



UNIVERSITÀ
DEGLI STUDI
FIRENZE

FLORE

Repository istituzionale dell'Università degli Studi di Firenze

NOE-less protein structures-an alternative paradigm in Highly Paramagnetic Systems

Questa è la versione Preprint (Submitted version) della seguente pubblicazione:

Original Citation:

NOE-less protein structures-an alternative paradigm in Highly Paramagnetic Systems / Trindade Ines, Invernici Michele, Cantini Francesca, Louro Ricardo, Piccioli Mario. - ELETTRONICO. - (2020), pp. 0-0. [10.48550/arXiv.2002.01228]

Availability:

This version is available at: 2158/1205006 since: 2024-06-25T11:50:50Z

Published version:

DOI: 10.48550/arXiv.2002.01228

Terms of use:

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

Publisher copyright claim:

(Article begins on next page)

NOE-less protein structures-an alternative paradigm in Highly Paramagnetic Systems

Inês Trindade[§], Michele Invernici[‡], Francesca Cantini[‡], Ricardo Louro^{§*}, Mario Piccioli^{‡*}

[§] Instituto de Tecnologia Química e Biológica António Xavier (ITQB-NOVA), Universidade Nova de Lisboa, Av. da República (EAN), 2780-157 Oeiras, Portugal, louro@itqb.unl.pt.

[‡] Magnetic Resonance Center and Department of Chemistry, University of Florence, Via L. Sacconi 6 50019 Sesto Fiorentino, Italy.

ABSTRACT: The current paradigm for protein structure determination by NMR spectroscopy has thus far been based on collecting distance restraints between pairs of nuclei in the form of nuclear Overhauser enhancements (NOE). These, have in some circumstances been supplemented with other information sources such as paramagnetic relaxation enhancements or residual dipolar couplings. Here, we report for the first time a protein structure determination by NMR without the use of NOEs. The protein PioC from *Rhodospseudomonas palustris* TIE-1 is an HIPIP where the 4Fe4S cluster is paramagnetic and provides the source of Paramagnetic Relaxation Enhancements (PRE) used as alternative distance constraints. Comparison of the family of structures obtained by NOE structural restraints, with that obtained by PRE and with the family of structures obtained by combining NOEs and PREs reveals that the pairwise RMSD between them are similar and comparable with the precision within each family. This work sets the stage for the structural characterization of small and dynamic paramagnetic metalloproteins opening a new paradigm in the use of NMR in structural biology.

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 stability and structural properties³⁻⁸. 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, amplitude and time-scale of local dynamics^{9,10} and hints on the electronic structure and oxidation states of the metal center¹¹, in conditions that mimic the physiological context. However, a significant part of the metalloproteome contains a paramagnetic metal ion that enhances nuclear relaxation in its vicinity posing a challenge for signal detection¹². In consequence, NMR solution structures of paramagnetic macromolecules have required a combination of classical structural constraints (NOE, residual dipolar couplings and secondary structure predictions) and paramagnetism-based constraints (cross-correlation, pseudocontact shifts and paramagnetic relaxation enhancements)¹³⁻¹⁶. High resolution NMR structures were obtained for paramagnetic metalloproteins¹⁷⁻²¹ and for proteins containing paramagnetic metal

binding tags²²⁻²⁶. The substitution of the paramagnetic center with a diamagnetic analogue has also been used to obtain NMR structures of metalloproteins^{27,28}. A number of alternative approaches have been proposed, in which sparse NOE and chemical shifts based structure prediction methods succeeded to obtain well-defined solution structures^{29,30}. Indeed, structures without NOEs have been solved when accurate orientational constraints are available from at least two full sets of rdc^{31,32}. However, NOEs still constitute the fundamental building block to sample the protein conformational space and provide a unique fold³³⁻³⁵. Indeed, networked interatomic distances are more powerful and robust than orientation constraints and dihedral angles. In the case of Ln³⁺-substituted Calbindin D_{9k}, it was found that even in the presence of a very large number of rdc and pcs a minimal number of about 30 NOEs was necessary to obtain a structure³⁶. Cross correlation rates, and especially that originating from the interaction between Curie-spin and dipolar coupling, may also contribute to solution structure calculation^{37,38}; in combination with orientational constraints measured on a diamagnetic analog, they succeeded to obtain a structure without NOEs³⁹. PREs are dipole-dipole restraints, like NOEs. Therefore, if a sufficient number of them is available and if they can be measured throughout the entire protein, they should sample the protein conformational space with an efficiency comparable to NOEs, even if they are a set of distance constraints all involving a single point, i.e. the paramagnetic center. An important drawback of PREs is that NMR experiments currently available to measure ¹H *R*₁ and *R*₂ rates⁴⁰ fail to provide PRE values in the proximity of the paramagnetic center, where many signals are broadened beyond detection. This can be overcome by developing experiments capable to measure with good precision fast relaxation rates. For longitudinal relaxation, the IR-HSQC-AP experiment⁴¹ is very effective to identify and to measure *R*₁ rates of fast relaxing signals. However, it has been reported that ¹H *R*₁ values are more susceptible than *R*₂ to internal motions and cross relaxation⁴². Therefore, accurate measurements of both rates are important to obtain reliable information on the metal-to-proton distances.

To explore this issue, the NMR solution structure of the small Iron-Sulfur protein PioC from *Rhodospseudomonas palustris* TIE-1⁴³ represents a paradigmatic challenge. PioC is a HiPIP (High Potential Iron Protein) and thus, it contains a 4Fe-4S cluster with a very high reduction potential (*E*⁰ = +450 mV vs SHE) being stable in the [Fe₄S₄]²⁺ oxidation state. The protein has 54 amino acids and it is the smallest HiPIP ever isolated. The structure is unknown, but homology modelling with other HiPIPs suggests that the protein has a compact globular structure. It is characterized by the absence of topologically relevant secondary structure elements and is essentially composed by a series of loops and turns wrapped around the 4Fe-4S cluster^{44,45}. The electronic properties of [Fe₄S₄]²⁺ in HiPIPs have been studied in detail since about 40 years^{4,46-52}. The magnetic coupling within the Fe₄S₄ cluster makes the electronic correlation times of the individual iron ions much shorter than isolated high spin Fe³⁺ or Fe²⁺ ion; nevertheless, paramagnetic

contributions to nuclear relaxation are significant for nuclei within a 10 Å sphere from the cluster⁵³. Therefore, PioC is a suitable system to address an important issue for inorganic structural biologists: can we improve methods for measuring relaxation rates in paramagnetic systems and use relaxation based NMR restraints⁵⁴ (or PREs, as they are now commonly defined^{40,55}) as the sole source of dipole-dipole restraints? Can we break the paradigm that NOEs are an essential step for solution structure calculation?

