (Arnesano et al. 2001, 2004).

Discussion

There have been several previous attempts to model the structure of a CPx-type ATPase. Fatemi et al. presented a model of the overall structure of the Wilson copper ATPase, ATP7B, solely based on sequence similarity to Serca1 (Fatemi and Sarkar 2002). While it was possible to model the large domain from TM3 to TM7 and, individually, the six N-terminal copper-binding domains, their position as well as the positions of TM1, TM2 and TM8 remained unresolved. We probed the CopA structure by intra-molecular cross-linking of neighboring peptides and identification of the cross-links by MALDI-TOF MS. Due to the hydrophobicity and the scarcity of lysine residues in the transmembrane helices of

CopA, no cross-links could be found in the membrane domain or between the membrane domain and other domains. The necessary presence of detergent in the cross-linking reaction may additionally have hindered cross-linking. There were six hotspots for cross-links (cf. Fig. 3b) and all cross-links were found in the putative functional centre of CopA.

The model of CopA was based on sequence and structure similarities to Serca1. The E2 structure of Serca1, represented by the PDB entry 1IWO, was chosen as modelling template because it is most compatible with the experimental conditions of cross-linking of CopA, namely the absence of the Cu^+ and ATP. Figure 3a represents a composite model, a large portion of which was generated by similarity modeling of distinct partial fragments for which structures are already available in the PDB database. Other parts of the model are speculative, especially the position and orientation of the transmembrane segments and their connecting loops. Although the membrane helices of Serca1 do not exhibit recognizable sequence similarity to the membrane helices of CopA, TM1–TM6 of Serca1 must evolutionarily be related to TM3–TM8 of CopA. The evolution of membrane helices underlies different constraints (hydrophobic α-helices) than functional protein domains (e.g., ATP binding) and are generally more divergent than catalytic domains in distantly related proteins. TM1, TM2 and the loops of CopA have no apparent correlation with Serca1 sequences. Thus, the assignment of TM1 and TM2 is somewhat arbitrary; they were placed such as to bridge the HMA domain reasonably to the rest of the molecule.

The model presented in Fig. 3 was checked with VERIFY3D (Eisenberg et al. 1997), which determines the compatibility of the sequence 3D profile with the calculated molecular coordinates: the well conserved and compactly organized HMA-, A-, P-, and N-domains are characterized by relatively high scoring functions, suggesting a higher level of confidence for these regions. In contrast, the arrangement of the transmembrane regions and the polar connecting loops are much less certain. These regions exhibit lower VERIFY3D scores since they are less constrained due to low sequence similarity. Cross-links between lysine residues formed by homobi-functional cross-linkers were predominantly located in the hydrophilic portion of CopA.

A novel feature of the present model is the placement of the HMA domain in the immediate neighborhood of the A-domain. This assignment could not be performed by homology-based techniques, because the HMA domain is absent in structurally resolved P-type-ATPases. Combination of experimental cross-linking and computational methods was necessary: (1) Initial docking results by 3DGARDEN providing lowest pseudo-energy scores, which were checked for consistency by fitting to the constraints imposed by the crosslinked lysine-pairs 213-49 (26 Å) and 505-45 (40 Å) and 505-79 (30 Å, see Table 1). (2) The respective molecular surfaces exhibited a complementary electrostatic fit, which was fixed for another round of fine-tuned docking calculations with HADDOCK. This provided the same structural clusters which were independent of the start conditions. Upon choosing up to 180° differing starting positions, we always obtained the representative interaction surface presented in Figs. 3a and 4a, and verified it to be in accordance with the experimental distance constraints. Other orientations of the HMA domain could be ruled out, because they would violate the cross-linking geometry, resulting in false lengths. In the structural model of Wu et al. (2008) derived by cryoelectron microscopy of tubular crystals of the CopA copper ATPase of A. fulgidus, the HMA domain was placed on the opposite flank of the ATPase core. Such an orientation is incompatible with our experimental data: the formation of cross-links between lysine pairs 505-45 and 505-79 would require the crosslinker to penetrate the protein. (3) In addition to the different location of the HMA domain proposed by Wu et al. our model also predicts an orientation of the CxxC copper binding motif facing to the outside and thus accessible to copper chaperones. Such an arrangement is biologically much more plausible than that of Wu et al. in which the CxxC metal binding motif is buried in the protein and thus not accessible for metallochaperones.

