PRE-driven Protein NMR Structures: an Alternative Approach in Highly Paramagnetic Systems
Inês B. Trindade[a], Michele Invernici[b], Francesca Cantini[b], Ricardo O. Louro[a]*, Mario Piccioli[b]*

[a] I.B. Trindade and Dr. R.O. Louro, Instituto de Tecnologia Química e Biológica António Xavier (ITQB-NOVA), Universidade Nova de Lisboa, Av. da República (EAN), 2780-157 Oeiras, Portugal.
E-mail: louro@itqb.unl.pt; phone: +351214469332

[b] Dr. M. Invernici, Dr. F. Cantini, Prof. M. Piccioli, Magnetic Resonance Center and Department of Chemistry, University of Florence, Via L. Sacconi 6 50019 Sesto Fiorentino, Italy.
E-mail: piccioli@cerm.unifi.it; phone: +39 055 4574188

Keywords: HIPIP, iron-sulfur proteins, paramagnetic NMR, NMR solution structure, Paramagnetic Relaxation Enhancement, Structural Biology, metalloproteins  Running Title: PRE-driven NMR Structure of HIPIP PioC

Abreviations
HIPIP- High Potential Iron-sulfur Protein
HSQC- Heteronuclear Single Quantum Coherence
NOE- Nuclear Overhauser Enhancement
NOESY- Nuclear Overhauser Enhancement SpectroscopY
PRE- Paramagnetic Relaxation Enhancement
RMSD- Root Mean Square Deviation

Databases
The 20 conformers with the lowest target function constituting the final family obtained using the full set of NMR restraints were deposited to the Protein Data Bank (PDB ID 6XYV). The 20 conformers with the lowest target function obtained using NOEs only (PDB ID: 7A58) and PREs only (PDB ID: 7A4L) were also deposited to the Protein Data Bank. The chemical shift assignments were deposited to the BMRB (code 34487).

Abstract

Metalloproteins play key roles across biology, and knowledge of their structure is essential to understand their physiological role. For those metalloproteins containing paramagnetic states, the enhanced relaxation caused by the unpaired electrons often makes signal detection unfeasible near the metal center, precluding adequate structural characterization right where it is more biochemically relevant. Here, we report a protein structure determination by NMR where two different sets of restraints, one containing Nuclear Overhauser Enhancements (NOEs) and another containing Paramagnetic Relaxation Enhancements (PREs), are used separately and eventually together. The protein PioC from Rhodopseudomonas palustris TIE-1 is a High Potential Iron-Sulfur Protein (HiPIP) where the [4Fe-4S] cluster is paramagnetic in both oxidation states at room temperature providing the source of (PREs) used as alternative distance restraints. Comparison of the family of structures obtained using NOEs only, PREs only and the combination of both reveals that the pairwise root mean square deviation (RMSD) between them are similar and comparable with the precision within each family. This demonstrates that, under favorable conditions in terms of protein size and paramagnetic effects, PREs can efficiently complement and eventually replace NOEs for the structural characterization of small paramagnetic metalloproteins and denovo designed metalloproteins by NMR.
Introduction

Metalloproteins represent 40 to 47% of all known enzymes [1, 2] and, for all of them, the metal center(s) are essential for catalysis, electron transfer, metal storage/transport, or they play a crucial role in determining stability and structural properties [3-12]. Structural biologists are mainly interested in obtaining detailed information in the proximity of the metal center(s), where the biochemically relevant events occur. NMR is a privileged approach for characterizing metalloproteins since it can provide the structure in solution at atomic resolution, information about the amplitudes and time-scales of internal dynamics [13-15], as well as hints on the electronic structure and oxidation states of the metal center [16] in conditions that mimic the physiological context. A significant part of the metallo-proteome contains paramagnetic metal ions. Their presence perturbs the chemical shift and relaxation of NMR signals in ways that can be converted into structural and dynamic information [17, 18].