A standard ¹⁵N HSQC experiment on a PioC sample shows only 39 amide resonances out of expected 49 non-Proline residues; however a ¹⁵N IR-HSQC-AP experiment shows additional 10 resonances, this demonstrated that all HN signals can be detected, provided that experiments specifically designed to observe fast relaxing resonances are used (**Figure 1**). In order to perform the complete resonance assignment

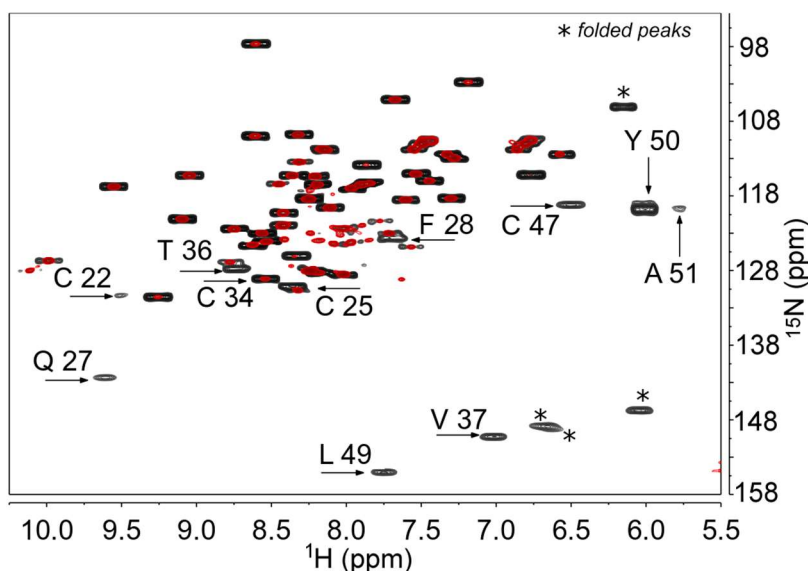


Figure 1. 500 MHz 298K, ¹⁵N- HSQC spectrum on PioC collected using the HSQC-AP experiment (black), overlaid with a standard ¹⁵N HSQC spectrum (red). Labeled signals are observable only in the HSQC-AP spectrum.

of the protein, we combined the conventional approach based on triple resonance experiments (**Table S1**) with a non-systematic procedure using a combination of 1D NOEs, ¹³C direct detection, double and triple resonance experiments recorded with parameters optimized “a-la-carte”⁵⁶. The non-systematic protocol for the complete NMR assignment is described in detail in SI (**Tables S2-S3** and **Figures S1-S4**).

Eventually, these experiments provided the complete NMR assignment

(**Table S4**). We assigned (excluding the N-ter Val 1) 100% of backbone ¹H_N, ¹³C, and ¹⁵N resonances, 98% of Hα, 86% and 91% of ¹H and ¹³C side chains atoms, respectively. However, even if the ¹H assignment is almost completed, ¹⁵N and ¹³C HSQC-NOESY experiments at high magnetic field gave only 344 meaningful NOEs that, without any additional information on cluster binding mode, were insufficient to obtain a structure. Three factors quench NOE intensities: i) the small rotational correlation time (3.4×10^{-9} s from ¹⁵N relaxation, **Figure S5**); ii) paramagnetic relaxation affecting at least 50% of the protein; iii) the absence of secondary structure elements, typical of HiPIPs. A few, crucial, structural constraints were obtained from residues binding the cluster, covalently or via H-bonds. Cys βCH₂ hyperfine shifts were converted into four χ₂ dihedral angle constraints defining the cluster binding topology⁵⁷, seven crucial 1D NOEs between Cys βCH₂ and neighboring residues were quantified (**Figure S2**), large ¹⁵N shifts values,

observed for Gln27, Val37 and Leu49 were taken as an evidence of three H-bonds, respectively linking HN to S γ of the preceding (i-2 or i-3) cluster-bound Cys. These fourteen constraints were included into structure calculation, together with the geometrical parameters of the cluster, obtained as previously de-

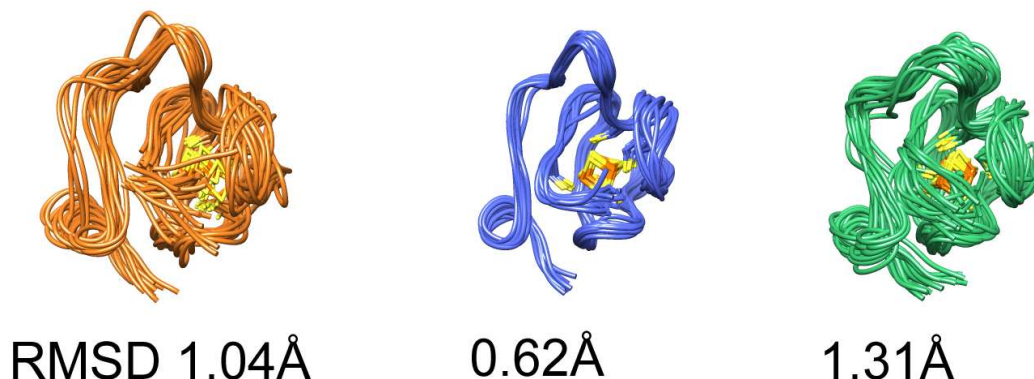


Figure 2. Solution structure of PioC obtained using NOEs only (orange), the full set of NMR restraints (blue), PREs only (green). In all cases, the families of 20 conformers were obtained from Torsion Angle Dynamics (CYANA2.1) and refinement using molecular dynamics (AMBER-16 package).

scribed¹⁹. We obtained, using torsion angle dynamics and refinement via molecular dynamics (see SI for details), a structure, shown in **Figure 2A**, with backbone and heavy atoms RMSD of 1.04 ± 0.30 Å and 1.81 ± 0.30 Å (residues 5-50). As expected, the cluster is the essential element to drive the fold of the polypeptide chain and the 14 constraints from cluster-bound residues are important to frame the cluster within the protein. Indeed, without them, we obtained structures with significantly higher backbone and heavy atoms RMSD of 1.27 ± 0.20 Å and 1.95 ± 0.20 Å. We now considered PREs: to this end, a novel R_2 -weighted ^{15}N -HSQC-AP experiments (**Figure S6**) was developed and used, together with the already described ^{15}N -IR-HSQC AP experiment⁸, to measure R_1 and R_2 values of all amide protons, as summarized in **Figure 3**. A ^{13}C version of the IR-HSQC-AP experiment was also developed and allowed to record ^1H relaxation rates for H α backbone protons as well as for aliphatic and aromatic side chains

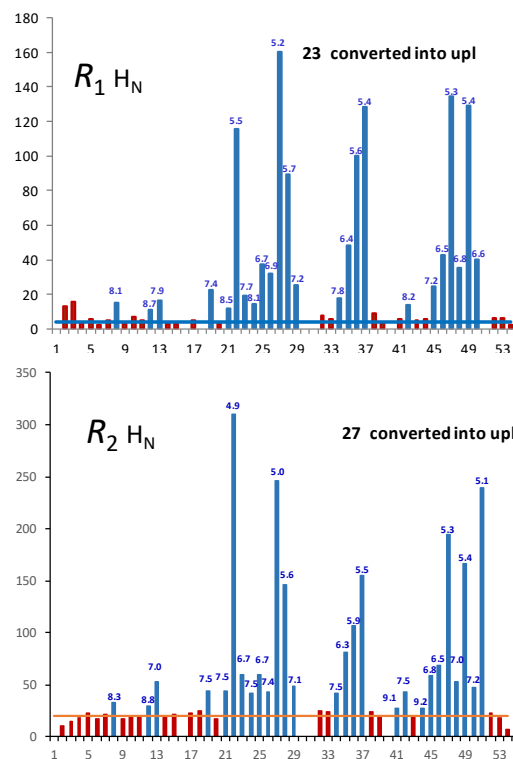


Figure 3. Longitudinal (upper) and transverse (lower) relaxation rates of amide H N signals. Horizontal lines show the threshold values taken as average diamagnetic value. Rates exceeding the threshold (blue-colored histograms) were converted into upper limit distance values, reported in Angstrom for each histogram.

