

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Revisiting the CooJ family, a potential chaperone for nickel delivery to [NiFe]-carbon monoxide dehydrogenase

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Elisabeth Darrouzet, C.R. (2021). Revisiting the CooJ family, a potential chaperone for nickel delivery to [NiFe]-carbon monoxide dehydrogenase. JOURNAL OF INORGANIC BIOCHEMISTRY, 225, 1-9 [10.1016/j.jinorgbio.2021.111588].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/836139> since: 2024-03-19

*Published:*

DOI: <http://doi.org/10.1016/j.jinorgbio.2021.111588>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

The carbon monoxide dehydrogenase accessory protein CooJ is a histidine-rich multidomain dimer containing an unexpected Ni(II)-binding site

**Marila Alfano<sup>1</sup>, Julien Pérard<sup>1</sup>, Philippe Carpentier<sup>1</sup>, Christian Basset<sup>1</sup>, Barbara Zambelli<sup>2</sup>, Jennifer Timm<sup>1, #</sup>, Serge Crouzy<sup>1</sup>, Stefano Ciurli<sup>2</sup> and Christine Cavazza<sup>1\*</sup>**

From the <sup>1</sup>Univ. Grenoble Alpes, CEA, CNRS, BIG, CBM, F-38000 Grenoble; <sup>2</sup>Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Via Giuseppe Fanin 40, I-40127, Bologna (Italy)

Running title: *A novel Ni(II)-binding site in the nickel chaperone CooJ*

#. Present address: Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA.

\* To whom correspondence should be addressed: [christine.cavazza@cea.fr](mailto:christine.cavazza@cea.fr),  
Tel.: +33 4 38 78 91 16  
ORCID: [0000-0002-3657-1302](https://orcid.org/0000-0002-3657-1302)

**Keywords:** chaperone, metal ion-protein interaction, small-angle X-ray scattering (SAXS), X-ray crystallography, molecular modeling, enzyme maturation, histidine-rich cluster, carbon monoxide dehydrogenase, nickel-binding protein, structural biology.

---

## ABSTRACT

Activation of nickel enzymes requires specific accessory proteins organized in multiprotein complexes controlling metal transfer to the active site. Histidine-rich clusters are generally present in at least one of the metallo-chaperones involved in Ni delivery. The maturation of carbon monoxide dehydrogenase (CODH) in the proteobacterium *Rhodospirillum rubrum* requires three accessory proteins CooC, CooT and CooJ dedicated to Ni insertion into the active site, a distorted [NiFe<sub>3</sub>S<sub>4</sub>] cluster coordinated to an Fe site. Previously, CooJ from *R. rubrum* (RrCooJ) has been described as a Ni chaperone with 16 histidines and 2 cysteines at its C-terminus. Here, the X-ray structure of a truncated version of RrCooJ, combined with small-angle X-ray scattering (SAXS) data and a modelling study of the full-length protein, revealed a homodimer comprising a coiled-coil with two independent and highly flexible His tails. Using isothermal calorimetry, we characterized several metal-binding sites (four per dimer), involving the His-rich motifs and having similar metal affinity ( $K_D = 1.6 \mu\text{M}$ ). Remarkably, biophysical approaches, site-directed mutagenesis and X-ray crystallography uncovered an additional Ni-binding site at the dimer interface, that binds Ni(II) with an affinity of 380 nM. Although RrCooJ was initially

thought to be a unique protein, a proteome database search identified at least 46 bacterial CooJ homologs. These homologs all possess two spatially separated nickel-binding motifs: a variable C-terminal histidine tail and a strictly conserved “H-(W/F)-X<sub>2</sub>-H-X<sub>3</sub>-H” motif, identified in this study, suggesting a dual function for CooJ both as a Ni chaperone and as a Ni storage protein.

---

*Rhodospirillum rubrum* is a purple non-sulfur photosynthetic bacterium, able to obtain energy through a variety of mechanisms depending on the growth conditions. This versatile bacterium is a facultative anaerobe, capable of using CO as a sole energy source during anaerobic growth in darkness through the water-gas shift (WGS) reaction ( $\text{H}_2\text{O} + \text{CO} \rightarrow \text{H}_2 + \text{CO}_2$ ) (1). The biologically-mediated WGS reaction is catalyzed by two nickel-dependent enzymes, namely a monofunctional [NiFe]-carbon monoxide dehydrogenase (CODH) coupled to a CO-tolerant energy-conserving [NiFe]-hydrogenase (2). CO is oxidized to CO<sub>2</sub> by CODH, releasing electrons used for the reduction of two protons to yield H<sub>2</sub>, the latter reaction being catalyzed by the hydrogenase. This CO-oxidizing complex contains also CooF,

a ferredoxin that transfers electrons from CODH to hydrogenase. As a result, the implication of these two essential Ni-enzymes in CO metabolism requires a highly controlled and effective nickel supply to ensure their correct activation and catalytic efficiency.

The crystal structures of CODH from *R. rubrum* (*RrCODH*) (3) and *Carboxydotherrmus hydrogenoformans* (4) were first reported in 2001, revealing the unique architecture of its active site, called C-cluster, constituted of a [NiFe<sub>3</sub>S<sub>4</sub>]-cluster and a mononuclear Fe(II) site. Previous studies have shown that the “Fe/S unit” is inserted prior to nickel. Indeed, when *R. rubrum* is grown in nickel-depleted media, the purified *RrCODH* corresponds to a stable, but inactive, nickel-deficient form (5),(6). The addition of nickel salts to nickel-free CODH in the presence of CO under reductive conditions is sufficient to convert the Ni-free C-cluster into a functional C-cluster, proving that nickel insertion is the key step of CODH activation (7). However, the mechanism of nickel import, trafficking and accumulation prior to its insertion into CODH remains a subject that has been poorly investigated to date.

In *R. rubrum*, a CODH operon, named *cooFSCTJ* (8), is located downstream the hydrogenase operon, named *cooMKLXUH* (9). Both operons are CO-induced via the CO-sensing transcriptional activator, *CooA*, found downstream of the *cooFSCTJ* operon (10). The five open-reading frames present in the CODH operon encode the ferredoxin *CooF*, CODH (*CooS*), and three additional nickel-dependent proteins (*CooC*, *CooT* and *CooJ*) shown to function specifically for nickel insertion into CODH (8). However, the precise role of the latter three proteins in the maturation pathway, from nickel acquisition to its transfer into the enzyme, is not well understood.

*CooC* (11) is widely distributed in microorganisms possessing a [NiFe]-CODH. The protein is an ATPase, analogous to the GTPases *HypB* (12) and *UreG* (13) required for nickel-processing for hydrogenase and urease, respectively. Today, NTPases are generally recognized as common components of the maturation processes of nickel enzymes, potentially playing a regulatory role by affecting protein-protein interactions in multimeric chaperonin complexes and/or modulating nickel ion affinity for intermediate metal binding sites.

*CooT* from *R. rubrum* (*RrCooT*) was recently shown to be a homodimer of 14 kDa that

specifically binds one Ni(II) per dimer (14). Initially identified in *R. rubrum* and *Rhodopseudomonas palustris* genomes, phylogenetic analyses have identified 111 *CooT* homologues in anaerobic bacteria and Archaea (14). In all cases, the presence of *CooT* is related to an anaerobic and metal-related metabolism.

*RrCooJ* is a metallochaperone with only few biochemical data and no structural information available to date (15). This 12.5 kDa protein contains a C-terminal nickel-binding domain with 16 histidines and 2 cysteines in the final 34 amino acids (**Fig. 1A**). The protein purified directly from *R. rubrum* cultures was reported to bind four Ni(II) per monomer with a *K<sub>D</sub>* value of 4.3  $\mu$ M, as shown by equilibrium dialysis (15). During the purification, *RrCooJ* co-eluted with *RrCODH* and other additional proteins, suggesting the formation of a multiprotein complex bound to CODH (15). Interestingly, the histidine-rich region of *RrCooJ* is partially dispensable for its physiological function: a truncated version of *RrCooJ* with only 6 histidines and no cysteine in the C-terminal part is still able to activate CODH *in vivo*, suggesting that the His-tail could play another role in the bacterium, in addition to CODH maturation (8). Ni(II) transport occurs in *R. rubrum* grown either in the presence or in the absence of CO, excluding a role for *CooC*, *CooT* and *CooJ* in Ni(II) import. However, <sup>63</sup>Ni(II) accumulates more rapidly in CO-induced cells, essentially because of nickel accumulation in CODH, which is highly dependent on the presence of *CooC*, *CooT* and *CooJ*, proving that all three proteins must be present for *in vivo* nickel insertion into CODH (16).

Here, we report a detailed characterization of *RrCooJ*, which reveals the existence of an unexpected nickel-binding site in the N-terminal region. Biophysical approaches using the wild-type protein and site-directed mutants allowed the determination of nickel-binding sites in the N-terminal region and in the C-terminal histidine-rich tail. Structural analyses using X-ray crystallography, small-angle X-ray scattering (SAXS) and modelling provided detailed information on the structure of *RrCooJ* and allowed the characterization of the novel Ni(II)-binding site. Based on the presence of two spatially separated Ni(II)-binding sites, this study suggests a double function for *RrCooJ* as a Ni(II) chaperone and a Ni(II) storage protein.

## Results

***RrCooJ is an  $\alpha$ -helical dimeric protein***

Wild type *RrCooJ* is a 12,581 Da protein (comprising Met1) possessing a histidine-rich tail that features 16 histidines and 2 cysteines. The recombinant protein (12,507Da) (**Fig. 1A**) was overproduced in *E. coli* and purified to homogeneity, as confirmed by ESI-MS (**Fig. S1**). Size-exclusion chromatography coupled with multiple-angle laser light scattering and refractive index (SEC-MALLS-RI) reveals a homo-dimeric state of the apoprotein in solution, with the peak eluting at ~25 kDa (**Fig. 2A**). The results of secondary structure prediction using PredictProtein (**Fig. 1B**), show four  $\alpha$ -helical motifs in the N-terminal region from residue 4 to 68, followed by a mostly disordered region from 69 to 115, as expected from its high content in histidines. However, a weak  $\alpha$ -helical and  $\beta$ -strand secondary structure propensity is also expected in the C-terminal portion. The circular dichroism (CD) spectrum of apo*RrCooJ* in the far-UV region is consistent with this prediction (**Fig. 2B**). A truncated form of the protein, lacking the C-terminal region from Asp69 to Pro115 (*RrCooJ*- $\Delta$ , **Fig. 1A**), was also constructed, as confirmed by ESI-MS (**Fig. S1B**). Similar to *RrCooJ*, *RrCooJ*- $\Delta$  behaved as a homodimer in solution, as shown by SEC-MALLS-RI (**Fig. 2A**) with an eluting peak at ~15 kDa, and it mainly contains  $\alpha$ -helices as revealed by the far-UV CD spectrum (**Fig. 2B**).

***The nickel binding properties of RrCooJ and its mutants in vitro***

ICP-AES measurements showed that purified *RrCooJ* and its mutants were devoid of any bound metal ion (**Table S2**). CD and Isothermal titration calorimetry (ITC), coupled with site-directed mutagenesis, were used to investigate the Ni(II)-binding properties of *RrCooJ*. CD spectra in the near-UV/visible range revealed the appearance of a ligand-to-metal charge transfer (LMCT) transition at 270 (-) nm, with saturation at ~ 7 Ni(II) molar equivalent per dimer upon addition of Ni(II) to *RrCooT*. A second signal at 330 (+) nm appeared when the Ni(II) concentration reached 4 molar equivalents per dimer, and saturated at about 7 molar equivalents per dimer (**Fig. 3A**). The signal at 330 (+) nm most likely corresponds to a Ni-thiolate LMCT, because it is not observed in the CD spectrum of the *RrCooJ*-C2S mutant (**Fig. 3B**), in which Cys109 and Cys111 are mutated to serine (**Fig. 1A**). However, the 270 (-) nm signal is still present in the CD spectrum of *RrCooJ*-C2S, even though at

lower intensity (**Fig. 3B**). This suggests that ~ 4 Ni(II) ions initially bind to the dimeric *RrCooJ* via histidine residues, followed by the binding of additional Ni(II) ions with the involvement of the two cysteine residues. The truncation of the histidine-rich tail in *RrCooJ*- $\Delta$ , does not completely abolish Ni(II)-binding, as indicated by the appearance of the CD signal at 270 (-) nm, saturating at 1 Ni(II) molar equivalent per dimer upon Ni(II) addition (**Fig. 3C**). To investigate the role of the three putative nickel-coordinating residues present in the N-terminal part of *RrCooJ*, namely His18, His22 and His26, the triple mutant *RrCooJ*- $\Delta$ 3HA (**Fig. 1A**) was constructed and purified. This mutant does not show any CD signal (**Fig. 3D**), demonstrating that the observed transition band in the dimeric *RrCooJ*- $\Delta$  mutant is due to the coordination of a single Ni(II) to His18, His22 and/or His26.