Early attempts to convert paramagnetic nuclear relaxation rates in NMR restraints and include them into solution structure protocols achieved the first solution structures of paramagnetic proteins [19]. The introduction of residual dipolar couplings (RDCs) arising from self-oriented paramagnetic proteins [20] combined with pseudocontact shifts (PCS) and cross correlation rates (CCR) succeeded, in the case of cytochrome c’, to obtain the first backbone structure of a protein in the absence of NOE measurements [21] and, at the same time, opened the possibility to use orientational restraints also in diamagnetic proteins. Nitroxide radicals used as site-directed spin labels allowed the conversion of Paramagnetic Broadening Effects into distance restraints for large molecular weight proteins [22]. This work led to the introduction of the “concept” of Paramagnetic Relaxation as long-range restraints for non-native metalloproteins and the acronym PRE (Paramagnetic Relaxation Enhancements) was coined [23]. Since they are long-range constraints, PREs are a good alternative/complement to RDCs to obtain solution structures when NOES and scalar couplings are unable/insufficient to define the structure. Extrinsic paramagnetic centers can be attached via conjugation to specific, solvent exposed sites [24] and, indeed, the use of PRE soon flourished. Macromolecular structures using PRE data have been characterized not only for soluble proteins [23, 25], but also for protein-protein [26-28] and protein-nucleic acid complexes [29-31], membrane proteins [32], unfolded or partially unfolded states [33-36], proteins in
living cells [37, 38]. Also in solid state protein NMR spectroscopy, PREs coupled with PCS [39, 40] provided accurate solid state structures in the absence of conventional distance or dihedral angle restraints [41, 42] as well as information on the oligomerization interface of large membrane proteins [43]. Actually, applications of PREs go beyond their use as structural constraints for obtaining “static” NMR structures: they may unravel structural information on transient, invisible, intermediates [44], provide information on encounter complexes [45-47], interdomain motions [48], transient protein associations [49-51], nonspecific protein-DNA interactions [52] and also in drug discovery [53-56].

However, several aspects, such as the accuracy of PRE data [57] and the weak correlation observed between distance derived PREs and those found in crystallographic structures [58, 59] deserve further investigations. For example, it has been shown for lanthanide ions that magnetic anisotropy provides a substantial angular dependence of nuclear relaxation rates [60, 61]. Non-specific intermolecular PREs and cross correlations between Curie Spin Relaxation and dipolar spin-spin couplings [62, 63] might be responsible for the deviation of paramagnetic relaxation rates from the \( r^6 \) dependency (\( r \) being the metal-to-nucleus distance) [58]. These factors limit the accuracy of PREs especially at longer metal to proton distances; on the other hand, at shorter metal to proton distances, the paramagnetism-induced line broadening prevents the detection of NMR signals limiting the information available at the close proximity of the paramagnetic center.