(**Figure S7**). Finally, ^1H and ^{13}C resonances of cluster bound Cys residues were identified and assigned using rapid recycling experiments and from these twelve R_1 and R_2 values were obtained from inversion recovery and linewidth analysis of one dimensional ^1H and ^{13}C experiments. Overall, 312 relaxation rates, amounting to ca. six rates per residue, were measured, as summarized in **Table S7**. From the full set of relaxation data, $R_{1\text{para}}$ and $R_{2\text{para}}$ were found and converted into upper distance limits, following a procedure described in **SI**. In the case of $^1\text{H}_\text{N}$ signals, for which both R_1 and R_2 were available, the conversion of $R_{1\text{para}}$ and $R_{2\text{para}}$ into upper distance constraints was done, independently for each of the two measurements, according to an r^{-6} dependence. Where different upper limit distances were found from R_1 and R_2 , the less restrictive among them is considered, in order to minimize errors in upper limit conversion due to internal motions that may affect R_1 and R_2 to a different extent. For amide H_N resonances, 33 upper distance limits were obtained out of 49 non-proline residues, indicating that about 66% of the protein is affected by paramagnetism, while 137 upper distance restraints are obtained from H_C R_1 values and 5 from ^{13}C R_1 values. Overall, 175 constraints were obtained from relaxation rates. The number of distance constraints was increased to a total of 533. The use of relaxation based NMR restraints, together with those used previously (the summary of constraints is reported in **Table S5**), gave a high resolution NMR structure with backbone and side chain RMSD values of $0.62\pm 0.11\text{\AA}$ and $1.14\pm 0.13\text{\AA}$ respectively (**Figure 2B**). PREs improve 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**).

We now address whether an NMR structure obtained without NOEs is able to achieve good accuracy and precision. Indeed, **Figure 2C** shows the family of structures obtained without the 344 NOEs from ^{13}C and ^{15}N -NOESY-HSQC experiments. The structure has backbone and heavy atoms RMSD of $1.31\pm 0.27\text{\AA}$ and $2.00\pm 0.32\text{\AA}$, respectively. The overall structure precision is obviously lower than that obtained in previous calculations but still lies within an acceptable quality range. The per-residue comparison of backbone RMSD (**Figure 4A**) shows that the family obtained with the full set of constraints has always the lowest RMSD (but Thr24), indicating that the combination of NOEs and PREs improves the precision in all the protein regions. PREs provide information exactly where 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 constraints, indicating that NOEs drive the structure towards a minimum. Conversely, in other regions the precision is significantly improved by the use of PREs, thus we can define them as 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 H-bonded residues. We can obtain clues on the accuracy of the structures by comparing the three mean structures, the *PRE-only*, the

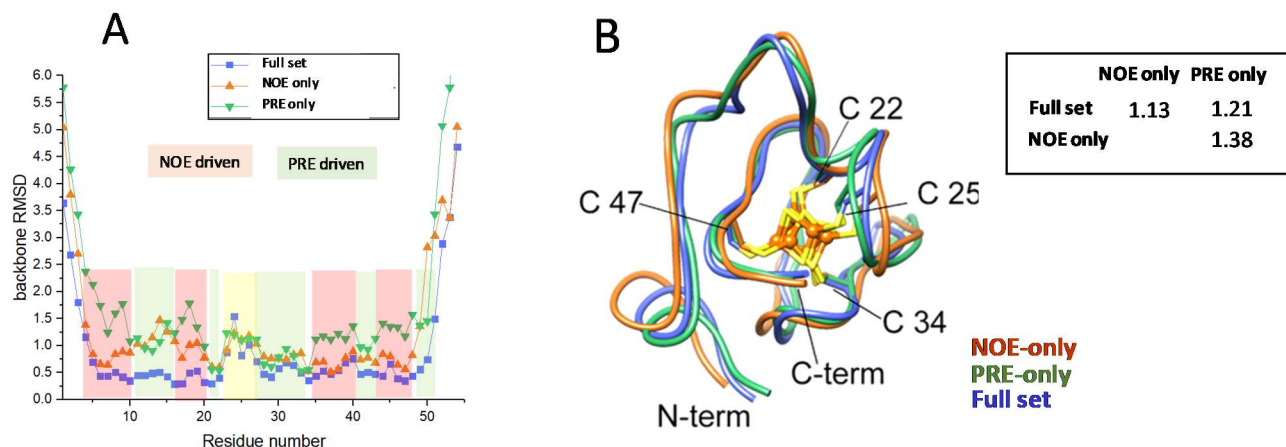


Figure 4. **A** per residue RMSD values of the three different families. The relative contribution of the different set of constraints on a per residue basis is shown with the color code indicated in the Figure. **B** Superimposing of the mean structures obtained with different sets of constraints. Figure reports also pairwise backbone RMSD values.

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 *PRE-only* structure is essentially the same, although with a larger RMSD, as the one obtained with the full-set of constraints. The latter is, for most of the protein, in an average position between *PRE-only* and *NOE-only*.

Notwithstanding the exciting perspectives opened by computational biologists⁵⁸⁻⁶¹, the quest for novel NMR restraints is still of primary importance in structural biology. Up to date, a dense network of NOEs has always been considered essential for structures, because constraints between residues that are far apart in the primary sequence define the relative orientation of different structural motifs⁶². Many authors introduced the use of PREs in this context and evoked the idea that they do have the property to replace NOEs in those protein regions where the paramagnetic effect is operative. However, until now, there are no cases reported in which NOE could be completely replaced by PREs. The NMR structure of PioC is a proof of concept of the above propositions: 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 ¹H spins. In this case, an extended network where all the ¹H 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 ¹H-¹H couplings. Noteworthy, PREs are unique also with respect to other paramagnetism-based restraints such as pseudocontact shifts and residual dipolar couplings. Noteworthy, PREs are unique

also with respect to other paramagnetism-based restraints such as pseudocontact shifts and residual dipolar couplings¹². The latter have a combination of angular and spatial dependence and thus, even when a large number of them are available, they are not able to replace NOEs completely. In PioC, the $[\text{Fe}_4\text{S}_4]^{2+}$ cluster provides upper distance limits up to 13 Å, while the average protein radius is about 15 Å, thus being an ideal case for obtaining the first NOE-less NMR structure. Besides the serendipity case of PioC, these results suggest the systematic use of PREs in structure calculations of metalloproteins because they provide distance restraints in protein regions where NOEs are sparse due to paramagnetism. Factors such as protein size, electronic correlation times of metal ion(s) and internal mobility will modulate the interplay between paramagnetism based and conventional NMR restraints and their relative contribution to the final structure. Whatever the combination will be, we may anticipate that, a combination of the largest possible set of constraints from both paramagnetic and conventional restraints will be the optimal solution in terms of structure quality.

DEPOSITION

The 20 conformers with the lowest target function constituted the final family and the chemical shift assignment were deposited to the Protein Data Bank (PDBID 6XYV) and to the BMRB (code 34487) respectively.

ACKNOWLEDGMENT

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 centre. 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. Project LISBOA-01-0145-FEDER-007660 (Microbiologia Molecular, Estrutural e Celular) funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI), Fundação para a Ciência e a Tecnologia (FCT) Portugal Grant PD/BD/135187/2017 to IBT