The proposed orientation of the HMA domain we propose here is in accord with the experimental data: interpretation of peptide fragment patterns after tryptic digestion of the Thermotoga maritima CopA in various functional states has provided evidence that its HMA- and short A-domain together are functionally equivalent to the longer A-domain of Ca-ATPase alone (Hatori et al. 2007). It has been suggested that the HMA-domain substitutes for the N-terminal missing portion of the shortened A-domain in CPx-type ATPases, which obviously requires close proximity of both domains. The extended part of the A-domain (Fig. 1, "phosphatase domain I" and Fig. 4b, domain colored in purple) of Ca-ATPase is positioned at the N-terminus, next to the first transmembrane helix TM1 (Kühlbrandt 2004). An analogous situation was found with CopA here, where the HMA domain is directly followed by TM1 (Figs. 3a, 4a). The critical role of the close connection between TM1 and the N-terminus has been described as "transmitting the movement of the A-domain to the transmembrane gate, by pulling up helix M1 toward the cytoplasm" (Hatori et al. 2007). Consistent with this notion, TM1 and TM2 in our model are placed in the neighborhood of TM3, which avoids crossing of the connecting loops. This model contrasts in the orientation of transmembrane helices with the recently published model of Chintalapati et al. for the CPx-type ATPase CtrA3 of Aquifex aeolicus (Chintalapati et al. 2008). The model was derived by fitting known structures to a projection map derived by electron diffraction of 2D crystals. Indeed, the authors admit that an alternative assignment of their electron density data might not be ruled out at the present state.

Our model of CopA also provides insight into the architecture of the transmembrane high-affinity copper binding sites, whose liganding amino acid residues on TM6–TM8 have been determined recently (Mandal et al. 2004). Figure 5 displays the copper sites I and II, modeled by using the topology of the membrane helices of Model-2. The copper site I is formed by C381, C383, and N686 and the copper site II by Y685, M707, and S711. This ligand assignment differs from that of Gonzalez-Guerrero et al. (2008). By measuring copper binding in multiple mutants of A. fulgidus CopA, they concluded that copper site I is formed by C380, C382, and Y682 and site II by N683, M711, and S715 Such a ligand assignment would only become compatible with our Model-2 if TM8 (blue helix in Fig. 5) would be substantially shifted and rotated. Gonzalez-Guerrero et al., do admit that their ligand assignment is not consistent with previous models of the CPx-ATPase although transmembrane metal binding sites were not explicitly discussed in these reports. To clarify the issue of transmembrane helix orientation and ligand

assignment, contacts between different helices could be determined experimentally by a disulfide cross-linking method (Bragg 1998; van der Sluis et al. 2002; Vandevuer et al. 2006), for which Model-2 would suggest potential residues to be replaced by cysteines.

Ideally, all distance information provided by the cross-linking experiments reported here should be taken into account to construct a model and all distances should be consistent with the model. However, some of the distance constraints had to be ignored. In a rigid structure, the maximal distance between cross-linked lysine residues is limited by the stretched linear states of the crosslinker and the lysine side-chains, resulting in a Cα–Cα spacing of 25 Å. Some distances exceed this value significantly (Table 1), which may be explained in several ways: first, it could be related to the inherent conformational dynamics of CopA, which will undergo thermal fluctuations even in the absence of substrates. Thus, asynchronous bond formation of the homobifunctional cross-linker at very distant positions may take place by loop and/or domain movements. It is furthermore possible, that cross-links connect lysines located in partially unfolded regions. In cases where the model yielded large deviations from the theoretical distance of 25 Å between cross-linked lysine residues, we weighted evidence based on structural similarity between CopA and Serca1 more heavily. Taken together, by combination of theoretical calculations and experimental data we obtained a structural model for CopA, which in many respects resembles the template of Serca1. The model provides a consistent view of the extra features which distinguish CPx-type ATPases from Serca1 and related ATPases. The model could be experimentally tested, for example by the determining transmembrane helical associations as discussed above, or by mutagenesis of residues proposed to be in electrostatic contact between the HMA and the actuator domain.