For the above reasons, small sized metalloproteins are interesting cases to study the behavior of paramagnetic relaxation when the paramagnetic center is not affected by local mobility and represents the crucial part of the protein. Moreover, tailored experiments increased the availability of PRE values also at shorter metal to nucleus distances [64]. In these conditions, PREs are usable not only as long-range restraints but also as medium and short-range restraints. A \([Fe_4S_4]^{2+}\) cluster possesses a negligible magnetic anisotropy and it is buried within the protein. This rules out the many effects that give rise to relaxation anisotropy. Therefore, PioC, a small (54 aa residues) HiPIP (High Potential Iron Protein) from \( R.\) palustris TIE-1 represents a peculiar and interesting system to assess nuclear relaxation properties, analyze the relative contribution of NOE and PREs and discuss the use of PREs as alternative to NOE for the portion of the protein where the presence of paramagnetic metal ions poses a challenge for the detection of classical structural restraints and even for signal detection [21].
The interplay between NOEs and paramagnetism-based restraints has been addressed by many groups, and there are many evidences of the fact that the replacement of NOEs with other restraints is feasible, when the alternative restraints considered, arise from different complementary sources [89-92]. Sparse NOE and chemical shifts can be used together with sophisticated modeling methods to obtain well-defined solution structures [93, 94]. Protein structures have also been determined without NOEs using orientational restraints from at least two full sets of RDC [21, 95-97]. Tagging a protein with lanthanide ions at four different sites succeeded to obtain a backbone structure without NOEs [98], whereas the use of different lanthanide ions on a single metal center required a minimal number of NOEs to obtain a structure [99]. A backbone structure was obtained using a combination of PRE, RDC, PCS, CCR and backbone NOEs, which were used to properly define $\alpha$ helices [100]. Paramagnetic Relaxation Enhancements (PRE) are dipole-dipole restraints, like NOEs. Therefore, if a sufficient number of PRE restraints are available throughout the entire protein, they should restrain the conformational space with efficiency comparable to NOEs, even when, like in a metalloprotein, the distance restraints all involve a single point, i.e. the native paramagnetic center of the system. To explore these issues, the NMR solution structure of the small iron-sulfur protein PioC from *R. palustris* TIE-1 [101] was pursued as a paradigmatic challenge. PioC mediates the electron transfer between the reaction center and the iron-oxidase in the photoferrotrophic metabolism of *R. palustris* TIE-1. It contains a $[4\text{Fe}-4\text{S}]^{2+}$ cluster with a very high reduction potential ($E^0 = +450$ mV vs SHE), being stable in the $[\text{Fe}_4\text{S}_4]^{2+}$ oxidation state. With 54 amino acids it is the smallest HiPIP ever isolated. Its 3D structure is unknown, but homology modeling with other HiPIPs suggests that the protein has a compact globular structure, characterized by the absence of topologically relevant secondary structure elements; it is instead predicted to be composed essentially by a series of loops and turns wrapped around the $[4\text{Fe}-4\text{S}]$ cluster [102, 103]. The electronic properties of $[\text{Fe}_x\text{S}_y]^{3+}$ in HiPIPs have been studied in detail over the past 40 years [3, 104-112]. The magnetic coupling within the $[4\text{Fe}-4\text{S}]$ cluster makes the electronic correlation times of the individual iron ions much shorter than isolated high spin Fe$^{3+}$ or Fe$^{2+}$ ion; nevertheless, paramagnetic contributions to nuclear relaxation are significant for nuclei within a 10 Å sphere from the cluster [113, 114]. Therefore, PioC is a suitable system to address an important issue for inorganic biochemistry: is it possible to improve methods for measuring relaxation rates in paramagnetic systems to
the point that PREs can be used as the sole source of restraints to define the structure of a metalloprotein?

**Results and discussion**

NMR experiments currently available for measuring $^1$H $R_1$ and $R_2$ rates [23, 57] fail to provide PRE data in the proximity of the paramagnetic center, where many signals are broadened beyond detection. This limits the use of PREs in metalloproteins, in which the paramagnetic center is also the core of the protein. Recently, we developed experiments [64, 115] that provided accurate $^1$H $R_1$ and $R_2$ values in the range 50-500 s$^{-1}$ and substantially improved the amount of PRE restraints in the close proximity of the paramagnetic center. Accurate measurements of both $R_1$ and $R_2$ rates are important to obtain reliable information on the metal-to-proton distances and to use PREs also as short-range restraints. A standard $^{15}$N HSQC experiment on a PioC sample shows only 39 amide resonances out of expected 49 non-proline residues. However, a $^{15}$N IR-HSQC-AP experiment, specifically designed to observe fast relaxing resonances, shows additional 10 resonances, demonstrating that all HN signals of PioC can be detected (**Figure 1A**).