REFERENCES

- 1 Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. & Thornton, J. M. Metal ions in biological catalysis: from enzyme databases to general principles. *J. Biol. Inorg. Chem* 13, 1205-1218 (2008).
- 2 Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. Metalloproteins and metal sensing. *Nature* 460, 823-830 (2009).
- 3 Mulliez, E., Duarte, V., Arragain, S., Fontecave, M. & Atta, M. On the Role of Additional [4Fe-4S] Clusters with a Free Coordination Site in Radical-SAM Enzymes. *Frontiers in chemistry* 5, 17, doi:10.3389/fchem.2017.00017 (2017).
- 4 Beinert, H. Iron-sulfur proteins: ancient structures, still full of surprises. *J Biol. Inorg. Chem* 5, 2-15 (2000).
- 5 Rouault, T. A. The indispensable role of mammalian iron sulfur proteins in function and regulation of multiple diverse metabolic pathways. *Biometals* 32, 343-353, doi:10.1007/s10534-019-00191-7 (2019).
- 6 Ciofi-Baffoni, S., Nasta, V. & Banci, L. Protein networks in the maturation of human iron-sulfur proteins. *Metallomics* 10, 49-72 (2018).
- 7 Perard, J. & Ollagnier de Choudens, S. Iron-sulfur clusters biogenesis by the SUF machinery: close to the molecular mechanism understanding. *J Biol Inorg Chem* 23, 581-596, doi:10.1007/s00775-017-1527-3 (2018).
- 8 Cai, K. et al. Structural/Functional Properties of Human NFU1, an Intermediate [4Fe-4S] Carrier in Human Mitochondrial Iron-Sulfur Cluster Biogenesis. *Structure* 24, 2080-2091, doi:10.1016/j.str.2016.08.020 (2016).
- 9 Arthanari, H., Takeuchi, K., Dubey, A. & Wagner, G. Emerging solution NMR methods to illuminate the structural and dynamic properties of proteins. *Curr Opin Struct Biol* 58, 294-304, doi:10.1016/j.sbi.2019.06.005 (2019).
- 10 Bax, A. & Clore, G. M. Protein NMR: Boundless opportunities. *J Magn Reson* 306, 187-191, doi:10.1016/j.jmr.2019.07.037 (2019).
- 11 Piccioli, M. & Turano, P. Transient iron coordination sites in proteins: Exploiting the dual nature of paramagnetic NMR. *Coordination Chemistry Reviews* 284, 313-328, doi:10.1016/j.ccr.2014.05.007 (2015).
- 12 Parigi, G., Ravera, E. & Luchinat, C. Magnetic susceptibility and paramagnetism-based NMR. *Prog Nucl Magn Reson Spectrosc* 114-115, 211-236, doi:10.1016/j.pnmrs.2019.06.003 (2019).

- 13 Arnesano, F., Banci, L. & Piccioli, M. NMR structures of paramagnetic metalloproteins. *Q. Rev. Biophys* 38, 167-219 (2006).
- 14 Koehler, J. & Meiler, J. Expanding the utility of NMR restraints with paramagnetic compounds: background and practical aspects. *Prog. Nucl. Magn Reson Spectrosc* 59, 360-389, doi:S0079-6565(11)00041-0 [pii];10.1016/j.pnmrs.2011.05.001 [doi] (2011).
- 15 Hass, M. A. & Ubbink, M. Structure determination of protein-protein complexes with long-range anisotropic paramagnetic NMR restraints. *Curr. Opin. Struct. Biol* 24, 45-53, doi:S0959-440X(13)00206-6 [pii];10.1016/j.sbi.2013.11.010 [doi] (2014).
- 16 Sala, D. et al. Protein structure prediction assisted with sparse NMR data in CASP13. *Proteins* 87, 1315-1332, doi:10.1002/prot.25837 (2019).
- 17 Louro, R. O., Pacheco, I., Turner, D. L., LeGall, J. & Xavier, A. V. Structural and functional characterization of cytochrome c3 from *D. desulfuricans* ATCC 27774 by 1H-NMR. *FEBS Lett* 390, 59-62 (1996).
- 18 Turner, D. L., Brennan, L., Chamberlin, S. G., Louro, R. O. & Xavier, A. V. Determination of solution structures of paramagnetic proteins by NMR. *Eur Biophys J* 27, 367-375 (1998).
- 19 Banci, L. et al. The three dimensional structure in solution of the paramagnetic protein high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* through nuclear magnetic resonance. *Eur. J. Biochem* 225, 715-725 (1994).
- 20 Bertini, I. et al. Solution structure of the oxidized 2[Fe4S4] ferredoxin from *Clostridium pasteurianum*. *Eur. J. Biochem* 232, 192-205 (1995).
- 21 Kudhair, B. K. et al. Structure of a Wbl protein and implications for NO sensing by *M. tuberculosis*. *Nature communications* 8, 2280, doi:10.1038/s41467-017-02418-y (2017).
- 22 Iwahara, J., Schwieters, C. D. & Clore, G. M. Ensemble approach for NMR structure refinement against H-1 paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J. Am. Chem. Soc* 126, 5879-5896 (2004).
- 23 Miao, Q. et al. A Double-Armed, Hydrophilic Transition Metal Complex as a Paramagnetic NMR Probe. *Angew Chem Int Ed Engl* 58, 13093-13100, doi:10.1002/anie.201906049 (2019).
- 24 Liu, W.-M., Overhand, M. & Ubbink, M. The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins. *Coord. Chem. Rev* 273-274, 2-12, doi:10.1016/j.ccr.2013.10.018 (2014).
- 25 Nitsche, C. & Otting, G. Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags. *Prog Nucl Magn Reson Spectrosc* 98-99, 20-49, doi:10.1016/j.pnmrs.2016.11.001 (2017).

- 26 Cetiner, E. C. et al. Paramagnetic-iterative relaxation matrix approach: extracting PRE-restraints from NOESY spectra for 3D structure elucidation of biomolecules. *J Biomol NMR* 73, 699-712, doi:10.1007/s10858-019-00282-0 (2019).
- 27 Ramelot, T. A. et al. Solution NMR structure of the iron-sulfur cluster assembly protein U (IscU) with zinc bound at the active site. *J Mol Biol* 344, 567-583, doi:10.1016/j.jmb.2004.08.038 (2004).
- 28 Lamosa, P., Brennan, L., Vis, H., Turner, D. L. & Santos, H. NMR structure of *Desulfovibrio gigas* rubredoxin: a model for studying protein stabilization by compatible solutes. *Extremophiles* 5, 303-311, doi:10.1007/s007920100206 (2001).
- 29 Tang, Y. et al. Protein structure determination by combining sparse NMR data with evolutionary couplings. *Nat. Methods* 12, 751-754, doi:nmeth.3455 [pii];10.1038/nmeth.3455 [doi] (2015).
- 30 Raman, S. et al. NMR structure determination for larger proteins using backbone-only data. *Science* 327, 1014-1018 (2010).
- 31 Al-Hashimi, H. M. et al. Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J. Magn. Reson* 143, 402-406 (2000).
- 32 Kontaxis, G., Delaglio, F. & Bax, A. Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol* 394, 42-78, doi:10.1016/s0076-6879(05)94003-2 (2005).
- 33 Williamson, M. P., Havel, T. F. & Wüthrich, K. Solution Conformation of Proteinase Inhibitor IIA from Bull Seminal Plasma by ¹H Nuclear Magnetic Resonance and Distance Geometry. *J. Mol. Biol* 182, 295-315 (1985).
- 34 Wüthrich, K. *NMR of Proteins and Nucleic Acids*. (Wiley, 1986).
- 35 Billeter, M., Wagner, G. & Wüthrich, K. Solution NMR structure determination of proteins revisited. *J. Biomol. NMR* 42, 155-158 (2008).
- 36 Bertini, I. et al. Paramagnetism-based Versus Classical Constraints: An Analysis of the Solution Structure of Ca Ln Calbindin D9k. *J. Biomol. NMR* 21, 85-98 (2001).
- 37 Bertini, I. et al. Cross Correlation Rates Between Curie Spin and Dipole-Dipole relaxation in Paramagnetic Proteins: the Case of Cerium Substituted Calbindin D9k. *J. Biomol. NMR* 23, 115-125 (2002).
- 38 Kateb, F. & Piccioli, M. New routes to the detection of relaxation allowed coherence transfer in paramagnetic molecules. *J. Am. Chem. Soc* 125, 14978-14979 (2003).
- 39 Hus, J. C., Marion, D. & Blackledge, M. De novo Determination of Protein Structure by NMR using Orientational and Long-range Order Restraints. *J. Mol. Biol* 298, 927-936 (2000).