ITC was applied to measure the Ni(II) binding properties of *RrCooJ* and its variants. In all cases but *RrCooJ*- $\Delta$ 3HA, Ni(II) titrations produce negative peaks, indicating the occurrence of an exothermic binding reaction. A fit of the integrated heat data for *RrCooJ* (**Fig. 4A, S2A**) using the “one set of sites” model was not satisfactory ( $\chi^2_v = 2.01 \text{ kJ}^2 \text{ mol}^2$ ) as the curve showed two inflection points, suggesting the presence of two Ni(II) binding events. Therefore, the “two sets of sites” model was applied with significantly improved statistics ( $\chi^2_{vWT} = 0.453 \text{ kJ}^2 \text{ mol}^2$ ). The results of the fit indicate that one Ni(II) ion binds to the protein dimer with higher affinity ( $K_{D1WT} = 380 \pm 30 \text{ nM}$ ), followed by the interaction of ~ 4 Ni(II) ions with lower affinity ( $K_{D2WT} = 1.60 \pm 0.02 \text{ }\mu\text{M}$ ). Both events are characterized by favorable enthalpic values ( $\Delta H_{1WT} = -33 \pm 3 \text{ kJ mol}^{-1}$  and  $\Delta H_{2WT} = -54.4 \pm 0.7 \text{ kJ mol}^{-1}$ ) and different entropic contributions ( $\Delta S_{1WT} = +12.9 \text{ J mol}^{-1} \text{ K}^{-1}$  and  $\Delta S_{2WT} = -71.5 \text{ J mol}^{-1} \text{ K}^{-1}$ ). As observed for CD titrations, ITC experiments on *RrCooJ*- $\Delta$  confirmed that the truncated protein retains Ni(II) binding properties. However, differently from the full-length protein, the integrated heat data show a single inflection point, indicating the occurrence of a single binding event (**Fig. 4B, S2B**). A fit of the integrated data with the “one set of sites” model ( $\chi^2_{v\Delta} = 1.20 \text{ kJ}^2 \text{ mol}^2$ ) shows that a single Ni(II) ion binds to the protein dimer with  $K_{D\Delta} = 13 \pm 2 \text{ nM}$ ,  $\Delta H_{\Delta} = -36.3 \pm 0.4 \text{ kJ mol}^{-1}$ ,  $\Delta S_{\Delta} = +29.3 \text{ J mol}^{-1} \text{ K}^{-1}$ . The calculated thermodynamic parameters are similar to those provided for the first binding event in *RrCooJ*, suggesting that the latter does not involve the C-terminal tail, while

the subsequent binding of four Ni(II) ions, absent in the *RrCooJ*- $\Delta$  protein, occurs at the C-terminal His-rich tail. These results are consistent with the CD-based titrations and provide improved stoichiometry for the different events as well as a quantitative determination of the equilibrium binding constants. The affinity of the internal site in the truncated protein is one order of magnitude higher as compared to the same site in the full length protein, suggesting that the C-terminal region somehow decreases the Ni(II) affinity for the internal binding site, possibly competing for Ni(II) binding. Mutation of the two Cys residues (*RrCooJ*-C2S mutant) (**Fig. 4C, S2C**), led to a decreased stoichiometry from  $\sim 5$  to  $\sim 4$  Ni(II) ions per protein dimer. In particular, the thermodynamic parameters obtained with a fit using a “two sets of sites model” ( $\chi^2_{C2S} = 1.72 \text{ kJ}^2 \text{ mol}^2$ ) are  $K_{D1C2S} = 380 \pm 10 \text{ nM}$ ,  $\Delta H_{1C2S} = -62 \pm 2 \text{ kJ mol}^{-1}$ ,  $\Delta S_{1C2S} = -84.2 \text{ J mol}^{-1} \text{ K}^{-1}$  and  $K_{D2C2S} = 1.07 \pm 0.01 \text{ }\mu\text{M}$ ,  $\Delta H_{2C2S} = -82 \pm 2 \text{ kJ mol}^{-1}$ ,  $\Delta S_{2C2S} = -159 \text{ J mol}^{-1} \text{ K}^{-1}$ . On the other hand, mutation of the three internal histidines (His18, 22 and 26) produces the *RrCooJ*-3HA variant that lacks the N-terminal binding site, and only retains the ability to bind four Ni(II) per protein dimer (**Fig. 4D, S2D**). Coherently, the binding isotherm shows a single inflection point and can be fitted with a “one set of sites” model ( $\chi^2_v = 1.01 \text{ kJ}^2 \text{ mol}^2$ ) and with  $K_{D3HA} = 940 \pm 40 \text{ nM}$ ,  $\Delta H_{3HA} = -88.7 \pm 0.6 \text{ kJ mol}^{-1}$ ,  $\Delta S_{3HA} = -182 \text{ J mol}^{-1} \text{ K}^{-1}$ . Finally, as observed with CD, the triple mutant *RrCooJ*- $\Delta$ 3HA does not feature any significant heat of binding (**Fig. 4E, S2E**), further confirming that the first binding event involves Ni(II) coordination to the histidine residues of the folded N-terminal domain.

In order to study the impact of Ni(II) on *RrCooJ* conformation, SEC-MALLS-RI experiments were conducted. Increasing addition of Ni(II) from 0 to 4 Ni(II) eq. per protein dimer led to elution volume variations, while the molecular masses agree with a dimeric protein, regardless of Ni(II) concentration. This suggests an effect of metal binding on the protein conformation (**Fig. 5A**). Above  $\sim 100 \text{ }\mu\text{M}$   $\text{NiSO}_4$ , a slight modification of the total injected protein was observed, suggesting protein aggregation at high Ni(II) concentrations. Notably, the addition of 5 mM EDTA to a *RrCooJ* solution pre-incubated with 3 molar eq. of Ni(II) led to the recovery of the apo*RrCooJ* dimeric elution peak, demonstrating that Ni(II) binding and nickel-induced conformational change of *RrCooJ* are reversible processes (**Fig. 5B**). In the case of

*RrCooJ*- $\Delta$ , the elution volume is independent from Ni(II) concentration (**Fig. 5C**), revealing that the conformational change of *RrCooJ*-WT is induced by Ni(II)-binding to the His-rich tail.

### *X-ray structure of RrCooJ- $\Delta$ in complex with Ni(II)*

The crystal structure of Ni(II)-loaded *RrCooJ*- $\Delta$  (Ni-*RrCooJ*- $\Delta$ ) was solved at 2.04-Å resolution in the space groups P2<sub>1</sub>. Crystallographic data statistics of the Ni-*RrCooJ*- $\Delta$  structure are summarized in **Table S3**. In the structure, 8 molecules are present in the asymmetric unit. Ni-*RrCooJ*- $\Delta$  forms a homodimer (**Fig. 6A**) in agreement with the SEC-MALLS-RI analyses. Each monomer is composed of two  $\alpha$ -helices (from Glu3 to Ala36 and from Leu43 to Leu68) connected by a short turn (from Leu37 to Arg42). The homodimer adopts a coiled-coil architecture comprising four helices, with the two monomers arranged antiparallel to each other, indicating that the two histidine-tails of the full-length protein are expected to be top-to-tail. The degree of mobility (B-factor) varies along the helix bundle with a central more rigid region and two flexible edges (**Fig. 6B**). Remarkably, four methionine residues (Met 27A/B and Met 54A/B) form a methionine-rich cluster (**Fig. 6A**). This kind of cluster was previously observed in helix-bundle structures (17),(18) and plays a role in helix-helix packing patterns. According to the CATH classification (structural classification UCL), the *RrCooJ*- $\Delta$  monomer belongs to the “mainly alpha” class, with a helix hairpin topology (CATH code: 1.10.287). *RrCooJ*- $\Delta$  shares structural similarities with the “Helix hairpin bin” superfamily, although no function can be deduced from it. Moreover, the investigation of the function of *RrCooJ* from its sequence revealed that the protein does not belong to any CATH functional family and that no homologue domain was found in the PFAM database. Assembly analyses using PDBePISA (19) showed that the 8 amino-acid chains present in the asymmetric unit form 4 stable dimers, with interface surface areas of about 7000 Å<sup>2</sup>. One Ni(II) ion is bound to each dimer. In addition, two BCEG-Ni<sub>2</sub> and ADFH-Ni<sub>2</sub> tetramers, predicted to be unstable in solution, were suggested.

### *Structural characterization of the novel nickel-binding site*

The X-ray structure of Ni-*RrCooJ*- $\Delta$  allowed the identification and the characterization of the Ni(II)-binding site, found in a largely solvent-

exposed site. One peak attributed to nickel is present in the X-ray fluorescence spectrum (**Fig. S3**), corresponding to four peaks with Ni(II) at 100% occupancy in the anomalous difference electron density map, at the surface of the four dimers. For clarity, only the BC dimer will be discussed here. The Ni(II) ion is hexacoordinated and adopts an octahedral coordination. The metal is bound to the dimer via the four N $\epsilon$  of imidazole groups of His22<sub>B</sub>, His26<sub>B</sub>, His18<sub>C</sub> and His22<sub>C</sub>. The two additional solvent-exposed coordination sites are filled by the bidentate carboxylate group of Glu29<sub>G</sub> (**Fig. 7A-C**). Thus, in the X-ray structure, a tetramer “BCEG” is formed via the interaction of Glu29<sub>G</sub> with the Ni(II) ion bound to the dimer BC. In addition, a  $\pi$ -stacking between His18<sub>B</sub> and His18<sub>E</sub> is also observed (**Fig. 7A**). However, these interactions are insufficient to stabilize the tetrameric conformation in solution, as proven by SEC-MALLS-RI and should be regarded as a solid-state effect. This conclusion is reinforced by ITC experiments on the *RrCooJ*- $\Delta$ E29A mutant, which indicate that mutation of Glu29 to Ala does not affect Ni(II) binding in solution (**Fig. 4F, S2F**). Indeed, the fit of the binding isotherm ( $\chi^2_{AE29A} = 0.325 \text{ kJ}^2 \text{ mol}^2$ ) shows that one Ni(II) ion binds per protein dimer with thermodynamic parameters similar to those of *RrCooJ*- $\Delta$  ( $K_{DAE29A} = 18 \pm 2 \text{ nM}$ ,  $\Delta H_{AE29A} = -37.6 \pm 0.3 \text{ kJ mol}^{-1}$ ,  $\Delta S_{AE29A} = 22.4 \text{ J mol}^{-1} \text{ K}^{-1}$ ).

Unexpectedly, CooJ homologues are present in five different bacterial phyla, with size varying from 70 to 117 residues: they were identified in the phyla *Proteobacteria*, *Planctomycetes*, *Lentisphaerae*, *Chloroflexi* and *Nitrospirae*, and in groundwater and marine sediment metagenomes. In *proteobacteria*, they are found in alphaproteobacteria (*R. rubrum*, *Rhodopseudomonas palustris*, *Pleomorphomonas* sp., *Telmatospirillum siberiense*), in unclassified Deltaproteobacteria and in sulfate-reducing deltaproteobacteria (unclassified *Desulfobulbaceae*, unclassified *Desulfobacterales*, *Desulfofustis glycolicus*, *Desulforhopalus singaporensis*, *Desulfurivibrio alkaliphilus*, *Desulfobulbus* sp., *Desulfotalea* sp., *Desulfopila aestuarii*, *Desulfocapsa sulfexigens*). In *planctomycetes*, they are found in unclassified *Planctomycetes* and in *Candidatus scalindua*. In *Lentisphaerae*, *Chloroflexi* and *Nitrospirae*, they were identified in unclassified bacteria.

The Ni(II)-binding residues His18, His22 and His26 in the *R. rubrum* sequence are strictly conserved among at least 46 putative CooJ sequences (**Fig. S4**).

Moreover, an acidic residue (Asp or Glu) is also conserved (in position 29 in *R. rubrum*). The new “H-(W/F)-X<sub>2</sub>H-X<sub>3</sub>H” Ni(II)-binding motif is revealed by the present study as a signature of the CooJ family. This result highlights the importance of the newly characterized Ni(II)-binding site in *RrCooJ*. The alignment of their amino-acid sequences by ClustalW further shows that they all possess a histidine-rich motif in the C-terminal part, although their lengths are variable, containing from 3 to 18 histidines, while cysteines are rarely present (**Fig. 7D**).