The complete resonance assignment of the protein was obtained combining the conventional approach based on triple resonance experiments (**Table S1**) with a non-systematic procedure using a combination of 1D NOEs, $^{13}$C direct detection, double and triple resonance experiments recorded with parameters optimized à-la-carte (**Table S2**) [118]. These experiments provided the complete NMR assignment of PioC (BRMB entry 34487) [119]. We assigned (excluding the N-ter Val 1) 100% of backbone $^1$H, $^{13}$C, and $^{15}$N resonances, 98% of Hα, 86% and 91% of $^1$H and $^{13}$C side chains atoms, respectively. However, even though the $^1$H resonance were almost completely assigned, $^{15}$N and $^{13}$C HSQC-NOESY experiments at high magnetic field gave only 344 meaningful NOEs, that without any additional information on the [4Fe-4S] cluster binding mode, were insufficient to obtain a converged structure. Three factors quench these NOE intensities: i) the small rotational correlation time (3.4 x 10$^{-9}$ s, from $^{15}$N relaxation); ii) paramagnetic relaxation affecting at least 50% of the protein; iii) the absence of secondary structure elements, typical of HiPIPs. We introduced the cluster into structure calculations: bond distances and angles defining the geometry of the cubane cluster, that are not accessible via NMR, are
given by introducing a special residue into the CYANA library [73], as described in supplementary information. The introduction of the cluster gave structures with backbone and heavy atoms RMSD of 1.27±0.20 Å and 1.95±0.20 Å. Indeed, the cluster is the essential structural element to drive the fold of the polypeptide chain.

Paramagnetic NMR experiments provided structural constraints for cluster binding residues (through coordinative or hydrogen bonds) crucial to define their orientation: Cys βCH₂ hyperfine shifts were converted into four γ₂ dihedral angle constraints defining the cluster binding topology according to a procedure already described [120], seven crucial 1D NOEs provided distances between Cys βCH₂ and neighboring residues (Figure S1), large ¹⁵N contact shifts, observed for Gin27, Val37 and Leu49 [119] were taken as an evidence of three hydrogen-bonds, respectively linking Hα to the Sγ atoms of the preceding (i-2 or i-3) cluster-bound cysteine residues [121]. Only fourteen constraints of this type are available but they are extremely important to frame the cluster within the protein and to provide restraints where other experimental approaches fail to provide information. These cluster-derived restraints are shown in Figure 1B. When also these constraints were included into structure calculation, a well-converged structure was obtained (Figure 2A), with backbone and heavy atoms RMSD of 1.04 ±0.30 Å and 1.81±0.30 Å (residues 5-50).

We next considered the impact of PREs: to this end, we collected R₁ and R₂ values of all amide protons using the ¹⁵N-IR-HSQC AP (48 ¹H R₁ values) and the R₂-weighted ¹⁵N-HSQC-AP experiments (50 ¹H R₂ values) [64]. For non-exchangeable protons, a ¹³C-IR-HSQC AP provided R₁ values of 200 ¹H protons of backbone ¹Hα and side chains. Finally, ¹H and ¹³C resonances of cluster-bound Cys residues, identified and assigned using rapid recycling experiments, provided thirteen R₁ and R₂ values from inversion recovery and linewidth analysis of one dimensional ¹H and ¹³C experiments. Overall, 306 ¹H (amide protons plus backbone ¹Hα and aliphatic side chains) and 5 ¹³C relaxation rates, amounting to ca. six rates per residue, were measured (Table S3). The behavior of Hα, Hα and Hβ rates is shown in Figure 3 and points out that about 60 % of signals are affected by the paramagnetic center.