- 40 Donaldson, L. W. et al. Structural Characterization of Proteins with an Attached ATCUN Motif by Paramagnetic Relaxation Enhancement NMR Spectroscopy. *J. Am. Chem. Soc* 123, 9843-9847 (2001).
- 41 Ciofi-Baffoni, S., Gallo, A., Muzzioli, R. & Piccioli, M. The IR-15N-HSQC-AP experiment: a new tool for NMR spectroscopy of paramagnetic molecules. *J. Biomol. NMR* 58, 123-128 (2014).
- 42 Iwahara, J., Tang, C. & Clore, G. M. Practical aspects of ¹H transverse paramagnetic relaxation enhancement measurements on macromolecules. *J. Magn Reson* 184, 185-195 (2007).
- 43 Bird, L. J. et al. Nonredundant roles for cytochrome c2 and two high-potential iron-sulfur proteins in the photoferrotroph *Rhodospseudomonas palustris* TIE-1. *J Bacteriol* 196, 850-858, doi:10.1128/jb.00843-13 (2014).
- 44 Rayment, I., Wesenberg, G., Meyer, T. E., Cusanovich, M. A. & Holden, H. M. Three-dimensional structure of the high-potential iron-sulfur protein isolated from the purple phototrophic bacterium *Rhodocyclus tenuis* determined and refined at 1.5 Å resolution. *J. Mol. Biol* 228, 672 (1992).
- 45 Stelter, M. et al. Structure at 1.0 Å resolution of a high-potential iron-sulfur protein involved in the aerobic respiratory chain of *Rhodothermus marinus*. *J Biol Inorg Chem* 15, 303-313, doi:10.1007/s00775-009-0603-8 (2010).
- 46 Dickson, D. P. E. et al. Mossbauer Spectroscopic Studies of iron-Sulphur Proteins with four-iron active centres. *J. Physique Colloq* 37, C6-171-C176-175 (1976).
- 47 Emptage, M. H. et al. Nature of the Iron-Sulfur Centers in a Ferredoxin from *Azotobacter-Vinelandii* - Mossbauer Studies and Cluster Displacement Experiments. *J. Biol. Chem* 255, 1793-1796 (1980).
- 48 Blondin, G. & Girerd, J.-J. Interplay of electron exchange and electron transfer in metal polynuclear complexes in proteins or chemical models. *Chem. Rev* 90, 1359-1376 (1990).
- 49 Garcia-Serres, R., Clemancey, M., Latour, J. M. & Blondin, G. Contribution of Mossbauer spectroscopy to the investigation of Fe/S biogenesis. *J Biol Inorg Chem* 23, 635-644, doi:10.1007/s00775-018-1534-z (2018).
- 50 Mouesca, J.-M. & Lamotte, B. Iron-Sulfur clusters and their electronic and magnetic properties. *Coordination Chemistry Reviews* 178-180, 1573-1614 (1998).
- 51 Noodleman, L., Case, D. A., Mouesca, J.-M. & Lamotte, B. Valence electron delocalization in polynuclear iron-sulfur clusters. *JBIC* 1, 177-182 (1996).
- 52 Bertini, I. et al. Sequence specific assignment of ligand cysteine protons of oxidized, recombinant HiPIP I from *Ectothiorhodospira halophila*. *Inorg. Chem* 34, 2516-2523 (1995).

- 53 Banci, L., Camponeschi, F., Ciofi-Baffoni, S. & Piccioli, M. The NMR contribution to protein-protein networking in Fe-S protein maturation. *Journal of Biological Inorganic Chemistry* 23, 665-685, doi:10.1007/s00775-018-1552-x (2018).
- 54 Bertini, I. et al. The solution structure refinement of the paramagnetic reduced HiPIP I from *Ectothiorhodospira halophila* by using stable isotope labeling and nuclear relaxation. *Eur. J. Biochem* 241, 440-452 (1996).
- 55 Battiste, J. L. & Wagner, G. Utilization of site-directed spin labelling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited Nuclear Overhauser Effect data. *Biochemistry* 39, 5355-5365 (2000).
- 56 Balayssac, S., Jiménez, B. & Piccioli, M. Assignment Strategy for Fast Relaxing Signals: Complete Aminoacid Identification in Thulium Substituted Calbindin D9k. *J. Biomol. NMR* 34, 63-73 (2006).
- 57 Bertini, I., Capozzi, F., Luchinat, C., Piccioli, M. & Vila, A. J. The Fe₄S₄ centers in ferredoxins studied through proton and carbon hyperfine coupling. Sequence specific assignments of cysteines in ferredoxins from *Clostridium acidii urici* and *Clostridium pasteurianum*. *J. Am. Chem. Soc* 116, 651-660 (1994).
- 58 Yang, J. et al. Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci U S A* 117, 1496-1503, doi:10.1073/pnas.1914677117 (2020).
- 59 Cong, Q., Anishchenko, I., Ovchinnikov, S. & Baker, D. Protein interaction networks revealed by proteome coevolution. *Science* 365, 185-189, doi:10.1126/science.aaw6718 (2019).
- 60 Senior, A. W. et al. Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706-710, doi:10.1038/s41586-019-1923-7 (2020).
- 61 Grime, G. W. et al. High-Throughput PIXE as an Essential Quantitative Assay for Accurate Metalloprotein Structural Analysis: Development and Application. *J Am Chem Soc* 142, 185-197, doi:10.1021/jacs.9b09186 (2020).
- 62 Lopez-Mendez, B. & Guntert, P. Automated protein structure determination from NMR spectra. *J. Am. Chem. Soc* 128, 13112-13122 (2006).

NOE-less protein structures-an alternative paradigm in Highly Paramagnetic Systems

Inês Trindade[§], Michele Invernici[£], Francesca Cantini[£], Ricardo Louro^{§*}, Mario Piccioli^{£*}

[§] Instituto de Tecnologia Química e Biológica António Xavier (ITQB-NOVA), Universidade Nova de Lisboa, Av. da República (EAN), 2780-157 Oeiras, Portugal, louro@itqb.unl.pt.

[£] Magnetic Resonance Center and Department of Chemistry, University of Florence, Via L. Sacconi 6 50019 Sesto Fiorentino, Italy.

Supplementary Material

1. Protein expression and purification.

PioC was expressed and purified as previously reported [1]. 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 (¹⁵N₂, 99%) and [U-¹³C₆] D-glucose in the M9 minimal media when labeling was required. BL21 DE3 cells were double 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 OD_{600nm} of 0.6 where they were induced with arabinose and FeCl₃ and cysteine were added. Cells were again incubated until the OD_{600nm} 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 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.

2.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 ^1H experiments which were recorded at 400 MHz using a room temperature, selective 5mm ^1H 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**.

Table S1

Experiments	Time Domain Data Size			Spectral width			ns	Delay time	Magnetic field
	(points)			(ppm)				(s)	(MHz)
	t ₁	t ₂	t ₃	F ₁	F ₂	F ₃			
[¹ H- ¹ H]-NOESY ^a	1024	2048		14.6 (¹ H)	14.6 (¹ H)		48	1.2	900
[¹ H- ¹ H]-TOCSY	600	2048		14.0 (¹ H)	14.0 (¹ H)		48	2.0	600
¹ H- ¹⁵ N-HSQC	128	2048		40.0 (¹⁵ N)	13.2 (¹ H)		2	1.5	500
¹ H- ¹³ C-HSQC	320	1024		80.0 (¹³ C)	13.0 (¹ H)		20	2.0	600
cbcaconh	128	48	1800	80.0 (¹³ C)	40.0 (¹⁵ N)	13.2 (¹ H)	8	1.3	900
cbcanh	128	48	1800	80.0 (¹³ C)	40.0 (¹⁵ N)	13.2 (¹ H)	16	1.3	900
hnco	72	48	1800	16.0 (¹³ C)	40.0 (¹⁵ N)	13.2 (¹ H)	8	1.3	900
hncaco	72	48	1800	16.0 (¹³ C)	40.0 (¹⁵ N)	13.2 (¹ H)	24	1.3	900
hnca	96	48	1800	40.0 (¹³ C)	40.0 (¹⁵ N)	13.2 (¹ H)	16	1.3	900
hbhanh	128	40	2048	13.2 (¹ H)	40.0 (¹⁵ N)	13.2 (¹ H)	16	1.3	900