#### Analysis of apo*RrCooJ*- $\Delta$ and apo*RrCooJ* by SAXS

As expected, the presence of a large disordered region in the C-terminal part of *RrCooJ* hindered its crystallization. In order to get information on the conformation of the full-length protein in solution, SAXS experiments were performed. The final scattering curves of apo*RrCooJ* and apo*RrCooJ*- $\Delta$  are shown in **Figure 8A**. The linearity of the Guinier plots at several protein concentrations confirms that both proteins are monodisperse in solution (**Fig. 8B**). The radius of gyration ( $R_g$ ) for apo*RrCooJ* ( $R_g$ : 3.05  $\pm$  0.12 nm) and apo*RrCooJ*- $\Delta$  ( $R_g$ : 1.99  $\pm$  0.05 nm) were measured. Moreover, the calculated molecular weight determined by I(0) is in good agreement with SEC-MALLS-RI data ( $MW_{SAXS}$ =26  $\pm$  0.5 kDa for apo*RrCooJ* and  $MW_{SAXS}$ =14.9  $\pm$  0.1 kDa for apo*RrCooJ*- $\Delta$ ). The pair distance distribution patterns of both proteins indicate that apo*RrCooJ* with a  $D_{max}$  of 12  $\pm$  1 nm, is more elongated than apo*RrCooJ*- $\Delta$  that displays a  $D_{max}$  of 6.5  $\pm$  0.4 nm (**Fig. 8C**). The flexibility of each protein was investigated with the analysis of the Kratky plots ( $s^2 I(s)$  vs.  $s$ ). The bell-shaped curves of the plots indicate that both proteins are partially folded (**Fig. 8D**). However, apo*RrCooJ* shows a decays to zero with a  $Q_{max}$  of 0.7 nm<sup>-1</sup>, indicating the presence of more flexible parts compared to apo*RrCooJ*- $\Delta$  with a  $Q_{max}$  of 1 nm<sup>-1</sup>. In addition, the lower intensity of the plateau phase of apo*RrCooJ*- $\Delta$  compared to that of apo*RrCooJ* strongly suggests the presence of additional flexible regions in the full-length protein. Taken together, these data revealed that apo*RrCooJ* is less structured than apo*RrCooJ*- $\Delta$ .

In order to determine the three-dimensional model of apo*RrCooJ* in solution, *ab initio* models were generated without any constraint in slow mode. First, *ab initio* models of apo*RrCooJ*- $\Delta$  were built. A solution model was selected by

DAMAVAR (21), with normalized spatial discrepancy (NSD) values  $< 0.7$ , corresponding to a homogenous solution in agreement with a compact and folded protein (**Fig. 8E**). To go further, *ab initio* models of the full-length protein were generated using the same procedure. In this case, a heterogeneous solution was obtained, with the presence of several clusters with NSD values varying from 0.8 and 1.5, in favor of the existence of flexible regions in apo*RrCooJ* (**Fig. 8F**). Taking advantage of the X-ray structure of Ni-*RrCooJ*- $\Delta$ , the SAXS experimental scattering curve of apo*RrCooJ*- $\Delta$  was fitted with the theoretical scattering curve from the crystal structure generated by CRY SOL, giving a  $\chi^2$  of 1.8 (**Fig. 8F**). Moreover, the Ni-*RrCooJ*- $\Delta$  X-ray structure was superimposed with the best *ab initio* model of apo*RrCooJ*- $\Delta$  using SUPCOMB (**Fig. 8G**). Our data show that the truncated protein shares the same conformation in solution and in the crystalline state.

### Molecular modeling of apo*RrCooJ*

In order to better characterize the apo*RrCooJ* dimer in solution, a hybrid approach was adopted combining both the Ni-*RrCooJ*- $\Delta$  X-ray structure and the SAXS data using the Ensemble Optimization Method (EOM 2.0) (20). From our SAXS analyses, the X-ray structure of the truncated holo-protein can be confidently used to model the N-terminal region of the full-length protein. EOM 2.0 was thus used to generate a pool of 10,000 conformers using both random chains for the C-terminal part and the atomic coordinates of the Ni-*RrCooJ*- $\Delta$  crystal structure. The Gaussian statistics of the dataset resulted in a best final ensemble of five independent models (a  $\chi^2 = 0.82$ ) in a highly variable range of  $D_{\max}$  from 8.3 to 13.1 nm, describing the scattering properties of inherently dynamic macromolecules (**Fig. 9A-B**). This result indicates that the apo*RrCooJ* dimer contains two highly flexible regions corresponding to the two C-terminal histidine rich clusters modeled as poly-asparagine by EOM 2.0 (**Fig. 9C**). The anti-parallel spatial arrangement of the two monomers, observed in the X-ray structure of Ni-*RrCooJ*- $\Delta$  dimer, is confirmed for the full-length protein in solution.

In order to obtain more detailed information on the flexible histidine-rich C-terminal tails, a pool of individual models of apo*RrCooJ* was built using the CHARMM force field and the *RrCooJ* amino acid sequence, with the atomic coordinates of Ni-*RrCooJ*- $\Delta$  structure

and the biophysical parameters ( $R_g$ ,  $D_{\max}$ ) from SAXS data. To generate individual models, the central ordered domain of *RrCooJ* dimer was considered similar to the *RrCooJ*- $\Delta$  dimer. Each model was treated using dynamic replica exchange (RXGLD) between 300 and 400 K covering 8 ns. The last 3 ns of RXGLD dynamics (3000 frames recorded every ps) were analyzed. The calculated average  $R_g$  was  $3.14 \pm 0.04$  nm compatible with the 3.12 nm target determined by SAXS, meaning that the modeled trajectory yields a correct representation of the flexibility of the C-terminal region. Frames were grouped into nine clusters corresponding to coordinate RMSD larger than 6.0 Å. The minimum energy models of each cluster are shown superimposed in **Fig. 9D**, indicating the appearance of transient  $\alpha$ -helix and  $\beta$ -sheet motifs in the C-terminal region. In particular,  $\alpha$ -helices appear between His92 and Cys109, while transient  $\beta$ -sheets are found in the Leu68-Asp69, Pro84-Phe85, His96-Gly99, His104-Asp107, and His110-Asp113 regions.

### Discussion

Nickel is an important trace element for many prokaryotic microorganisms, where it participates to a variety of cellular processes, being part of the active site of essential enzymes. As the free metal ion is toxic in the intracellular environment, the cells control its availability with transporters, storage proteins and metalloregulators, working in concert with accessory proteins to deliver the metal ion to its specific target sites. In *R. rubrum*, two key enzymes involved in CO metabolism, namely hydrogenase and CODH, require nickel for their activity. Among the specific accessory proteins expressed by this bacterium, *RrCooJ* is an active player in nickel homeostasis and, more specifically, in CODH maturation. In this work, the use of a multidisciplinary approach provided a detailed characterization of the nickel chaperone *RrCooJ*.

*RrCooJ* forms a multidomain dimeric protein, with a coiled coil and two flexible regions. A combination of X-ray crystallographic and SAXS data revealed that the two histidine-rich tails at the C-termini are independent to each other. The role of histidine-rich regions in proteins is generally proposed to be related to nickel storage and/or detoxification, due to their ability to quickly bind and release nickel ions. When present, histidine-rich clusters are located either at the C-terminal position as in *CooJ* for CODH maturation, in *UreE* for urease maturation

(21) and in SlyD for hydrogenase maturation (22), or at the N-terminal region as in HypB for hydrogenase maturation (23), in UreG for urease maturation (24) and in LarC for lactate racemase maturation (25). Differently, *Helicobacter pylori* expresses two unique proteins, Hpn and Hpn-2, composed by histidines and cysteines for half of the total amino-acid sequence. The two proteins play a central role in controlling the intracellular nickel trafficking and participate in a common pathway to the maturation of urease, a virulence factor (26). All together, these studies suggest that histidine-rich motifs might satisfy a nickel storage role for the bacteria rather than being part of the mechanism of metal insertion into enzymes.

SAXS and modelling experiments showed that the C-terminal region of *RrCooJ*, initially predicted from the sequence as intrinsically disordered, is partly and transiently structured even in the absence of nickel. This region appears to be highly flexible and linked to a more rigid N-terminal domain corresponding to the coiled coil. The two His-rich regions are able to bind four Ni(II) ions per dimer with dissociation constants in the micromolar range. The effect of Ni(II) binding to the C-terminus tail onto the protein structure is difficult to evaluate because of the presence of several metal binding sites with similar affinity and the tendency of the protein to aggregate at high nickel concentrations.

In addition to the characterization of Ni(II) binding to the His-rich tail of *RrCooJ*, an unexpected Ni(II)-binding site was detected with a dissociation constant in the nanomolar range. This site involves conserved histidines located at the dimer interface of the protein. In most nickel-dependent accessory proteins, an exposed nickel-binding site is found, compatibly with fast metal transfer to the physiological partner(s). The Ni(II) ion is coordinated by four histidines, two from each monomer, with two additional ligands from a bidentate Glu29 of the neighboring dimer completing an octahedral coordination geometry. With a solvent exposed Ni(II)-binding site, water molecules can easily replace the glutamate residue and complete the coordination sphere in solution, with a N/O ligand environment.

With two independent types of Ni(II)-binding sites, *RrCooJ* thus resembles the known behavior of HypB and UreE nickel-chaperones. Indeed, in addition to the non-conserved histidine-rich cluster, two metal-binding sites have been identified in HypB: a conserved low-

affinity Ni(II)/Zn(II) binding site, proposed to be responsible for the direct Ni(II) transfer to its physiological partner HypA (27), and a non-conserved high affinity Ni(II)-binding site at the N-terminus (28). Similarly, the dimeric UreE presents a conserved metal-binding site at the dimer interface, involving a conserved His residue from each monomer located in the C-terminal domains, as well as a second His residue in the disordered C-terminal tail, in addition to other metal binding sites located in the histidine-rich segment, found in some UreE homologues (29).

Using the ever continuously updating of genome databases, several *CooJ* homologues were identified in recently sequenced genomes. They all present a histidine-rich tail, featuring highly variable size and composition, while the strict conservation of the N-terminal binding site reinforces its essential role in nickel metabolism. While its role was predicted to be minor and restricted to *R. rubrum*, the occurrence of *CooJ* in various bacteria also raises the question of its implication in the overall maturation pathway of CODH.

Altogether, this study has provided a structural and biophysical characterization of *RrCooJ* having more general implications in the field of nickel trafficking by metallo-chaperones containing His-rich regions. On the basis of the knowledge acquired on nickel chaperones, a double function can be proposed for *CooJ*, in both CODH maturation and nickel storage. Further studies, required to clarify the molecular details of protein-protein interactions involving *CooJ*, *CooT* and *CooC* leading to Ni(II) transfer and CODH activation, are underway in our laboratories.

## Experimental procedures

*RrCooJ* and its mutants were produced in *E. coli* and purified. Cell pellets from 2 L of *E. coli* culture over-expressing *RrCooJ*, *RrCooJ*-3HA or selenomethionine-labeled *RrCooJ*- $\Delta$  were thawed and resuspended in buffer QA (50 mM Tris-HCl pH 8.5, containing 1 mM EDTA and 5 mM DTT) complemented with one complete Protease Inhibitor cocktail tablet (Roche) per 50 mL buffer, and lysed by sonication. The cell debris were removed by centrifugation at 40,000 rpm for 30 min at 4 °C. The supernatant was loaded at 0.7 mL min<sup>-1</sup> onto a Q-sepharose HP 16/10 column (GE Healthcare) equilibrated with buffer QA. After loading, the column was washed with QA buffer until the



baseline was reached, and the protein was eluted with a linear gradient (0-50% in six column volumes) of buffer QB (50 mM Tris-HCl pH 8.5, 1 M NaCl, 1 mM EDTA, 5 mM DTT). The fractions of the main elution peak were analyzed by SDS-PAGE and those containing *RrCooJ* proteins were pooled and concentrated. The protein was then loaded onto a Superdex75 16/600 size exclusion column (GE Healthcare) equilibrated in buffer S (50 mM HEPES pH 7.5, containing 300 mM NaCl, 1 mM TCEP). Fractions were collected and analyzed by SDS-PAGE. The fractions containing *RrCooJ* were pooled and concentrated up to 10.0 mg mL<sup>-1</sup> before flash freezing in liquid nitrogen and stored at -80 °C. *RrCooJ*-Δ, *RrCooJ*-Δ3HA, *RrCooJ*-2CS and *RrCooJ*-Δ-E29A were purified using the same protocol except that DTT and TCEP were excluded from Buffers QA, QB and S, respectively. Experimental extinction coefficient at 280 nm was calculated from the concentrations determined by SEC-MALLS-RI on three independent samples for apo*RrCooJ* (4.62 mM<sup>-1</sup>·cm<sup>-1</sup>) and apo*RrCooJ*-Δ (1.52 mM<sup>-1</sup>·cm<sup>-1</sup>).

Inductively coupled plasma atomic emission spectroscopy (ICP-AES) (Shimadzu ICP 9000 instrument with Mini plasma Torch in axial reading mode) was used to measure the metal content. Samples were incubated in 65% HNO<sub>3</sub> overnight at 95°C before measurement(30).

Secondary structure elements of *RrCooJ* were predicted using the PredictProtein program (<http://www.predictprotein.org>), an automatic server that searches up-to-date public sequence databases, creates alignments, and predicts aspects of protein structure and function.