The R₁,2para contributions are calculated according to:

\[
R_{1,2para} = R_{1,2obs} + R_{1,2para}
\]  

(1)
For backbone atoms, $R_{1,2}$ (red lines in Figure 3) were estimated from averaged values of residues 4-7, that do not belong to the flexible N-term loop (Figure S2) and are not affected by paramagnetism. For side chains protons, the $R_{1,2}$ values were taken by considering for each type of proton, the average value obtained by taking into account only values within the standard deviation. Then $R_{1,2}$ (blue histograms in Figure 3) are converted into distances ($r_{ab}$) according to the Solomon-Bloembergen equations [122]

$$R_i = \frac{2}{15} \left( \frac{\mu_0}{4\pi} \right) ^2 \frac{1}{r_{\text{null}}} \left( \frac{S+1}{1 + (\omega_1 - \omega_2)^2 \tau_7^3} + \frac{3\tau_7}{1 + \omega_1^2 \tau_7^3} + \frac{6\tau_7}{1 + \omega_1^2 \tau_7^3} \right)$$

(2)

$$R_2 = \frac{1}{15} \left( \frac{\mu_0}{4\pi} \right) ^2 \frac{1}{r_{\text{null}}} \left( \frac{S+1}{1 + (\omega_1 - \omega_2)^2 \tau_7^3} + \frac{3\tau_7}{1 + \omega_1^2 \tau_7^3} + \frac{6\tau_7}{1 + \omega_1^2 \tau_7^3} \right)$$

(3)

In the equations, we considered $\tau_7 = 6 \times 10^{-12}$ s, arising from an estimated value of $\tau_7$ for each iron ion, $\tau_7 = 3 \times 10^{-9}$ s obtained from $^{15}$N relaxation, and $S=1$, arising from the lowest excited state of the electron spin ladder of the $[\text{Fe}_4\text{S}_4]^2^+$ cluster. Distances obtained from eqs (2-3) were then converted into upper limit distances (upl). For non-exchangeable protons, we obtained overall 122 upl. For amide HN resonances, when $R_1$ and $R_2$ provided different upper distance limits for the same $^1$H$_N$ proton, the upper limit value was taken by considering the less restrictive value among the two. In these cases, the upper limit value was given a weighting factor 2. For the 49 non-proline residues, 30 upper distance limits were used.

Since we do not have a priori the information on which iron of the cluster is causing the dipole coupling with each nucleus, in a first structure calculation all PRE-based upper limit distance restraints were assigned to the mass center of the cluster, rather than to a specific iron ion of the cluster. This was obtained by adding at the end of the protein sequence a special linker made of 100 pseudo-residues called LL2. The "atoms" of LL2 pseudo-residues have zero mass and zero Van-der-Waals radii, thus the linker can freely pass through the structure during simulated annealing. The last residue of the link-
er is an ION residue (CYANA library) which has been subsequently linked at fixed distances with the four iron ions and with the four sulfur ions of the cluster, with van der Waals contact taken to zero in order to avoid distortions or additional contribution to the overall energy. In this construct, the ION residue represents therefore the mass center of the cubane. Additional 1.4 Å (where 1.4 Å is the distance between each iron and the mass center of the cubane in the typical [4Fe-4S] structure) was added to all upl values obtained from PREs and all restraints were given to the center of mass of the cubane. To perform the final refinement via AMBER, for each PRE restraint the center of mass of the cubane has been replaced with the closest iron ion of the cluster as resulting from the structure obtained with CYANA (or with the two closest iron ions when ambiguous metal-to-proton distances occur) and the upper limit distance taken from eqs (2-3).