References

Achila D, Banci L, Bertini I, Bunce J, Ciofi-Baffoni S, Huffman DL (2006) Structure of human Wilson protein domains 5 and 6 and their interplay with domain 4 and the copper chaperone HAH1 in copper uptake. Proc Natl Acad Sci USA 103:5729–5734. doi:10.1073/pnas.0504472103

Arnesano F, Banci L, Bertini I, Cantini F, Ciofi-Baffoni S, Huffman DL, O'Halloran TV (2001) Characterization of the binding interface between the copper chaperone Atx1 and the first cytosolic domain of Ccc2 ATPase. J Biol Chem 276:41365–41376. doi:10.1074/jbc.M104807200

Arnesano F, Banci L, Bertini I, Bonvin AM (2004) A docking approach to the study of copper trafficking proteins; interaction between metallochaperones and soluble domains of copper ATPases. Structure 12:669–676

Banci L, Bertini I, Simone CB, Huffman DL, O'Halloran TV (2000) Solution structure of the yeast copper transporter domain Ccc2a in the apo and Cu(I) loaded states. J Biol Chem 276:8415–8426. doi:10.1074/jbc.M008389200

Banci L, Bertini I, Ciofi-Baffoni S, D'Onofrio M, Gonnelli L, Marhuenda-Egea FC, Ruiz-Duenas FJ (2002) Solution structure of the N-terminal domain of a potential copper-translocating P-type ATPase from *Bacillus subtilis* in the apo and Cu(I) loaded states. J Mol Biol 317:415–429. doi:10.1006/jmbi.2002.5430

Bissig K-D, Wunderli-Ye H, Duda P, Solioz M (2001) Structure-function analysis of purified *Enterococcus hirae* CopB copper ATPase: effect of Menkes/Wilson disease mutation homologues. Biochem J 357:217–223. doi:10.1042/0264-6021:3570217

Bragg PD (1998) Site-directed mutagenesis of the proton-pumping pyridine nucleotide transhydrogenase of *Escherichia coli*. Biochim Biophys Acta 1365:98–104. doi:10.1016/S0005-2728(98)00049-8

Chintalapati S, Al Kurdi R, Terwissscha van Scheltinga AC, Kühlbrandt W (2008) Membrane Structure of CtrA3, a Copper-transporting P-type-ATPase from *Aquifex aeolicus*. J Mol Biol 378:581–595. doi:10.1016/j.jmb.2008.01.094

Cobine P, Wickramasinghe WA, Harrison MD, Weber T, Solioz M, Dameron CT (1999) The *Enterococcus hirae* copper chaperone CopZ delivers copper(I) to the CopY repressor. FEBS Lett 445:27–30. doi:10.1016/S0014-5793(99)00091-5

Cobine PA, George GN, Jones CE, Wickramasinghe WA, Solioz M, Dameron CT (2002) Copper transfer from the Cu(I) chaperone, CopZ, to the repressor, Zn(II)CopY: metal coordination environments and protein interactions. Biochemistry 41:5822–5829. doi:10.1021/bi025515c

DeLano WL (2002) The PyMOL molecular graphics system. http://www.pymol.org

Dominguez C, Boelens R, Bonvin AM (2003) HADDOCK: a protein–protein docking approach based on biochemical

or biophysical information. J Am Chem Soc 125:1731–1737. doi:10.1021/ja026939x

Eisenberg D, Luthy R, Bowie JU (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol 277:396–404. doi:10.1016/S0076-6879(97)77022-8