Purified and frozen *RrCooJ* and *RrCooJ*-Δ were thawed, diluted to ~50 μM (dimer), and incubated in the absence and presence of NiSO<sub>4</sub> (0.5, 1.0, 2.0, 3.0 or 4.0 molar equivalents per dimer) for 15 min at room temperature prior to injection onto the SEC-MALLS-RI system (Wyatt Dawn HELEOS-II 18-angle light scattering detector and Wyatt Optilab rEX refractive index monitor linked to a Shimadzu HPLC system comprising a LC-20AD pump, a SPD20A UV/Vis detector, and a Superdex 200 10/300 column (GE Healthcare)). Injections were carried out using a 20 μL loop. The size exclusion column was equilibrated using a 50 mM HEPES pH 7.5, 300 mM NaCl and 1 mM TCEP. Protein concentration in all samples was determined by integration of the differential refractive index

(dRI) peak. It is important to note that all used samples are coming from the same stock dilution, thus a modification on the final protein concentration can be attributed to protein aggregation. The data were analyzed using the ASTRA software (version 6).

CD spectra were recorded using a J-1500 circular dichroism spectrometer from JASCO Analytical Instruments. A stock solution of 10 mM NiSO<sub>4</sub> was used to monitor metal-ligand charge transfer and d-d transitions of *RrCooJ* and its mutants upon Ni(II) titration (from 0 to 10 molar equivalents of NiSO<sub>4</sub>). Spectra were recorded from 250 to 700 nm using a 10 mm cuvette, with five accumulations to increase the signal-to-noise ratio. Proteins were thawed and diluted to 20-100 μM dimer in CD buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP). CD spectra were also recorded using a 1 mm cuvette in the 190-250 nm range to determine the secondary structure content of apo*RrCooJ* and apo*RrCooJ*-Δ, with ten accumulations on samples containing 7.5 μM protein dimer in CD buffer (8 mM HEPES pH 7.2, 10 mM NaCl and 1 mM TCEP). The spectra were analyzed using the JASCO spectra manager software (version 2).

*RrCooJ* and its variant (7-10 μM dimer), and *RrCooJ*-Δ and its variants (12-17 μM dimer), diluted in 50 mM HEPES pH 7.2 300 mM NaCl, were titrated with NiSO<sub>4</sub> (100-500 μM) diluted in the same buffer. The data were analyzed using the MicroCal PEAQ-ITC analysis software to provide the isotherm of binding and to derive the thermodynamic parameters. Integrated data obtained for each titration were fitted using a nonlinear least-squares minimization algorithm to a theoretical titration curve, with either the “one set of sites” or the “two sets of sites” models. N (stoichiometry of binding), ΔH (reaction enthalpy change, cal mol<sup>-1</sup>) and K<sub>a</sub> (binding constant, M<sup>-1</sup>) were the thermodynamic fitting parameters. The reaction entropy was calculated using the relationships ΔG = -RTlnK<sub>a</sub> (R = 1.9872 cal mol<sup>-1</sup> K<sup>-1</sup>, T = 298 K) and ΔG = ΔH - TΔS. The reliability of the obtained fits was evaluated using the reduced χ<sup>2</sup> (χ<sub>v</sub><sup>2</sup>) parameter.

Initial crystallization screens were carried out using the HTXLab facilities at EMBL Grenoble set up using a sitting-drop vapor diffusion 1:1 ratio format with 100 nL protein solution ± 1 molar eq. of Ni(II) at 400 μM dimer. Crystals were optimized manually in the laboratory by the hanging drop method, using 2 μL holo-protein solutions at 400 μM dimer added with an equal volume of precipitant. The best

diffracting crystal of Ni-*RrCooJ*- $\Delta$  was obtained with a protein concentration of 400  $\mu$ M dimer in the presence of one equivalent of Ni(II) per dimer, in 16% (v/v) PEG3350 and 0.2 M calcium chloride. Micro-seeding was carried out, according to the Hampton protocol ("seed-bead" kit), to increase the size of the crystals. A cryo-protection solution was made starting from the mother liquor composition by addition of 25 % glycerol before flash-cooling in liquid nitrogen.

Diffraction data were collected on the ID30B beamline at the European Synchrotron (ESRF, Grenoble France) (31). Fluorescence spectra were recorded to confirm the presence of Se and Ni atoms in the crystals. The best crystal diffracted at 2.04 Å resolution, and a single anomalous diffraction dataset was recorded above the Se edge ( $\lambda = 0.9786\text{\AA}$ ) to benefit from the anomalous contributions of both Se and Ni atoms ( $f'' \sim 3.8e^-$  and  $2e^-$  respectively). The data were integrated in the P21 space group using XDS and scaled/merged using XSCALE (32). The asymmetric unit was estimated to contain 8 *RrCooJ*- $\Delta$  molecules (68 residues per molecule) based on the Matthews coefficient using the CCP4 suite (33). The structure was solved by the SAD method using the SHELX suite (34). The whole heavy atoms substructure, including 16 Se and 4 Ni atoms, was solved using SHELXD. The experimental phases were calculated and further improved by alternating density-modification and auto-tracing using SHELXE. Then, Phenix-Autobuilt (35) was able to automatically build 78% of the model (422 out of 544 residues located, with  $R_{\text{work}} = 28\%$  and  $R_{\text{free}} = 32\%$ ). The model was subsequently rebuilt and corrected manually using COOT (36) and finally refined using Phenix-Refine (37) resulting in a final model with  $R_{\text{work}} = 20.0\%$  and  $R_{\text{free}} = 25.9\%$  (refinement statistics in Table S3).

Apo*RrCooJ* and apo*RrCooJ*- $\Delta$  were filtered extemporaneously before each experiment, using size exclusion chromatography with a Superdex 200 increase column (GE Healthcare) equilibrated in 50 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP. Protein concentration were 33-240  $\mu$ M of dimer. SAXS data were collected at the European Synchrotron Radiation Facility (Grenoble, France) on the BM29 beamline at BioSAXS. Data were processed using standard procedures with the ATSAS v2.8.3 suite of programs (38). The *ab initio* determination of the molecular shape of the proteins was done as previously described, using DAMMIF(39). Radius of gyration ( $R_g$ ) and

forward intensity at zero angle ( $I(0)$ ) were determined with the programs PRIMUS (40) by using the Guinier approximation at low Q value, in a  $Q \cdot R_g$  range  $< 1.5$ :

$$\ln I(Q) = \ln I(0) - \frac{R_g^2 Q^2}{3}$$

Porod volumes and Kratky plot were determined using the Guinier approximation and the PRIMUS programs (40). The radius of gyration and the pairwise distance distribution function  $P(r)$  were calculated by indirect Fourier transform with the program GNOM (41). The maximum dimension ( $D_{\text{max}}$ ) value was adjusted in order that the  $R_g$  value obtained from GNOM agreed with that obtained from Guinier analysis.

In order to build *ab initio* models, several independent DAMMIF (39) models were calculated in slow mode with pseudo chain option and merged using the program DAMAVER (42). The program CRY SOL (43) was used to generate the theoretical scattering curves from the PDB coordinates form of Ni-*RrCooJ*- $\Delta$ .

The program suite EOM 2.0 (Ensemble Optimisation Method)(20) was used to get information on the predicted intrinsically disordered region: an ensemble of 10,000 conformers was generated using genetic algorithms in order to study the flexibility of the C-terminal domain of *RrCooJ*. A starting model, containing the atomic coordinates of Ni-*RrCooJ*- $\Delta$  structure and a poly-asparagine mimicking the C-terminal region, was generated as input to the EOM program. The genetic algorithm compared the averaged theoretical scattering intensity from N independent ensembles of conformations against the scattering data of *RrCooJ*. The ensemble that best fitted the experimental SAXS data was selected with a maximum number of 10 individual models.

In order to study the flexibility of the C-terminal part of the protein in relation with SAXS data, molecular dynamics simulations were run using the molecular dynamics program CHARMM (44). A dimer of *RrCooJ* was modelled with the all-atom force field *all22* for proteins (45)-(46). Initial coordinates were taken from the X-ray structure of the Ni-*RrCooJ*- $\Delta$  dimer solved from E3 to L68 for chain A and P5 to L68 for chain B. All missing atom coordinates were built with CHARMM using internal coordinates and the structure was energy minimized down to a gradient of 0.1 kcal/mol/Å. With 100 residues unresolved in the dimer, the resulting initial model needed a powerful algorithm able to sample the conformational

space of the protein like Replica eXchanging Self-Guided Langevin Dynamics (47). SGLD enhances conformational search efficiency through acceleration of low frequency motions in a molecular system (48). Different types of restraints were applied on the protein to maintain the position of backbone atoms resolved in the X-ray structure and prevent fraying of  $\alpha$ -helices at their N- and C-terminal during high temperature dynamics. Finally, the radius of gyration of the solvated proteins were estimated to 19.8 and 34 Å, respectively for *RrCooJ*- $\Delta$  and *RrCooJ* dimer by SAXS. The corresponding radius of gyration calculated by CHARMM for the  $\Delta$ -Cter unsolvated system being 17 Å, the radius of gyration of the *RrCooJ* unsolvated dimer was estimated to  $34 - 2.8 = 31.2$  Å. In order to reproduce the experimental observations, a radius

of gyration restraint (*RGYR*) of 2 kcal/mol/Å<sup>2</sup>, calculated on  $\alpha$ -carbon atoms, was thus introduced in the simulations. A restraining radius of 32.6 Å was chosen. (This value was shown in test simulations to best approach the 31.2 Å target). The Screened Coulomb Potentials Implicit Solvent Model (SCPISM)(49), was used in this work involving 16 replicas. Additional details are provided in the Supplementary Information.

**Data deposition:** the atomic coordinates for Ni-CooJ- $\Delta$  have been deposited with the protein databank (<http://www.rcsb.org>) with PDB accession code 6HK5.

**Acknowledgments:** This work was supported by “the ITERLIS PhD program, CEA Life sciences” for MA’s PhD funding, the “FUNBIOCO” project (IDEX-UGA, Initiatives de Recherche stratégiques), the “COSYNBIO” project (Projets exploratoires, Cellule énergie-CNRS), and the Department of Pharmacy and Biotechnology of the University of Bologna. This work has been partially supported by Labex ARCANE and CBH-EUR-GS (ANR-17-EURE-0003). The research leading to these results has received funding from the European Community's Seventh Framework Programme H2020 under iNEXT (H2020 Grant # 653706) and networking support from the COST Action FeSBioNet (Contract CA15133). We thank the High Throughput Crystallization Laboratory (HTX Lab) at the EMBL Grenoble for RrCooJ- $\Delta$  crystallization. We thank the staffs from the BM-30A, BM29-BioSAXS and ID-30B beamlines of the European Synchrotron Facility in Grenoble. We thank Dr Luca Signor (IBS, Grenoble) for mass spectrometry experiments in denaturing conditions (Integrated Structural Biology Grenoble platform).

**Conflicts of interest:** The authors declare that they have no conflicts of interest with the contents of this article.

## References

- Schultz, J. E., and Weaver, P. F. (1982) Fermentation and anaerobic respiration by *Rhodospirillum rubrum* and *Rhodopseudomonas capsulata*. *J. Bacteriol.* **149**, 181–190
- Alfano, M., and Cavazza, C. (2018) The biologically mediated water–gas shift reaction: structure, function and biosynthesis of monofunctional [NiFe]-carbon monoxide dehydrogenases. *Sustain. Energy Fuels*. **2**, 1653–1670
- Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E., and Ludden, P. W. (2001) Life on carbon monoxide: X-ray structure of *Rhodospirillum rubrum* Ni-Fe-S carbon monoxide dehydrogenase. *Proc. Natl. Acad. Sci.* **98**, 11973–11978
- Dobbeek, H., Svetlitchnyi, V., Gremer, L., Huber, R., and Meyer, O. (2001) Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster. *Science*. **293**, 1281–1285
- Bonam, D., McKennat, M. C., Stephenst, P. J., and Ludden, P. W. (1988) Nickel-deficient carbon monoxide dehydrogenase from *Rhodospirillum rubrum*: In vivo and in vitro activation by exogenous nickel (hydrogenase/iron-sulfur protein/electron paramagnetic resonance). *Biochemistry*. **85**, 31–35
- Spangler, N. J., Lindahl, P. A., Bandarian, V., and Ludden, P. W. (1996) Spectroelectrochemical characterization of the metal centers in carbon monoxide dehydrogenase (CODH) and nickel-deficient CODH from *Rhodospirillum rubrum*. *J. Biol. Chem.* **271**, 7973–7977
- Jeon, W. B., Singer, S. W., Ludden, P. W., and Rubio, L. M. (2005) New insights into the mechanism of nickel insertion into carbon monoxide dehydrogenase: Analysis of *Rhodospirillum rubrum* carbon monoxide dehydrogenase variants with substituted ligands to the [Fe<sub>3</sub>S<sub>4</sub>] portion of the active-site C-cluster. *J. Biol. Inorg. Chem.* **10**, 903–912
- Kerby, R. L., and Ludden, P. W. (1997) In vivo nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum* : molecular and physiological characterization of cooCTJ . *J. Bacteriol.* **179**, 2259–2266
- Fox, J. D., Yiping, H. E., Shelver, D., Roberts, G. P., and Ludden, P. W. (1996) Characterization of the region encoding the CO-induced hydrogenase of *Rhodospirillum rubrum*. *J. Bacteriol.* **178**, 6200–6208
- He, Y., Shelver, D., Kerby, R. L., and Roberts, G. P. (1996) Characterization of a CO-responsive transcriptional activator from *Rhodospirillum rubrum*. *J. Biol. Chem.* **271**, 120–123
- Jeon, W. B., Cheng, J., and Ludden, P. W. (2001) Purification and Characterization of Membrane-associated CooC Protein and Its Functional Role in the Insertion of Nickel into Carbon Monoxide Dehydrogenase from *Rhodospirillum rubrum*. *J. Biol. Chem.* **276**, 38602–38609
- Sydor, A. M., Lebrette, H., Ariyakumaran, R., Cavazza, C., and Zamble, D. B. (2014) Relationship between Ni(II) and Zn(II) coordination and nucleotide binding by the helicobacter pylori [NiFe]-hydrogenase and urease maturation factor HypB. *J. Biol. Chem.* **289**, 3828–3841
- Zambelli, B., Stola, M., Musiani, F., De Vriendt, K., Samyn, B., Devreese, B., Van Beeumen, J., Turano, P., Dikiy, A., Bryant, D. A., and Ciurli, S. (2005) UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn<sup>2+</sup>. *J. Biol. Chem.* **280**, 4684–4695
- Timm, J., Brochier-Armanet, C., Perard, J., Zambelli, B., Ollagnier-de-Choudens, S., Ciurli, S., and Cavazza, C. (2017) The CO dehydrogenase accessory protein CooT is a novel nickel-binding protein. *Metallomics*. **9**, 575–583
- Watt, R. K., and Ludden, P. W. (1998) The Identification, Purification, and Characterization of CooJ. *J. Biol. Chem.* **273**, 10019–10025
- Watt, R. K., and Ludden, P. W. (1999) Ni<sup>2+</sup> Transport and Accumulation in *Rhodospirillum rubrum*. **181**, 4554–4560
- Sutton, R. B., Fasshauer, D., Jahn, R., and Brunger, A. T. (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature*. **395**, 347
- Sliz, P., Engelmann, R., Hengstenberg, W., and Pai, E. F. (1997) The structure of enzyme II<sub>lactose</sub> from *Lactococcus lactis* reveals a new fold and points to possible interactions of a multicomponent system. *Structure*. **5**, 775–788
- Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
- Tria, G., Mertens, H. D. T., Kachala, M., and Svergun, D. I. (2015) Advanced ensemble modelling