PREs that are violated in a significant number of structures were critically analyzed, taking also into account the NOES. It was found that, while no violations have been observed for PREs obtained from non-exchangeable protons, some of the PRE values of exchangeable HN gave rise to consistent violations. Some of them are due to local internal dynamics (15N relaxation shown in Figure S2). However, some of the R1 and R2 HN values belonging to the residues 5-20 (i.e. the N-terminal part preceding the cluster binding region) and to residues 42-44 (part of the long loop between Cys 34 and Cys 47) showed consistent violations that are not accounted by 15N relaxation. All these residues have calculated iron-to-proton distances in the range 7-11 Å, while no violations were observed for HN distances in the range 4-7 Å. This effect is observed only for exchangeable amide protons; indeed, aliphatic protons values were in very good agreement with the structure also at distances >10 Å. This suggests that magnetic susceptibility anisotropy is not the only factor responsible for the deviation from eqs 2-3 [58, 61], as also previously observed in Cu(II) proteins [59]. Intermolecular effects and/or solute-solvent interactions could be possible factors affecting the quantitative analysis of HN relaxation rates. PREs giving rise to consistent violations were excluded from structure calculation. Overall, 175 PRE restraints were retained from relaxation-rate data (Table S4), and the total number of distance restraints (NOE and PRE) was increased to 533. A summary of these restraints is reported in Table S5.
The addition of relaxation-based NMR restraints gave a more tightly converged NMR structure, with backbone and side chain RMSD values of 0.62±0.11Å and 1.14±0.13Å respectively (Figure 2B). PRE data improved the quality of the structure not only in the proximity of the cluster, but throughout the entire protein. The combination of both type of restraints led to statistical parameters that are indicative of a highly precise structure of a well-folded protein of small/medium size (Table S6). Finally, we addressed the question of whether an NMR structure obtained without NOEs is able to achieve good accuracy and precision. Figure 2C shows the family of structures obtained without the 344 NOEs from 13C and 15N-NOESY-HSQC experiments. The structure has backbone and heavy atoms RMSD of 1.31±0.27Å and 2.00±0.32Å, respectively. The overall precision is obviously lower than that obtained with the full set of restraints and it is also lower than that obtained with NOE-only, because of the lower number of restraints, but still lies within an acceptable structure quality range. The per-residue comparison of backbone RMSD (Figure 4A) shows that the family obtained with the full set of restraints has always the lowest RMSD (except Thr24), indicating that the combination of NOEs and PREs improves the precision in all the protein regions. PREs provide information exactly were NOEs are missing, thus complementing NOE data. In several protein regions, the NOE-only family has an RMSD similar to the family obtained with the full set of restraints, indicating that NOEs drive the structure towards a minimum. Conversely, in other protein regions, the precision is improved by the use of PREs; in these regions, the structure quality is PRE driven. The loop surrounding the cluster and containing Cys 22 and Cys 25 of the CXXC binding motif in HiPIPs has a different trend. Here, the RMSD values of the three families are similar and higher than average values. This is the situation in which, not only NOEs but also PREs are missing due to the close proximity to the paramagnetic center. Essentially, for this fragment the structure is given by the cluster binding topology, by the dihedral angles of Cys bound residues, and by the hydrogen-bonds identified by 15N contact shift. We can obtain clues on the accuracy of the structures by comparing the most representative structures of the three families: the PRE-only, the NOE-only and the full-set structures. As shown in Figure 4B, the pairwise RMSD between them are all similar and comparable with the precision within each family. The representative conformer selected from the ensemble obtained with the full-set of restraints is, for most of the protein, in an average position among the representative conformers of the PRE-only and NOE-only ensemble. The PRE-only
structure ensemble is essentially the same, although with a larger RMSD, as the one obtained with the full-set of restraints. Finally, Figure 4C points out the relative impact of PREs vs long-range NOEs. As expected, residues surrounding the cluster have a dominance of PREs, however also regions from a larger sphere, such as stretch 10-13 experiences the contribution of PREs. Noteworthy, an opposite behavior is observed for the aromatic residue surrounding the cluster Trp46, which experiences both NOEs and PREs. This large hydrophobic residue has the role of maintaining the hydrophobicity of the cluster and protecting it from solvent accessibility. Therefore, it gives rise to many long-range NOEs, extending from the close proximity of the paramagnetic center to the diamagnetic region.