Fatemi N, Sarkar B (2002) Insights into the mechanism of copper transport by the Wilson and Menkes disease copper-transporting ATPases. Inorg Chim Acta 339:179–187. doi:10.1016/S0020-1693(02)00949-0

Gitschier J, Moffat B, Reilly D, Wood WI, Fairbrother WJ (1998) Solution structure of the fourth metal-binding domain from the Menkes copper-transporting ATPase. Nat Struct Biol 5:47–54. doi:10.1038/nsb0198-47

Gonzalez-Guerrero M, Eren E, Rawat S, Stemmler TL, Arguello JM (2008) Cu^+ transporting ATPases: structure of the transmembrane Cu^+ transport sites. J Biol Chem (in press)

Hatori Y, Majima E, Tsuda T, Toyoshima C (2007) Domain organization and movements in heavy metal ion pumps: papain digestion of CopA, a Cu^+-transporting ATPase. J Biol Chem 282:25213–25221. doi:10.1074/jbc.M703520200

Holm L, Park J (2000) DaliLite workbench for protein structure comparison. Bioinformatics 16:566–567. doi:10.1093/bioinformatics/16.6.566

Jacobsen RB, Sale KL, Ayson MJ, Novak P, Hong J, Lane P, Wood NL, Kruppa GH, Young MM, Schoeniger JS (2006) Structure and dynamics of dark-state bovine rhodopsin revealed by chemical cross-linking and high-resolution mass spectrometry. Protein Sci 15:1303–1317. doi:10.1110/ps.052040406

Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580. doi:10.1006/jmbi.2000.4315

Kühlbrandt W (2004) Biology, structure and mechanism of P-type ATPases. Nat Rev Mol Cell Biol 5:282–295. doi:10.1038/nrm1354

Laemmli UK, Favre M (1973) Maturation of the head of bacteriophage T4. J Biol Chem 80:575–599

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. doi:10.1093/bioinformatics/btm404

Lesk VI, Sternberg MJ (2008) 3D-Garden: a system for modelling protein-protein complexes based on conformational refinement of ensembles generated with the marching cubes algorithm. Bioinformatics 24:1137–1144. doi:10.1093/bioinformatics/btn093

Linder MC, Hazegh Azam M (1996) Copper biochemistry and molecular biology. Am J Clin Nutr 63:797S–811S

Linz R, Lutsenko S (2007) Copper-transporting ATPases ATP7A and ATP7B: cousins, not twins. J Bioenerg Biomembr 39(5–6):403–407

Lübben M, Güldenhaupt J, Zoltner M, Deigweiher K, Haebel P, Urbanke C, Scheidig AJ (2007) Sulfate acts as phosphate analog on the monomeric catalytic fragment of the CPx-ATPase CopB from Sulfolobus solfataricus. J Mol Biol 369:368–385. doi:10.1016/j.jmb.2007.03.029

Lutsenko S, Kaplan JH (1995) Organization of P-type ATPases: significance of structural diversity. Biochemistry 34:15607–15613. doi:10.1021/bi00048a001

Mandal AK, Yang Y, Kertesz TM, Arguello JM (2004) Identification of the transmembrane metal binding site in Cu^+-transporting PIB-type ATPases. J Biol Chem 279:54802–54807. doi:10.1074/jbc.M410854200

Morth JP, Pedersen BP, Toustrup-Jensen MS, Sorensen TL, Petersen J, Andersen JP, Vilsen B, Nissen P (2007) Crystal structure of the sodium-potassium pump. Nature 450:1043–1049. doi:10.1038/nature06419

Odermatt A, Solioz M (1995) Two trans-acting metalloregulatory proteins controlling expression of the copper-ATPases of Enterococcus hirae. J Biol Chem 270:4349–4354. doi:10.1074/jbc.270.16.9217

Odermatt A, Suter H, Krapf R, Solioz M (1992) An ATPase operon involved in copper resistance by Enterococcus hirae. Ann N Y Acad Sci 671:484–486. doi:10.1111/j.1749-6632.1992.tb43836.x