- of flexible macromolecules using X-ray solution scattering. *IUCrJ.* **2**, 207–217
21. Colpas, G. J., and Hausinger, R. P. (2000) In vivo and in vitro kinetics of metal transfer by the *Klebsiella aerogenes* urease nickel metallochaperone, UreE. *J. Biol. Chem.* **275**, 10731–10737
22. Zhang, J. W., Butland, G., Greenblatt, J. F., Emili, A., and Zamble, D. B. (2005) A role for SlyD in the *Escherichia coli* hydrogenase biosynthetic pathway. *J. Biol. Chem.* **280**, 4360–4366
23. Olson, J. W., and Maier, R. J. (2000) Dual roles of *Bradyrhizobium japonicum* nickelin protein in nickel storage and GTP-dependent Ni mobilization. *J. Bacteriol.* **182**, 1702–1705
24. Real-Guerra, R., Staniscuaski, F., Zambelli, B., Musiani, F., Ciurli, S., and Carlini, C. R. (2012) Biochemical and structural studies on native and recombinant Glycine max UreG: a detailed characterization of a plant urease accessory protein. *Plant Mol. Biol.* **78**, 461–475
25. Desguin, B., Soumillion, P., Hols, P., and Hausinger, R. P. (2016) Nickel-pincer cofactor biosynthesis involves LarB-catalyzed pyridinium carboxylation and LarE-dependent sacrificial sulfur insertion. *Proc. Natl. Acad. Sci.* **113**, 5598–5603
26. Vinella, D., Fischer, F., Vorontsov, E., Gallaud, J., Malosse, C., Michel, V., Cavazza, C., Robbe-Saule, M., Richaud, P., Chamot-Rooke, J., Brochier-Armanet, C., and De Reuse, H. (2015) Evolution of *Helicobacter*: Acquisition by Gastric Species of Two Histidine-Rich Proteins Essential for Colonization. *PLoS Pathog.* **11**, e1005312
27. Lacasse, M. J., Douglas, C. D., and Zamble, D. B. (2016) Mechanism of Selective Nickel Transfer from HypB to HypA, *Escherichia coli* [NiFe]-Hydrogenase Accessory Proteins. *Biochemistry.* **55**, 6821–6831
28. Khorasani-Motlagh, M., Lacasse, M. J., and Zamble, D. B. (2017) High-affinity metal binding by the *Escherichia coli* [NiFe]-hydrogenase accessory protein HypB is selectively modulated by SlyD. *Metallomics.* **9**, 482–493
29. Banaszak, K., Martin-Diaconescu, V., Bellucci, M., Zambelli, B., Rypniewski, W., Maroney, M. J., and Ciurli, S. (2012) Crystallographic and X-ray absorption spectroscopic characterization of *Helicobacter pylori* UreE bound to Ni(2)(+) and Zn(2)(+) reveals a role for the disordered C-terminal arm in metal trafficking. *Biochem. J.* **441**, 1017–1026
30. Pérard, J., Covès, J., Castellan, M., Solard, C., Savard, M., Miras, R., Galop, S., Signor, L., Crouzy, S., Michaud-Soret, I., and de Rosny, E. (2016) Quaternary Structure of Fur Proteins, a New Subfamily of Tetrameric Proteins. *Biochemistry.* **55**, 1503–1515
31. Mueller-Dieckmann, C., Bowler, M. W., Carpentier, P., Flot, D., McCarthy, A. A., Nanao, M. H., Nurizzo, D., Pernot, P., Popov, A., Round, A., Royant, A., de Sanctis, D., von Stetten, D., and Leonard, G. A. (2015) The status of the macromolecular crystallography beamlines at the European Synchrotron Radiation Facility. *Eur. Phys. J. Plus.* **130**, 70
32. Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D.* **66**, 125–132
33. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D.* **67**, 235–242
34. Sheldrick, G. M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. Sect. D.* **66**, 479–485
35. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D.* **66**, 213–221
36. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 486–501
37. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D.* **68**, 352–367
38. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., and Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350
39. Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape

- determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346
40. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282
41. Svergun, D. I. (2007) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503
42. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864
43. Svergun, D., Barberato, C., and Koch, M. H. J. (2007) CRY SOL— a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28**, 768–773
44. Brooks, B. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217
45. Mackerell, A. D., Feig, M., and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**, 1400–1415
46. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* **102**, 3586–3616
47. Wu, X., Hodoscek, M., and Brooks, B. R. (2012) Replica exchanging self-guided Langevin dynamics for efficient and accurate conformational sampling. *J. Chem. Phys.* **137**, 44106
48. Jiang, W., Hodoscek, M., and Roux, B. (2009) Computation of Absolute Hydration and Binding Free Energy with Free Energy Perturbation Distributed Replica-Exchange Molecular Dynamics. *J. Chem. Theory Comput.* **5**, 2583–2588
49. Hassan, S. A., Mehler, E. L., Zhang, D., and Weinstein, H. (2003) Molecular dynamics simulations of peptides and proteins with a continuum electrostatic model based on screened Coulomb potentials. *Proteins Struct. Funct. Bioinforma.* **51**, 109–125
50. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190
51. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38

## Legends to figures

### Figure 1. Amino acid sequences and secondary structure prediction of *RrCooJ*.

(A) Amino acid sequences of *RrCooJ* and mutants. Histidines are in red. Cys 109 and 111 are in green. In *RrCooJ*-3HA and *RrCooJ*-Δ3HA mutants, His18, 22 and 26 are mutated in Ala (in cyan). In *RrCooJ*-C2S mutant, Cys 109 and 111 are mutated in Ser (in blue). (B) Secondary structure elements of *RrCooJ*-WT predicted using the PredictProtein server (<http://www.predictprotein.org>).

### Figure 2. SEC-MALLS-RI and circular dichroism (CD) in far UV of apo*RrCooJ* and apo*RrCooJ*-Δ.

(A) SEC-MALLS-RI profile for apo*RrCooJ* (black) and apo*RrCooJ*-Δ (red). Both proteins are in a dimeric conformation in solution with a mass of  $25.7 \pm 0.2\%$  kDa for apo*RrCooJ* and  $15.7 \pm 0.2\%$  kDa for apo*RrCooJ*-Δ. The SEC-MALLS-RI spectra were recorded for a protein concentration of 750 μM dimer (apo*RrCooJ*) and 285 μM dimer (apo*RrCooJ*-Δ) in 50 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP. (B) The CD spectra in the 190-250 nm range shows the secondary structure composition of the protein in solution. The apo*RrCooJ* spectrum (black) was recorded at 12 μM dimer concentration in 8 mM HEPES pH 7.2 10 mM NaCl. The apo*RrCooJ*-Δ spectrum (red) was recorded at 14 μM dimer concentration in 9 mM HEPES pH 7.2 10 mM NaCl.

### Figure 3. Binding properties of Ni(II) to *RrCooJ* and its variants using CD in the near-UV visible region (250-700 nm).

(A) Ni(II) titration of *RrCooJ* at 20 μM dimer, (B) *RrCooJ*-C2S at 20 μM dimer, (C) *RrCooJ*-Δ at 100 μM dimer and (D) *RrCooJ*-Δ3HA at 50 μM dimer. Buffer: 50 mM HEPES pH 7.5, 300 mM NaCl (with the addition of 1 mM TCEP for spectra (A)).

### Figure 4. Binding properties of Ni(II) to *RrCooJ* and its variants determined using isothermal titration calorimetry.

A) Integrated heat data of Ni(II) titration over *RrCooJ* as a function of metal/protein molar ratio. The continuous line represents the best fit obtained using a two sets of sites model. (B) Integrated heat data of Ni(II) titration over *RrCooJ*-Δ as a function of metal/protein molar ratio. The continuous line represents the best fit obtained using a one sets of site model. (C) Integrated heat data of Ni(II) titration over *RrCooJ*-C2S. The continuous line represents the best fit obtained using a two sets of sites model. (D) Integrated heat data of Ni(II) titration over *RrCooJ*-3HA. The continuous line represents the best fit obtained using a one set of sites model. (E) Integrated heat data of Ni(II) titration over *RrCooJ*-Δ3HA. (F) Integrated heat data of Ni(II) titration over *RrCooJ*-ΔE29A. The continuous line represents the best fit obtained using a one set of sites model.

### Figure 5. Ni(II) impact on *RrCooJ* and *RrCooJ*-Δ conformation using SEC-MALLS-RI.

(A) Chromatogram of SEC-MALLS-RI analysis of *RrCooJ* in the presence of Ni(II) (from 0 to 4 molar eq). 20 μL of protein samples were injected at 45 μM of dimer in a Superdex 200 increase equilibrated in 50 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP. Apo *RrCooJ* (blue), in the presence of 22.5 μM of Ni(II) (0.5 eq.) (violet), 45 μM of Ni(II) (1 eq.) (green), 90 μM of Ni(II) (2 eq.) (yellow), 135 μM of Ni(II) (3 eq.) (orange) and 180 μM of Ni(II) (4 eq.) (red). The final protein concentration in the dimeric form was determined by ASTRA 6 program and found equal to 44 μM dimer (100 % of injected protein) for the apo *RrCooJ*, and *RrCooJ* pre-incubated with 0.5 eq Ni(II) or 1 eq Ni(II), 36 μM (82 % of injected protein) for *RrCooJ* pre-incubated with 2 eq Ni(II) and *RrCooJ* pre-incubated with 3 eq Ni(II) and 28 μM (64 % of injected protein) for pre-incubated with 4 eq Ni(II). (B) Chromatogram of SEC-MALLS-RI analysis of apo *RrCooJ* (blue), in the presence of 3 eq. Ni(II) (orange) and 3 eq. of Ni(II) + 5mM EDTA (magenta). Samples were injected at 50 μM dimer. (C) Chromatogram of SEC-MALLS-RI analysis of *RrCooJ*-Δ at 100 μM dimer in the absence (black) or pre-incubated with 5 molar eq. of Ni(II) (red).

### Figure 6. X-ray structure of Ni-*RrCooJ*-Δ homodimer.



(A) The two monomers are in magenta and cyan, Ni(II) is depicted as a green sphere. (B) B-factor diagram of Ni-*RrCooJ*- $\Delta$  dimer by the B-factor putty program in Pymol. The B-factors are illustrated by color, ranging from low (blue, B-factor=20) to high (red, B-factor=60).

**Figure 7. The novel Ni(II) binding site in *RrCooJ*.**

(A) Tetrameric conformation of Ni-*RrCooJ*- $\Delta$ , composed of the two homodimers BC (in magenta and cyan) and GE (in grey and red). (B) Structure of the Ni(II)-binding site. Ni(II) is depicted as a green sphere, and coordinating residues are shown in sticks. (C) In grey the 2Fo-Fc map contoured at  $2.0\sigma$ , in green the difference anomalous map contoured at  $3.0\sigma$ . (D) Logo showing the conservation of the 46 CooJ homologues identified by pBlast. For clarity, non-conserved residues forming gaps in ClustalW alignment were removed. The logo of CooJ homologues was generated using the WEBLOGO tool (50).