Conclusions
Notwithstanding the exciting perspectives opened by computational biologists [123-126], the quest for novel experimental restraints remains of primary importance for structural modeling. Up to date, a dense network of NOEs has always been considered essential for NMR structures, because restraints between residues that are far apart in the primary sequence define the relative orientation of different structural motifs [127]. Factors such as protein size, electronic correlation times of metal ion(s) and internal mobility modulate the interplay between paramagnetism-based and conventional NMR restraints and their relative contribution to the final structure. The NMR structure of PioC is a proof of concept that PREs may drive a solution structure and eventually act as the sole source of NMR restraints. When a protein is small enough to be affected by paramagnetism in a large percentage, then NOEs are not essential anymore, if relaxation rates are measured virtually for all 'H spins. In PioC, the [Fe₄S₄]²⁺ cluster provides upper distance limits for PREs up to 13 Å, while the average protein radius is about 15 Å, thus being an ideal case for obtaining an accurate and well converged PRE-driven NMR structure. In this case, an extended network where all the 'H spins are linked to a single point (the metal center) via long-range dipolar couplings can completely replace a network of short-range dipole-dipole 'H-'H couplings. The availability of PREs from aliphatic protons circumvent the lack of accuracy for long metal to proton distances of exchangeable protons. This represents an opportunity for the characterization and structural study of metallopeptides and de-novo designed and bio-inspired metalloenzymes [128-130]. These results argue for the systematic use of PREs in structure calculations of metalloproteins.
where metal substitution is not possible since they provide distance restraints in protein regions where NOEs are sparse due to paramagnetism, and where most biochemically relevant events occur.

Material and Methods

Protein expression and purification

E. coli BL21 DE3 cells were transformed with pET32h, a plasmid containing the construct thioredoxin–6xHis–thrombin cleavage site–PioC, and with pDB1281, a plasmid that carries the machinery for the assembly of iron-sulfur clusters. Cells were grown in Luria-Bertani (LB) supplemented with 100mg*dm⁻³ ampicillin and 35mg*dm⁻³ chloramphenicol until the OD600nm of 0.6 where they were induced with 1.0 mM arabinose and 20 μM FeCl₃ and 200 μM cysteine were added. Cells were again incubated until the OD600nm of 1 and then harvested and washed in M9 minimal media salts before resuspension in M9 minimal media. Once re-suspended, cells were incubated for one hour before induction with 0.5mM IPTG. After 4hr cells were harvested by centrifugation and disrupted using a French Press at 1000psi. The lysate was ultra-centrifuged at 204 709g for 90 min at 4°C to remove cell membranes and debris and the supernatant was dialyzed overnight against 50 mM potassium phosphate buffer pH 5.8 with 300mM NaCl before injection in a His-trap affinity column (GE Healthcare). The fraction containing Histag-PioC eluted with 250mM imidazole and was incubated overnight with Thrombin (GE Healthcare) for digestion. The final purified PioC (His-tag free) was then concentrated from the flow through of a 2nd passage through the His-trap column using an Amicon Ultra Centrifugal Filter (Millipore) with a 3kDa cutoff. The purity of PioC was confirmed by SDS-PAGE with Blue Safe staining (NzyTech) and by UV-Visible spectroscopy. Three samples of PioC were produced (unlabeled, single ¹⁵N-labeled, double ¹³C & ¹⁵N-labeled) and the expression and purification protocol was identical throughout except in the addition of ammonium sulfate (¹⁶N₂, 99%) and [U-¹³C₆] D-glucose in the M9 minimal media when labeling was required.