hnhha	48	128	2048	40.0 (¹⁵ N)	13.2 (¹ H)	13.2 (¹ H)	16	1.3	900
(H)CCH-TOCSY	128	64	1800	80.0 (¹³ C)	80.0 (¹⁵ N)	13.2 (¹ H)	16	1.5	900
¹⁵N-edited [¹H-¹H]- NOESY	128	40	2048	13.2 (¹ H)	40.0 (¹⁵ N)	13.2 (¹ H)	32	1.2	900
¹³C-edited [¹H-¹H]- NOESY	192	64	2048	14.0 (¹ H)	80.0 (¹³ C)	14.0 (¹ H)	32	1.2	950
¹⁵N R₁^b	128	1024		41.0 (¹⁵ N)	15.6 (¹ H)		16	3.0	500
¹⁵N R₂^c	128	1024		41.0 (¹⁵ N)	15.6 (¹ H)		16	3.0	500
¹⁵N- NOE	144	1024		41	15.6		56	5.0	500
¹HN R₁	220	1024		80.0 (¹⁵ N)	21.7 (¹ H)		16	4.0	500
¹HN R₂	156	1024		80.0 (¹⁵ N)	21.7 (¹ H)		64	4.0	500
¹HC R₁	320	1024		80.0 (¹³ C)	13.0 (¹ H)		20	2.0	600
CON	128	1024		40.6 (¹⁵ N)	47.0 (¹³ C)		64	1.0	700

3.Paramagnetism-tailored NMR Experiments

To identify signals affected by the hyperfine interaction, tailored experiments were performed. Most of them are experiments, commonly used in biomolecular NMR spectroscopy, where the standard set of parameters is modified to account for the shift and relaxation properties of paramagnetic systems. The pulse sequences are modified to choose those building blocks that are more “robust” with respect to fast nuclear relaxation, and the magnetic field is chosen accordingly. A few experiments have been specifically designed to identify resonances/dipolar or scalar couplings/relaxation properties of paramagnetic systems [2-4]. For each of them, typical parameter choices are summarized in **Table S2**

Table S2

Experi- ments	Time Domain Data Size		Spectral width		ns	Delays	Magnetic field
	(points)		(ppm)			(s)	(MHz)
	t ₁	t ₂	F ₁	F ₂			
¹ H-NOE	8192		80.0 (¹ H)		400000	aq= 133 ms saturation=110ms	400
IR- ¹⁵ N-HSQC-AP	220	1024	80.0	21.7 (¹ H)	256	aq=47ms recyc=0.15s invrec=0.2s δ=700us	500
R ₂ -weighted ¹⁵ N HSQC-AP	256	1024	80.0 (¹⁵ N)	21.7 (¹ H)	256	aq=47ms recyc=0.15s δ=2.0ms	500
¹ H- ¹³ C-HNCA	160	1024	100.0 (¹³ C)	25.0 (¹ H)	512	aq=40ms Recyc=0.5s δ=2.0ms	500
¹ H- ¹⁵ N-HNCA	128	1024	70.0 (¹⁵ N)	25.0 (¹ H)	512	aq=40ms Recyc=0.5s δ=2.0ms	500
1d-¹³C	8192		245.6 (¹³ C)		32768	aq= 71ms recyc=73ms	900
[¹³ C- ¹³ C]-COSY	500 (400)	2048 (1500)	258.2 (¹³ C)	258.2 (¹³ C)	480	aq=22ms recyc= 0.2s	700
IR- ¹³ C-HSQC-AP	256	2048 (512)	136.0 (¹³ C)	60.2 (¹ H)	1024	aq=34ms recyc=0.15s δ=2.0ms Invrec=30ms	500
¹ HC R ₁ IP acq	320	1024	136.0 (¹³ C)	60.2 (¹ H)	320	aq=17ms recyc=0.12s invrec=80ms δ=600us	500
¹ HC R ₁ AP acq	320	1024	136.0 (¹³ C)	60.2 (¹ H)	1024	aq=17ms recyc=34ms invrec=80ms δ=600us	500

4. Assignment of hyperfine shifted signals

In order to identify signals affected by the hyperfine interaction, we recorded a paramagnetic-tailored HNCA experiment (see previous section). Thanks to this HNCA experiment (**Figure S1A**), H_N , N_H and $C\alpha$ resonances of Cys22, Cys25 and Cys34 have been assigned due to the sequential connectivities observed with HN signals of residues Gly23, Arg26 and Ile35, unambiguously identified with the standard assignment strategy. **Figure S1B** shows the superimposition of two paramagnetic ^{13}C -HSQC-AP experiments, at 288 K (black lines) and at 298 K (red lines). The Anti-Curie temperature dependence provides the identification of $4-\beta\text{CH}_2$ and $3-\alpha\text{CH}$ of the ligand cysteines. Through the cross peaks obtained by ^{13}C -HSQC experiment and the cysteines $C\alpha$ assigned by HNCA, the $H\alpha$ resonances of Cys22, Cys25 and Cys34 have been assigned.

The ^1H NMR spectrum (**Figure S2A**) contain two signals experiencing paramagnetic line broadening, downfield hyperfine shift (16.6 ppm and 14.5 ppm) and antiCurie temperature dependence, showing the typical fingerprint of FeS protein binding a $[\text{Fe}_4\text{S}_4]^{2+}$ cluster, where the two peaks arise from Cys $H\beta$ signals [5]. Traces B-C show the NOE difference experiments performed upon selective irradiation of signal at 16.6 ppm and 14.5 ppm, respectively. They allow us to detect and measure Nuclear Overhauser enhancements (NOEs) between these two protons and their environment. From the selective saturation of the $H\beta$ signal at 16.6 ppm (**Figure S2B**), a strong NOE is observed with its geminal partner at 6.60 ppm (f). Other NOEs are observed with Gly 23 H_N (c), with Arg 21 $H\beta$ (h) and Arg 21 $H\gamma$ (i); this network of inter-residue connectivities allows us to assign the signal at 16.6 ppm as the Cys 22 $H\beta_2$ signal. The assignment of Cys 22 $H\beta_1$ at 6.60 ppm is also confirmed by a cross peak, observed in a 2D NOESY (data not shown), between the peak at 6.60 ppm and the previously assigned Arg 21 $H\epsilon$. Additionally, this NOE experiment also allows us to identify spatial proximity with $H\delta$ and $H\epsilon$ of Phe 28 (e,f) and with Glu 43 $H\alpha$ (g). Upon irradiation of $H\beta$ signal at 14.5 ppm (**Figure S2C**), a strong NOE is observed to the HN resonance of Cys 34 at 8.55 ppm (j), already identified from the HNCA experiment and identifies signal b as the $H\beta_2$ of Cys 34. In addition, Cys 34 $H\beta_2$ (14.5 ppm) gives NOEs to $H\alpha$ Val 37 and $H\gamma$ Ile 41. The assignment of Cys 22 and Cys 34 $H\beta$ signals allows to identify the $C\beta$ from the ^{13}C HSQC-AP experiments shown in **Figure S1B**. HNCO experiments (data not shown) provide the assignment of all C' from the cross peaks with the related already assigned $\text{HN}(i+1)$ nuclei.