**Figure 8. Study of apo*RrCooJ* and apo*RrCooJ*- $\Delta$  by SAXS**

(A) Scattering curves of experimental data of apo*RrCooJ* in cyan and apo*RrCooJ*- $\Delta$  in black in solution. (B) Linear dependence of  $\ln[I(Q)]$  vs  $Q^2$  determined by Guinier plot at several concentrations: O=56  $\mu\text{M}$  -  $\Delta$ = 112  $\mu\text{M}$  -  $\square$ = 240  $\mu\text{M}$  dimer for apo*RrCooJ* (cyan) and O= 34  $\mu\text{M}$  -  $\Delta$ = 190  $\mu\text{M}$  -  $\square$ = 340  $\mu\text{M}$  dimer for apo*RrCooJ*- $\Delta$  (black), with standard  $R_g$  limits (0.3 to 1.0), indicating the absence of aggregation for both proteins. (C) Pair distribution function  $p(R)$  in arbitrary units (arb.u) vs.  $r$  (nm) determined by GNOM of apo*RrCooJ* (cyan) with a  $D_{\text{max}}$  of 12nm and apo*RrCooJ*- $\Delta$  (black) with  $D_{\text{max}}$  of 6.5nm. (D) Kratky plot of apo*RrCooJ* (cyan) and apo*RrCooJ*- $\Delta$  (black). (E) Final *ab initio* model of apo*RrCooJ*- $\Delta$  generated with Dammif and merged with Damaver (NSD < 0.7). (F) Final *ab initio* model of apo*RrCooJ* generated with Dammif and merged with Damaver (NSD = 0.8-1.5). (G) Superimposition of the X-ray structure of Ni-*RrCooJ*- $\Delta$  on *ab initio* model of apo*RrCooJ*- $\Delta$  (done with the SUPCOMB software). Fit of the SAXS experimental scattering curve of apo*RrCooJ*- $\Delta$  (black) with the theoretical scattering curve from the crystal structure generated by CRY SOL (red).

**Figure 9. *in silico* modeling of apo*RrCooJ* using EOM2.0 and CHARMM**

(A) Distribution ( $D_{\text{max}}$ ) of end-to-end distances computed from pools containing 10,000 *RrCooJ* structural models generated by EOM2.0 in grey. The five selected models are depicted in red. (B) Fit of experimental data in black compared to final combination model of *RrCooJ* generated by EOM with a  $\chi^2$  of 0.8. (C) Superposition of the five final models of *RrCooJ*: coiled coil domains are in red. For the flexible regions, each color corresponds to an individual model. (D) Modeling of apo*RrCooJ*. Superposition of the minimum energy models of apo*RrCooJ* in each of the 9 clusters obtained after RXSGLD simulation with CHARMM. Each model is represented in a different color. The figure was drawn with VMD (51).

A

*RrCooJ* (12,507 Da)

10 20 30 40 50 60 70 80 90 100 110  
GTESPERGRK RLGIYLAHFL DHVEGHMGEI GVQRDALAED ARLGALIDRA LADMAVARAS LNAVLRDLG EAPAPASPEA VHSFFHSsHaH sHDHdHaHGH sHDHaHdHCH CHdHP

*RrCooJ*-3HA (12,309 Da)

10 20 30 40 50 60 70 80 90 100 110  
GTESPERGRK RLGIYLAHFL DAVEGAMGEI GVQRDALAED ARLGALIDRA LADMAVARAS LNAVLRDLG EAPAPASPEA VHSFFHSsHaH sHDHdHaHGH sHDHaHdHCH CHdHP

*RrCooJ*-C2S (12,607 Da)

10 20 30 40 50 60 70 80 90 100 110  
GTESPERGRK RLGIYLAHFL DHVEGHMGEI GVQRDALAED ARLGALIDRA LADMAVARAS LNAVLRDLG EAPAPASPEA VHSFFHSsHaH sHDHdHaHGH sHDHaHdHSH SHdHP

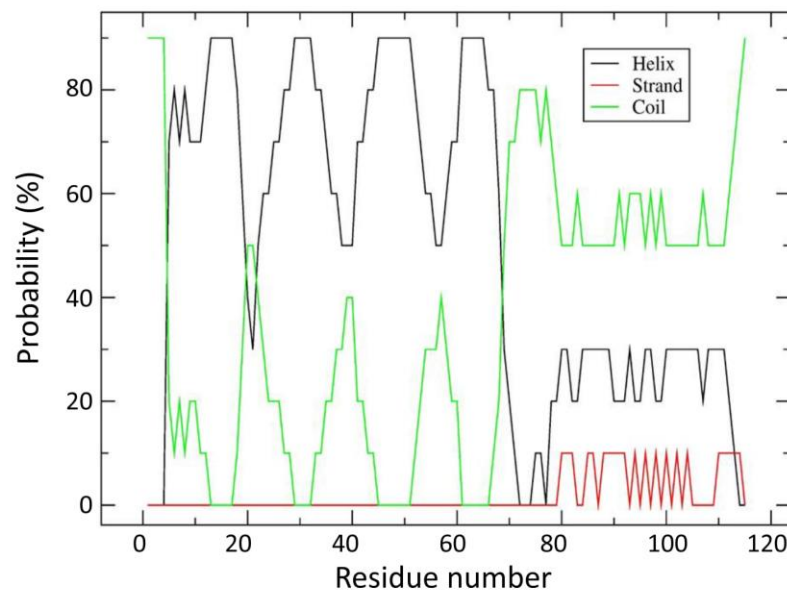
*RrCooJ*-Δ (7,378 Da)

10 20 30 40 50 60  
GTESPERGRK RLGIYLAHFL DHVEGHMGEI GVQRDALAED ARLGALIDRA LADMAVARAS LNAVLRDL

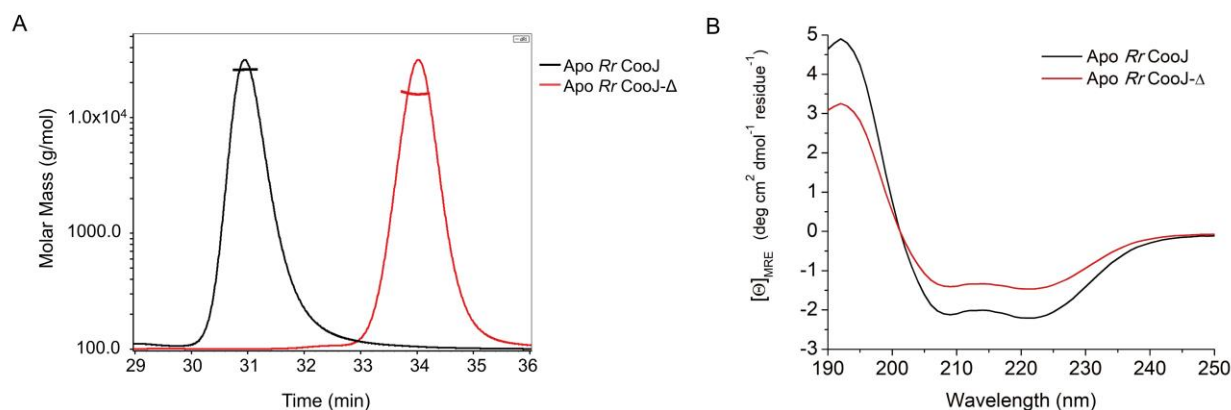
*RrCooJ*-Δ3HA (7,181 Da)

10 20 30 40 50 60  
GTESPERGRK RLGIYLAHFL DAVEGAMGEI GVQRDALAED ARLGALIDRA LADMAVARAS LNAVLRDL

B

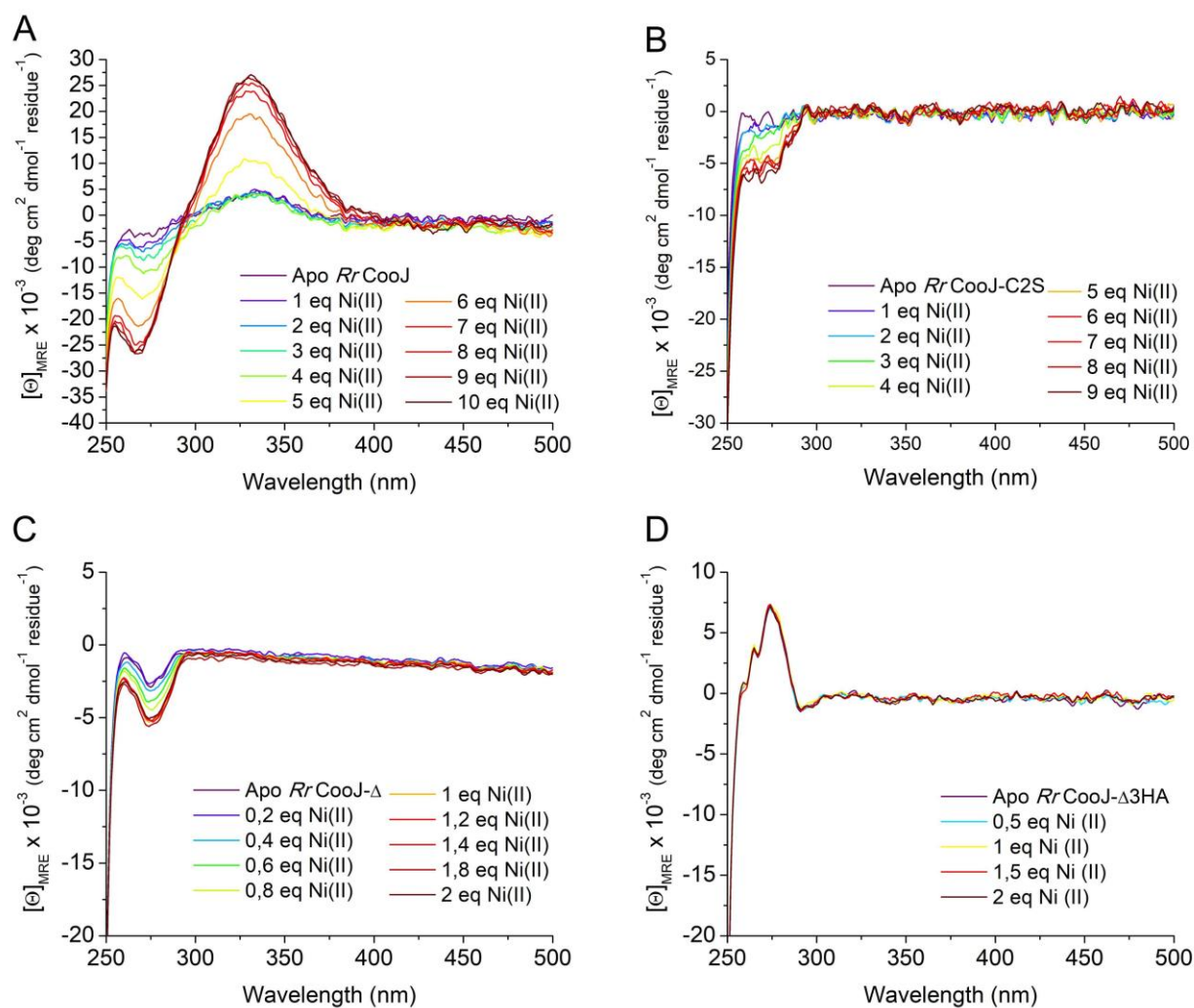


**Figure 1.** Amino acid sequences and secondary structure prediction of *RrCooJ*-WT. (A) Amino acid sequences of *RrCooJ* and mutants. Histidines are in red. Cys 109 and 111 are in green. In *RrCooJ*-3HA and *RrCooJ*-Δ3HA mutants, His18, 22 and 26 are mutated in Ala (in cyan). In *RrCooJ*-C2S mutant, Cys 109 and 111 are mutated in Ser (in blue). (B) Secondary structure elements of *RrCooJ*-WT predicted using the PredictProtein server (<http://www.predictprotein.org>).