NMR experiments
All experiments were recorded using Bruker AVANCE-NEO spectrometers, equipped with cryogenically cooled triple resonance inverse detection probeheads (CP-TXI), except $^{13}$C-detected experiments, which were acquired at 176.05 MHz using a cryogenically cooled probehead optimized for $^{13}$C direct detection (CP-TXO), and $^1$H experiments which were recorded at 400 MHz using a room temperature, selective 5mm $^1$H probe without pulsed field gradients. All spectra were processed using the Bruker software TopSpin. Standard radio frequency pulses and carrier frequencies for triple resonance experiments were used. The set of NMR experiments used for sequence specific assignment, NOE collection and $^{15}$N relaxation analysis is summarized in Table S1. To identify signals affected by the hyperfine interaction, tailored experiments were performed [64, 115, 118]. Experimental parameters are summarized in Table S2. Data analysis and resonances assignment were performed using CARA 1.9 [131]. The complete assignment has been submitted in BRMB entry 34487 [132].

**Structure Calculations**

Structure calculations were performed with the program CYANA 2.1 [133, 134]. NOEs were analyzed and converted into upper distance limits and used for manual structure calculation in CYANA 2.1. Backbone dihedral angle constraints were derived from $^{15}$N, $^{13}$C', $^{13}$C$\alpha$, $^{13}$C$\beta$, and H$\alpha$ chemical shifts, using TALOS+ and added as restraints. All structure ensembles presented here were refined with molecular dynamics using AMBER-16 [135] and force field parameters for the 4Fe4S cluster as previously described [136, 137] and validated using PDBstat and PSVS programs [138, 139]. Detailed description of the methodology used for structure calculation and refinement is reported in supplementary material.

**Author contributions:** R.L., M.P. designed research; I.T., M.I., F.C., and M.P. performed research; F.C., M.I., and I.T. analyzed data; and M.P., R.L., F.C., I.T., and M.I. wrote the paper.

**Acknowledgements**

We are grateful to Prof. Gaetano Montelione for the enlightening discussions during his stay at CERM as invited Professor. This work benefited from access to CERM/CIRMMP, the Instruct-ERIC Italy cen-
tre. Financial support was provided by European EC Horizon2020 TIMB3 (Project 810856) Instruct-ERIC (PID 4509). This article is based upon work from COST Action CA15133, supported by COST (European Cooperation in Science and Technology). Fondazione Ente Cassa di Risparmio di Firenze (CRF 2016 0985) is acknowledged for providing fellowship to MI. This work was funded by national funds through FCT– Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/04612/2020 and UIDP/04612/2020, Fundação para a Ciência e a Tecnologia (FCT) Portugal Grant PD/BD/135187/2017 to IBT.

References


Figure Legend

Fig. 1. A) 500 MHz 298K, 15N- HSQC spectrum on PioC collected using the HSQC-AP experiment (blue), overlaid with a standard 15N HSQC spectrum (red). Labeled signals are observable only in the HSQC-AP spectrum. Signals marked with asterisks are folded peaks arising from side chains. B) Cluster derived NMR restraints: dihedral angles $\chi_2$ of cluster-bound Cysteines (black), hydrogen bonds between HN residues and $S_2$ of Cysteines (red), $^1$H NOEs between well-resolved $^1$H fCH$_2$ Cys resonances and surrounding protons (green)

Fig. 2. Solution structure of PioC obtained using NOEs only (orange), the full set of NMR restraints (blue), PREs only (green). In all cases, the families of 20 conformers were obtained from Torsion Angle Dynamics (CYANA2.1) and refinement using molecular dynamics (AMBER-16 package). Residues 5-49 are shown.

Fig. 3. Longitudinal and transverse relaxation rates of amide and aliphatic protons. Horizontal red lines show the average diamagnetic values. Blue histograms are rates converted into PRE values. Values of Hβ are out of scale (see Table S3)

Fig. 4. (A) per residue RMSD values of the three different families. The relative contribution of the different set of constraints on a per residue basis is shown with the color code indicated in the Figure. (B) Superimposing of the most representative structures of each ensemble obtained with different sets of constraints. Figure reports also pairwise backbone RMSD values. (C) Regions where the number of PRE restraints exceeds that of long-range NOE restraints are shown in red and orange; regions were the opposite occurs are shown in blue and light blue.