^{13}C direct detection provides precious spectroscopic information. The 1D ^{13}C experiments are shown in **Figure S3A**. Signals belonging to Cys bound residues are expected to experience a strong antiCurie Temperature dependence, this allows to confirm the assignment of Cys 22 and Cys 34 $C\alpha$ and $C\beta$

signals and of Cys 25 C α . Additionally, two other Cys C β signals are observed at 111.1 and 102.65 ppm. C-C COSY has proved to be a powerful technique able to identify all C α /C' connectivities of cluster bond Cysteine. A tailored choice for $t_{1\max}$ and $t_{2\max}$ values provided [6], for the upper diagonal part, the spectrum shown in **Figure S3B**, which confirmed the assignment of C α of cysteine 22, 25 and 34 and identified the missing C α Cys 47, unobserved in the 1D ^{13}C experiments because in overlap with Cys34 and Cys22 C α and unobserved also in the ^{13}C -HSQC-AP experiment shown previously. The set of ^{13}C direct detected experiments is completed by a CON experiment (**Figure S3C**), which confirms the cysteine amide nitrogen assignment for Cys 22, 25 and 34 and provides the assignment for the amide nitrogen of Cys 47.

The information obtained from the paramagnetic ^{15}N -HSQC-AP experiment have been already discussed in the text. In particular, the amide HN of Cys 47 escaped detection also in the “optimized” version of HNCA and HNCO experiments. This accounts for a very efficient R_2 ^1H relaxation of the amide proton that would prevent the identification of its HNCA and HNCO peaks. Indeed, the ^{15}N -HSQC-AP experiment shows a very broad, upfield shifted HN amide signal (6.47/118.9 ppm), with a ^{15}N shift consistent with the assignment of Cys 47 amide, obtained with CON (C' of Trp 46 was assigned via HNCACO from the HN of Trp 46). Consistently, the H N signal at 6.50 ppm has ^1H T_2 values of 5.1 ms, accounting for an HN belonging to the first coordination sphere of the cluster without any connectivity in triple resonance experiments and, therefore, it can be safely assigned as Cys 47. The four Cysteines HN peaks are pointed out in **Figure S4**. All of them are affected by paramagnetic relaxation; however paramagnetism affects these four signals at different extent, as qualitatively observed here from relative peak intensities in the ^{15}N -HSQC-AP and quantitatively confirmed by ^1H R_1 and R_2 relaxation data (see later). Cys 25 and Cys 34 have relatively sharp HN resonances, observable even in a HNCA experiment, suggesting a Cysteine orientation in which the amide group is pointing far from the cluster. On the other hand, HN of Cys 47 and of Cys 22 are observable in a ^{15}N HSQC only at very short INEPT transfer delays, indicating that the HN vector is, in these cases, pointing towards the cluster. This is fully consistent with what observed in our NMR structure. Indeed, this feature is highly conserved throughout the very many HiPIPs of known structure [7-11].

The two βCH_2 protons of Cys 25 and Cys 47 can be discriminated by analyzing relaxation rates of H β protons, as measured from a ^{13}C IR-HSQCP-AP experiment. The ^{13}C HSQC-AP shows the βCH_2 groups (**Figure S2B**) of Cys 25 and Cys 47 that are not yet sequence specifically assigned. Measured R_1 ^1H rates are 313 s^{-1} and 282 s^{-1} for the two protons bound to the C β at 103.8 ppm and 450 s^{-1} and 130 s^{-1} for

the two protons bound to the C β at 88.3 ppm. Conversion of R_1 ^1H rates into upper distance limits gives values of 3.8 Å and 3.9 Å for the two protons bound to the C β at 103.8 ppm, while the two protons bound to the C β at 88.3 ppm give metal-to-proton distances of 3.6 Å and 4.5 Å. The inspection of X-ray and NMR structures of the many HiPIPs characterized so far, invariantly shows that the cluster binding topology of Cys II give rise to very similar metal-to-proton distances for the two H β protons, while the two H β of Cys IV are always asymmetric, with one H β close to the cluster and the other much farther apart. This is a strong indication that the symmetrical βCH_2 group with C β at 103.8 ppm belongs to Cys 25 while βCH_2 protons with asymmetric distances, with C β at 88.3 ppm, are assigned to Cys 47. Overall, the assignment of Cys resonances and the key experiments supporting their assignment is summarized in **Table S3**.

Table S3. Sequence-Specific Assignments of Cysteine resonances. For each resonance, color code indicates the experiment used to perform the assignment.

	Chemical shift (ppm)							
	HN	N	C α	H α	C β	H β 1	H β 2	CO
Cys22	9.47	131.4	89.5	8.18	102.65	6.6	16.6	177.6
Cys25	8.36	130.3	80.0	3.82	103.8	8.09	9.94	176.5
Cys34	8.55	129.1	90.3	3.83	111.1	5.8	14.5	173.9
Cys47	6.50	118.8	90.3		88.3	7.99	8.95	170.7

Table legend

HNCA	$^{13}\text{CHSQC}$	1D-NOE	2D-NOESY	Relaxation analysis
C-C COSY	CON	$^{15}\text{N-HSQC}$	HNCO	

4.Sequence Specific Assignment

Data analysis and resonances assignment were performed using CYANA 2.1. Proton resonances were calibrated with respect to the signal of 2,2-dimethylsilapentane-5-sulfonic acid (DSS). Nitrogen chemical shifts were referenced indirectly to the ^1H standard using a conversion factor derived from the ratio of NMR frequencies. Carbon resonances were calibrated using the signal of dioxane at 69.4 ppm (298 K) as secondary reference. The complete assignment is reported in **Table S4**. All parameters of the experiments used are reported in Tables S1 and S2.

Table S4 ^1H , ^{15}N and ^{13}C resonance assignments for PioC at 298 K, in 50 mM phosphate, 300 mM KCl at pH 5.8.

aa	N (HN)	CO	C α (H α)	C β (H β)	Others
Val-1		176.34	62.36	32.34	
Thr-2	117.2 (8.26)	174.0	61.8 (4.26)	69.78 (4.12)	21.67 (1.13)
Lys-3	123.7 (8.42)	176.0	55.7 (4.32)	33.0 (1.69)	C γ 24.3 (1.38) C δ 28.9 (1.63) C ϵ 41.68 (2.96)
Lys-4	121.8 (8.43)	176.0	56.7 (4.14)	34.7 (1.45, 1.37)	C γ 26.4 (1.32,0.95) C δ 29.9 (-) C ϵ 42.3 (2.91, 2.74)
Ala-5	122.3 (8.76)	177.5	51.4 (4.51)	19.2 (1.56)	
Ser-6	115.2 (9.08)	175.0	56.9 (4.41)	64.9 (4.32, 4.07)	
His-7	116.3 (9.64)	177.6	58.7 (4.24)	28.2 (3.39, 3.32)	C δ 2 117.5 (6.69) C ϵ 1 136.8 (8.24)
Lys-8	116.5 (8.49)	179.6	59.4 (4.19)	32.4 (1.90, 1.74)	C γ 24.3 (1.49,1.41) C δ 29.2 (1.71, 1.69) C ϵ 41.7 (2.98)
Asp-9	119.5 (8.13)	178.2	57.2 (4.36)	40.0 (2.85, 2.71)	
Ala-10	118.2 (8.23)	176.2	55.7 (4.28)	27.2 (1.91)	
Gly-11	105.0 (7.69)	175.1	46.8 (3.90, 3.73)		
Tyr-12	116.3 (8.19)	174.8	59.1 (4.51)	39.3 (2.56, 2.37)	C δ 135.25 (6.88) C ϵ 120,59 (7.33)
Gln-13	124.5 (8.67)	172.93	52.6 (4.66)	32.5 (2.11, 1.92)	C γ 32.7 (2.48, 2.77) N ϵ 2 112.36 (7.35, 6.58)
Glu-14	115.2 (8.22)	174.7	56.3 (4.09)	29.1 (2.28, 1.81)	C γ 36.8 (2.35)
Ser-15	112.8 (7.29)	171.8	56.2 (4.99)	63.9 (3.73, 3.68)	
Pro-16	130.36	176.1	63.0 (4.15)	33.3 (1.94,1.68)	C γ 27.6 (1.87,1.69) C δ 51.3 (3.40, 4.03)
Asn-17	116.4 (7.92)	175.7	50.0 (4.41)	36.2 (1.04, 0.69)	N δ 2 106.15 (6.17, 3.93)
Gly-18	111.6 (8.17)	174.1	46.6 (3.38,3.71)		
Ala-19	126.9 (8.78)	178.0	52.2 (4.21)	19.2 (1.3)	