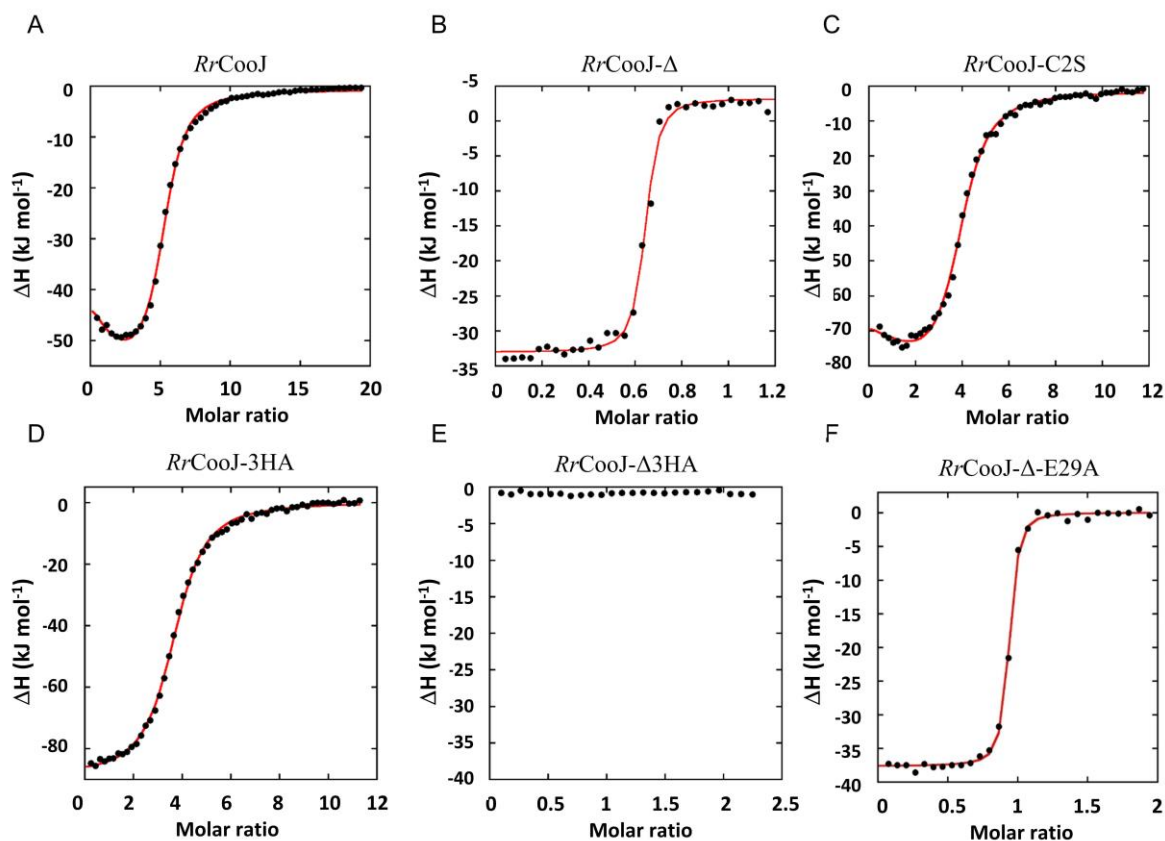


**Figure 2. SEC-MALLS-RI and circular dichroism (CD) in far UV of apoRrCooJ and apoRrCooJ-Δ.**

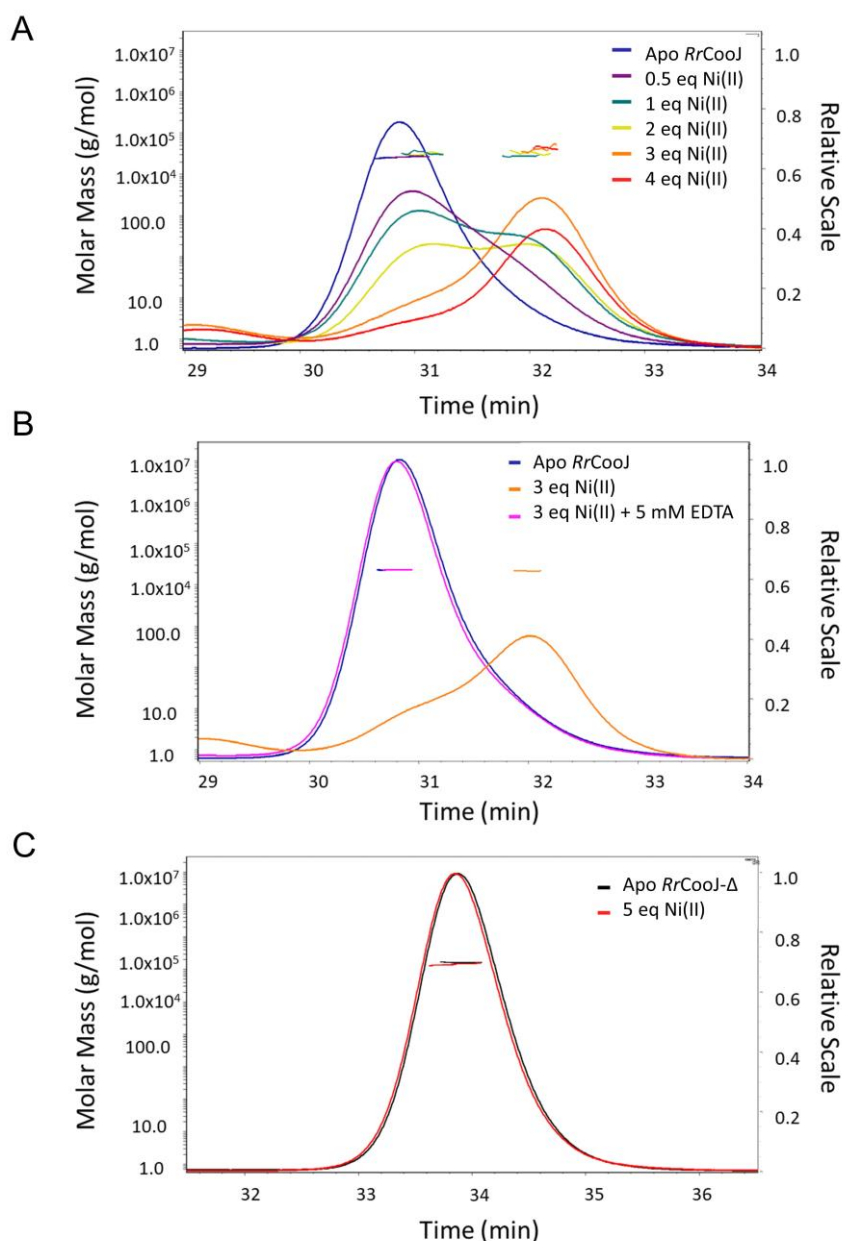
(A) SEC-MALLS-RI profile for apoRrCooJ (black) and apoRrCooJ-Δ (red). Both proteins are in a dimeric conformation in solution with a mass of  $25.7 \pm 0.2\%$  kDa for apoRrCooJ and  $15.7 \pm 0.2\%$  kDa for apoRrCooJ-Δ. The SEC-MALLS-RI spectra were recorded for a protein concentration of 750 μM dimer (apoRrCooJ) and 285 μM dimer (apoRrCooJ-Δ) in 50 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP. (B) The CD spectra in the 190-250 nm range shows the secondary structure composition of the protein in solution. The apoRrCooJ spectrum (black) was recorded at 12 μM dimer concentration in 8 mM HEPES pH 7.2, 10 mM NaCl. The apoRrCooJ-Δ spectrum (red) was recorded at 14 μM dimer concentration in 9 mM HEPES pH 7.2, 10 mM NaCl.



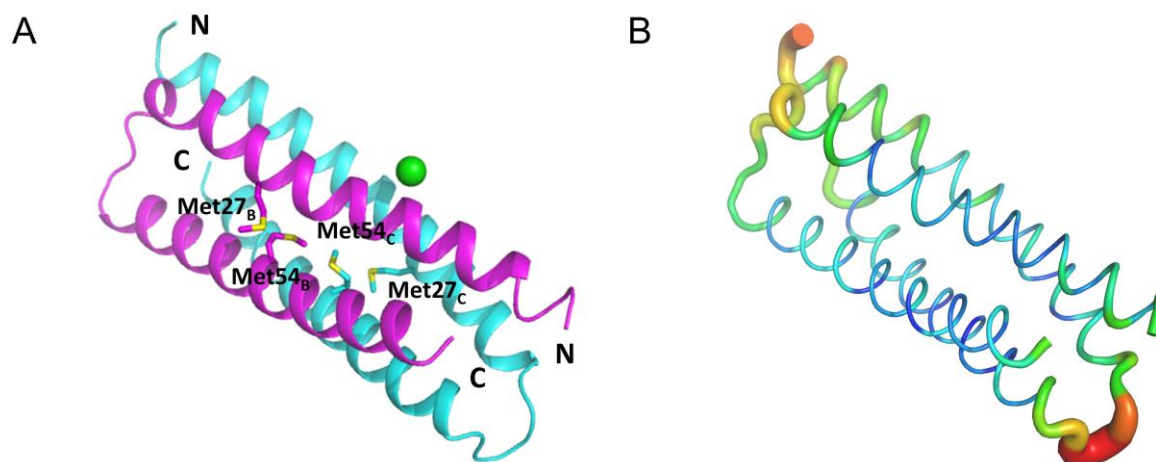
**Figure 3.** Binding properties of Ni(II) to *RrCooJ* and its variants using CD in the near-UV visible region (250-700 nm). (A) Ni(II) titration of *RrCooJ* at 20 μM dimer, (B) *RrCooJ*-C2S at 20 μM dimer, (C) *RrCooJ*-Δ at 100 μM dimer and (D) *RrCooJ*-Δ3HA at 50 μM dimer. Buffer: 50 mM HEPES pH 7.5, 300 mM NaCl (with the addition of 1 mM TCEP for spectra (A)).



**Figure 4.** Binding properties of Ni(II) to *RrCooJ* and its variants determined using isothermal titration calorimetry. (A) Integrated heat data of Ni(II) titration over *RrCooJ* as a function of metal/protein molar ratio. The continuous line represents the best fit obtained using a two sets of sites model. (B) Integrated heat data of Ni(II) titration over *RrCooJ-Δ* as a function of metal/protein molar ratio. The continuous line represents the best fit obtained using a one sets of site model. (C) Integrated heat data of Ni(II) titration over *RrCooJ-C2S*. The continuous line represents the best fit obtained using a two sets of sites model. (D) Integrated heat data of Ni(II) titration over *RrCooJ-3HA*. The continuous line represents the best fit obtained using a one set of sites model. (E) Integrated heat data of Ni(II) titration over *RrCooJ-Δ3HA*. (F) Integrated heat data of Ni(II) titration over *RrCooJ-ΔE29A*. The continuous line represents the best fit obtained using a one set of sites model.