Ramirez DC, Mejiba SE, Mason RP (2005) Copper-catalyzed protein oxidation and its modulation by carbon dioxide: enhancement of protein radicals in cells. J Biol Chem 280:27402–27411. doi:10.1074/jbc.M504241200

Rosenzweig AC, Huffman DL, Hou MY, Wernimont AK, Pufahl RA, O'Halloran TV (1999) Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. Structure 7:605–617. doi:10.1016/S0969-2126(99)80082-3

Sazinsky MH, Agarwal S, Argüello JM, Rosenzweig AC (2006a) Structure of the actuator domain from the Archaeoglobus fulgidus Cu(+)-ATPase. Biochemistry 45:9949–9955. doi:10.1021/bi0610045

Sazinsky MH, Mandal AK, Argüello JM, Rosenzweig AC (2006b) Structure of the ATP binding domain from the Archaeoglobus fulgidus Cu^+-ATPase. J Biol Chem 281:11161–11166. doi:10.1074/jbc.M510708200

Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160:65–73. doi:10.1016/S1090-7807(02)00014-9

Solioz M, Stoyanov JV (2003) Copper homeostasis in Enterococcus hirae. FEMS Microbiol Rev 27:183–196. doi:10.1016/S0168-6445(03)00053-6

Solioz M, Vulpe C (1996) CPx-type ATPases: a class of P-type ATPases that pump heavy metals. Trends Biochem Sci 21:237–241

Solioz M, Odermatt A, Krapf R (1994) Copper pumping ATPases: common concepts in bacteria and man. FEBS Lett 346:44–47. doi:10.1016/0014-5793(94)00316-5

Strausak D, Solioz M (1997) CopY is a copper-inducible repressor of the Enterococcus hirae copper ATPases. J Biol Chem 272:8932–8936. doi:10.1074/jbc.272.14.8932

Toyoshima C, Mizutani T (2004) Crystal structure of the calcium pump with a bound ATP analogue. Nature 430:529–535. doi:10.1038/nature02680

Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. Nature 405:647–655. doi:10.1038/35015017

Toyoshima C, Nomura H, Sugita Y (2003) Crystal structures of Ca^{2+}-ATPase in various physiological states. Ann N Y Acad Sci 986:1–8

van der Sluis EO, Nouwen N, Driessen AJ (2002) SecY-SecY and SecY-SecG contacts revealed by site-specific cross-linking. FEBS Lett 527:159–165. doi:10.1016/S0014-5793 (02)03202-7

Vandevuer S, Van Bambeke F, Tulkens PM, Prevost M (2006) Predicting the three-dimensional structure of human P-glycoprotein in absence of ATP by computational techniques embodying crosslinking data: insight into the mechanism of ligand migration and binding sites. Proteins 63:466–478. doi:10.1002/prot.20892

Wernimont AK, Huffman DL, Lamb AL, O'Halloran TV, Rosenzweig AC (2000) Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins. Nat Struct Biol 7:766–771. doi:10.1038/78999

Wimmer R, Herrmann T, Solioz M, Wüthrich K (1999) NMR structure and metal interactions of the CopZ copper chaperone. J Biol Chem 274:22597–22603. doi:10.1074/jbc. 274.32.22597

Wu CC, Rice WJ, Stokes DL (2008) Structure of a copper pump suggests a regulatory role for its metal-binding domain. Structure 16:976–985. doi:10.1016/j.str.2008.02.025

Wunderli-Ye H, Solioz M (2001) Purification and functional analysis of the copper ATPase CopA of *Enterococcus hirae*. Biochem Biophys Res Commun 280:713–719. doi: 10.1006/bbrc.2000.4176

Yoshida Y, Furuta S, Niki E (1993) Effects of metal chelating agents on the oxidation of lipids induced by copper and iron. Biochim Biophys Acta 1210:81–88

Young MM, Tang N, Hempel JC, Oshiro CM, Taylor EW, Kuntz ID, Gibson BW, Dollinger G (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. Proc Natl Acad Sci USA 97:5802–5806. doi:10.1073/pnas.090099097