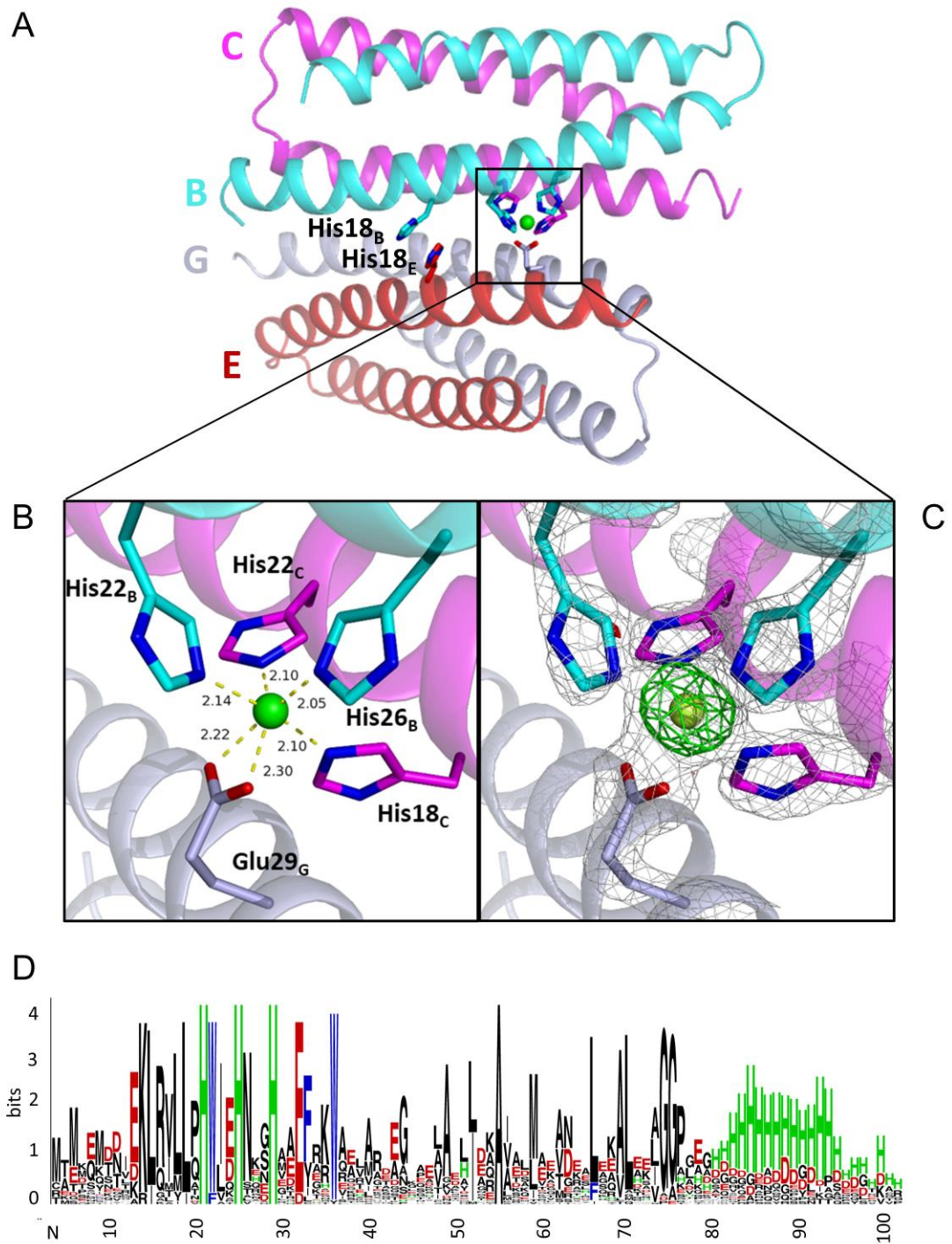


**Figure 5.** Ni(II) impact on *RrCooJ* and *RrCooJ*-Δ conformation using SEC-MALLS-RI. (A) Chromatogram of SEC-MALLS-RI analysis of *RrCooJ* in the presence of Ni(II) (from 0 to 4 molar eq). 20 μL of protein samples were injected at 45 μM of dimer in a Superdex 200 increase equilibrated in 50 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP. Apo*RrCooJ* (blue), in the presence of 22.5 μM of Ni(II) (0.5 eq.) (violet), 45 μM of Ni(II) (1 eq.) (green), 90 μM of Ni(II) (2 eq.) (yellow), 135 μM of Ni(II) (3 eq.) (orange) and 180 μM of Ni(II) (4 eq.) (red). The final protein concentration in the dimeric form was determined by ASTRA 6 program and found equal to 44 μM dimer (100 % of injected protein) for the apo*RrCooJ*, and *RrCooJ* pre-incubated with 0.5 eq Ni(II) or 1 eq Ni(II), 36 μM (82 % of injected protein) for *RrCooJ* pre-incubated with 2 eq Ni(II) and *RrCooJ* pre-incubated with 3 eq Ni(II) and 28 μM (64 % of injected protein) for pre-incubated with 4 eq Ni(II). (B) Chromatogram of SEC-MALLS-RI analysis of apo*RrCooJ* (blue), in the presence of 3 eq. Ni(II) (orange) and 3 eq. of Ni(II) + 5 mM EDTA (magenta). Samples were injected at 50 μM dimer. (C) Chromatogram of SEC-MALLS-RI analysis of *RrCooJ*-Δ at 100 μM dimer in the absence (black) or pre-incubated with 5 molar eq. of Ni(II) (red).



**Figure 6.** X-ray structure of Ni-*RrCooJ*-Δ homodimer. (A) The two monomers are in magenta and cyan, Ni(II) is depicted as a green sphere. (B) B-factor diagram of Ni-*RrCooJ*-Δ dimer by the B-factor putty program in Pymol. The B-factors are illustrated by color, ranging from low (blue, B-factor=20) to high (red, B-factor=60).

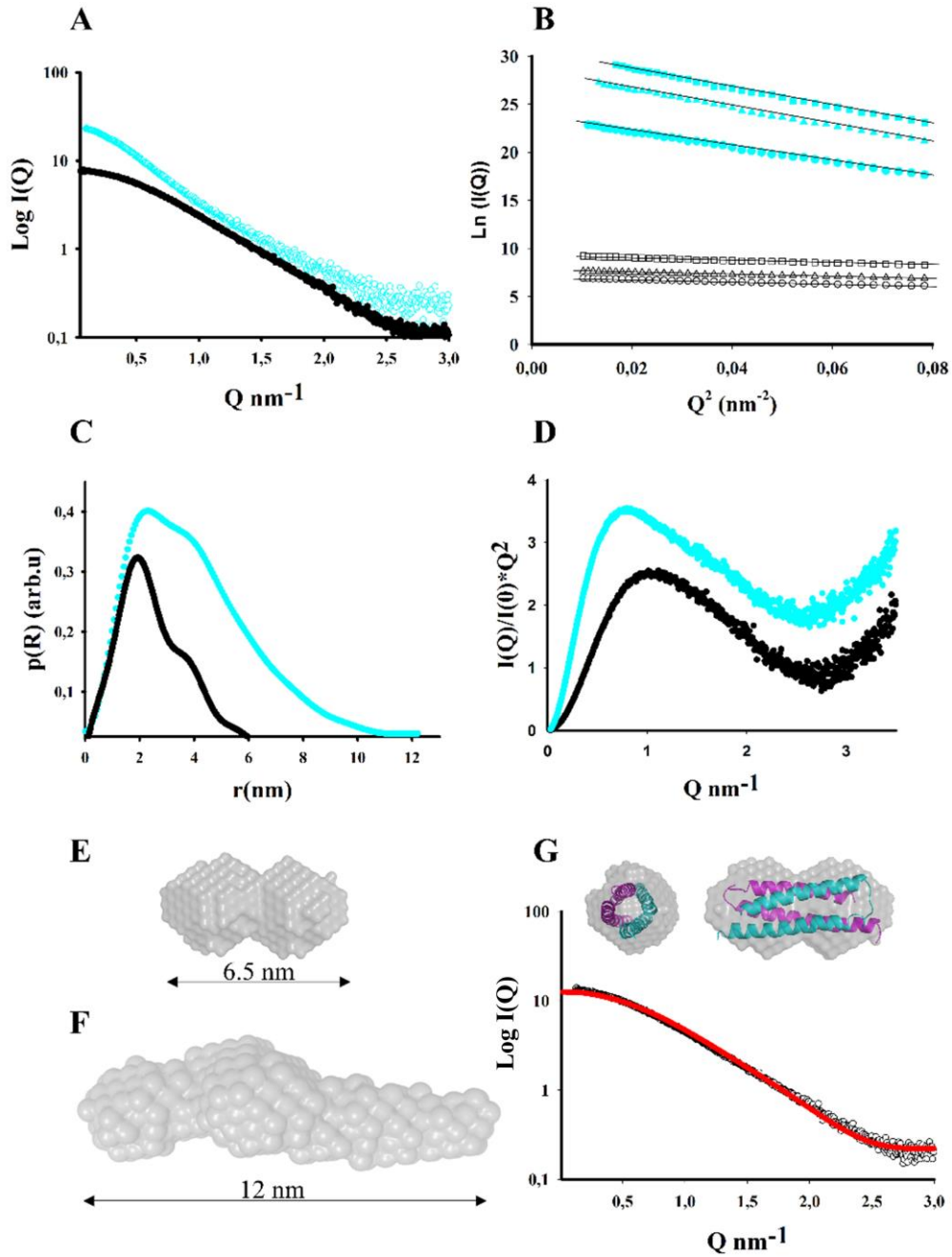




**Figure 7. The novel Ni(II) binding site in Rr-CooJ.**

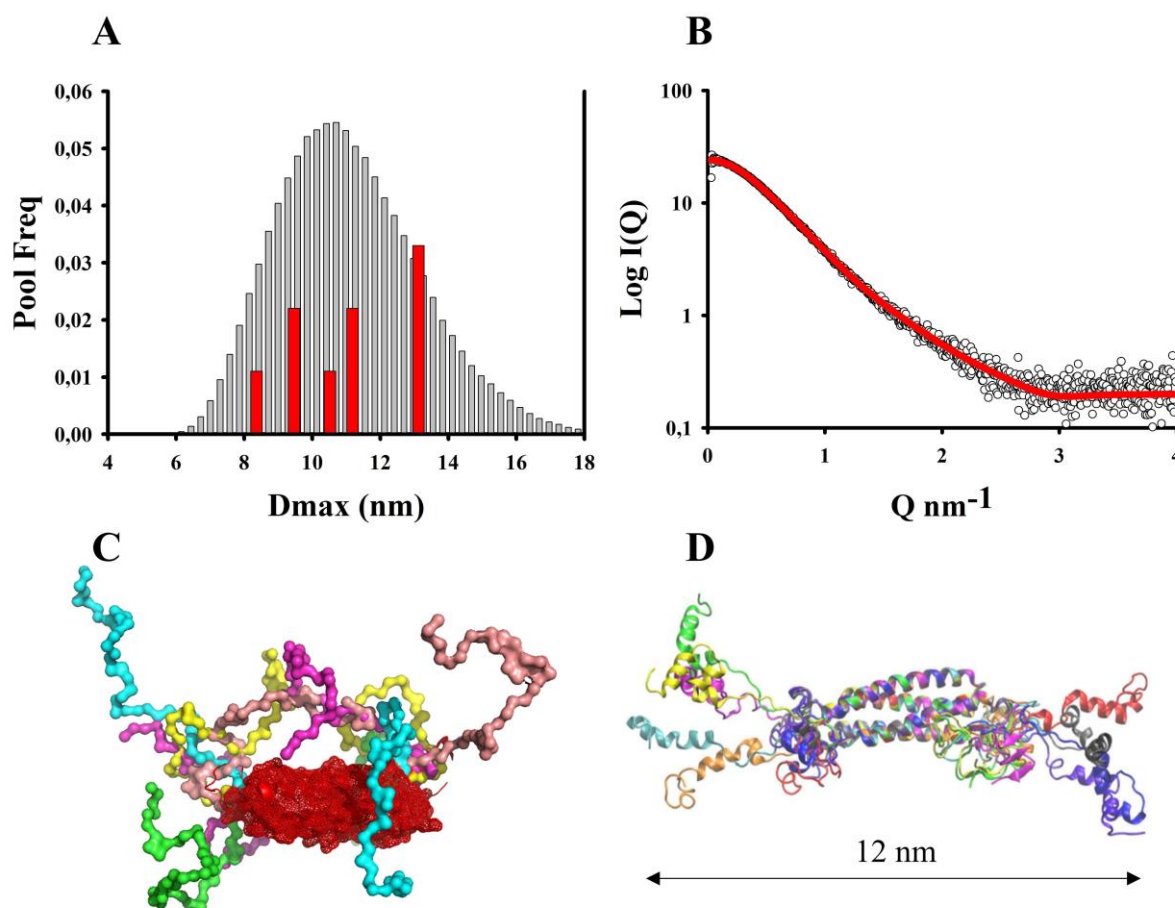
(A) Tetrameric conformation of Ni-RrCooJ-Δ, composed of the two homodimers BC (in magenta and cyan) and GE (in grey and red). (B) Structure of the Ni(II)-binding site. Ni(II) is depicted as a green sphere, and coordinating residues are shown in sticks. (C) In grey the 2Fo-Fc map contoured at 2.0σ, in green the difference anomalous map contoured at 3.0σ. (D) Logo showing the conservation of the 46 CooJ homologues identified by pBlast. For clarity, non-conserved residues forming gaps in ClustalW alignment were removed. The logo of CooJ homologues was generated using the WEBLOGO tool (50).





**Figure 8. Study of apoRrCooJ and apoRrCooJ-Δ by SAXS**

(A) Scattering curves of experimental data of apoRrCooJ in cyan and apoRrCooJ-Δ in black in solution. (B) Linear dependence of  $\ln[I(Q)]$  vs  $Q^2$  determined by Guinier plot at several concentrations: O=56  $\mu$ M -  $\Delta$ = 112  $\mu$ M -  $\square$ = 240  $\mu$ M dimer for apoRrCooJ (cyan) and O= 34  $\mu$ M -  $\Delta$ = 190  $\mu$ M -  $\square$ = 340  $\mu$ M dimer for apoRrCooJ-Δ (black), with standard  $R_g$  limits (0.3 to 1.0), indicating the absence of aggregation for both proteins. (C) Pair distribution function  $p(R)$  in arbitrary units (arb.u) vs.  $r$  (nm) determined by GNOM of apoRrCooJ (cyan) with a  $D_{max}$  of 12nm and apoRrCooJ-Δ (black) with  $D_{max}$  of 6.5nm. (D) Kratky plot of apoRrCooJ (cyan) and apoRrCooJ-Δ (black). (E) Final *ab initio* model of apoRrCooJ-Δ generated with Dammif and merged with Damaver (NSD < 0.7). (F) Final *ab initio* model of apoRrCooJ generated with Dammif and merged with Damaver (NSD = 0.8-1.5). (G) Superimposition of the X-ray structure of Ni-RrCooJ-Δ on *ab initio* model of apoRrCooJ-Δ (done with the SUPCOMB software). Fit of the SAXS experimental scattering curve of apoRrCooJ-Δ (black) with the theoretical scattering curve from the crystal structure generated by CRY SOL (red).



**Figure 9. *in silico* modeling of apoRrCooJ using EOM2.0 and CHARMM**

(A) Distribution (Dmax) of end-to-end distances computed from pools containing 10,000 *RrCooJ* structural models generated by EOM2.0 in grey. The five selected models are depicted in red. (B) Fit of experimental data in black compared to final combination model of *RrCooJ* generated by EOM with a  $\chi^2$  of 0.8. (C) Superposition of the five final models of *RrCooJ*: coiled coil domains are in red. For the flexible regions, each color corresponds to an individual model. (D) Modeling of apoRrCooJ. Superposition of the minimum energy models of apoRrCooJ in each of the 9 clusters obtained after RXSGLD simulation with CHARMM. Each model is represented in a different color. The figure was drawn with VMD (51).