Lys-20	118.4 (7.62)	173.1	56.5 (3.86)	32.1 (1.54, 0.26)	C γ 24.5 (1.30, 1.13) C δ 29.2 (1.59, 1.40) C ϵ 42.0 (2.95, 2.94)
Arg-21	115.87 (7.47)	176.8	53.5 (6.13)	32.9 (2.43, 1.80)	C γ 26.5 (1.69, 1.45) C δ 43.6 (3.50, 3.00) N ϵ 84.94 (7.34)
Cys-22	131.4 (9.47)	177.6	89.5 (8.18)	102.65 (6.60, 16.6)	
Gly-23	97.54 (8.62)	176.0	46.2 (3.91,3.81)		
Thr-24	109.8 (8.62)	173.6	60.8 (4.69)	69.1 (4.95)	C γ 21.03 (1.18)
Cys-25	130.0 (8.36)	176.5	80.0 (3.86)	103.8 (9.94,8.09)*	
Arg-26	131.4 (9.28)	179.3	58.9 (4.46)	31.0 (1.94,1.84)	C γ 27.1 (1.72,1.65) C δ 43.5 (3.15, 3.09) N ϵ 82.5 (6.96)
Gln-27	141.7 (9.62)	175.4	59.5 (4.36)	32.0 (2.36)	C γ 28.0 (2.47, 1.99)
Phe-28	123.4 (7.72)	178.0	63.7 (4.40)	52.3 (3.21)	C δ 134.0 (7.13) C ϵ 130.38 (7.0) C ζ 129.16 (7.12)
Arg-29	128.2 (8.19)	171.1	49.9 (4.59)	29.9 (1.28, 1.13)	C γ 25.5 (1.49, 1.40) C δ 42.8 (3.34, 2.91) N ϵ 83.3 (7.22)
Pro-30		176.2	61.2 (3.39)	30.3 (2.23, 1.71)	C γ 26.2 (1.88, 1.78) C δ 50.5 (3.52, 2.79)
Pro-31	128.9	177.9	64.3 (4.40)	35.0 (2.40, 2.12)	C γ 24.5 (1.91, 1.85) C δ 50.3 (3.59, 3.33)
Ser-32	111.8 (8.15)	174.34	57.4 (5.43)	64.5 (4.05, 3.83)	
Ser-33	114.9 (7.55)	169.5	57.2 (4.69)	66.0 (3.57, 3.47)	
Cys-34	129.0 (8.55)	173.8	90.0 (3.83)	111.1 (14.48,5.75)	
Ile-35	126.0 (8.36)	178.0	64.1 (4.16)	38.5 (2.03)	C δ 1 12.9 (0.91) C γ 1 26.9 (1.67, 1.46) C γ 2 18.8 (1.00)
Thr-36	127.6 (8.75)	173.1	62.3 (4.30)	69.7 (4.06)	C γ 2 21.7 (1.16)
Val-37	149.9 (7.00)	173.9	63.0 (5.02)	36.8	C γ 1 26.0 (--) C γ 2 23.8 (1.24)
Glu-38	122.8 (8.61)	174.8	56.9 (3.79)	31.1 (1.82, 1.73)	C γ 36.2 (2.19, 2.15)
Ser-39	115.7 (8.43)	171.4	59.0 (4.06)	62.7(4.07, 3.54)	
Pro-40	136.1	174.3	62.6 (4.58)	34.8 (2.34, 2.02)	C γ 24.2 (1.92, 1.89) C δ 50.6 (3.41, 3.28)
Ile-41	118.2 (7.28)	175.5	58.2 (4.10)	39.3 (1.34)	C δ 1 20.8 (0.89) C γ 1 26.9 (1.04, 0.28) C γ 2 21.0 (0.09)
Ser-42	116.4 (7.84)	178.3	55.2 (4.87)	65.0 (3.96,3.64)	
Glu-43	121.2 (9.12)	176.0	58.7 (3.32)	29.2 (2.11, 1.96)	C γ 34.9 (2.24, 2.04)
Asn-44	116.8 (7.98)	172.2	52.9 (5.00)	40.1 (3.00, 2.42)	N δ 2 111.7 (7.56, 6.88)
Gly-45	102.6 (7.20)	173.0	45.5(4.14,3.77)		
Trp-46	113.7 (7.86)	172.6	65.5 (5.19)	31.1 (2.85, 3.87)	N ϵ 1 124.0 (8.53) C δ 1 126.0 (6.66) C ζ 2 114.72 (7.39) C ζ 3 122.07 (6.91) C η 2 124.26 (6.88) C ϵ 3 119.9 (6.74) C ϵ 2 139.48 (--)

Cys-47	118.9 (6.47)	170.6	90.0	88.3 (7.99,8.95)	
Arg-48	115.2 (6.79)	180.22	57.4 (3.60)	29.8(1.62)	C γ 27.9 (1.53, 1.66) C δ 43.5 (3.18, 3.11) N ϵ 82.4 (7.26)
Leu-49	154.7 (7.74)	180.9	58.2 (4.13)	42.4	C γ 30.2 (1.30) 34.8 (1.40) 21.9
Tyr-50	119.5 (6.0)	173.7	62.4		C δ 137.2 (6.30) C ϵ 120.1 (7.15)
Ala-51	119.0 (5.46)	176.09	50.8 (4.26)	20.7 (1.01)	
Gly-52	109.7 (8.33)	174.5	45.8 (3.68,3.99)		
Lys-53	120.2 (8.43)	175.54	57.0 (4.14)	33.4 (1.79)	C γ 25.2 (1.38) C δ 29.7 (1.66) C ϵ 38.6 (-)
Ala-54	128.5 (8.04)	182.5	53.8(4.10)	20.2(1.29)	
*Stereo specific assignment of β CH ₂ not available					

5. Structural constraints and structure calculation

Diamagnetic NMR restraints. 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 , $^{13}\text{C}'$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, and $\text{H}\alpha$ chemical shifts, using TALOS+ and added as restraints. Overall, 344 meaningful upper distance constraints and 51 dihedral angles were used to calculate the structure. The data are summarized in **Table S5**.

Paramagnetic NMR restraints. Hyperfine shifts of Cysteines βCH_2 protons were converted into Cysteines dihedral angles χ_2 according to the procedure already described [12,13]. 1D NOEs observed from well resolved, hyperfine shifted signals were measured according to a procedure originally describe in [14] and converted into upper distance limits provided the known distance among geminal βCH_2 protons.

Hydrogen bond donor atoms were identified by considering ^{15}N shift values that are outliers by more than 20 ppm the average values according to BMRB Data Bank. Three H-bonding donor HN groups were identified. The three acceptors sulfur atoms were unambiguously identified from the NMR structure obtained without H-bonds and the three H-bonds added as structural constraints.

R_1 and R_2 rates of $^1\text{H}_\text{N}$ and $^1\text{H}_\text{C}$ were converted into upper distance limits as described in the following section. When R_1 and R_2 provided different upper distance limits for the same $^1\text{H}_\text{N}$ 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. Overall, 39 values were taken from H_N , out of which 19 were weighted by a factor 2; 31 upper distance limits were taken from $\text{H}\alpha$ and 90 upper distance limits were taken from side chains (HC groups). Additionally, 10 upper limit values were taken by relaxation rate measurements of βCH_2 and αCH from the four cluster-bound Cysteines. The total number of ^1H based PRE restraints was 170. In addition, the relaxation rates of Cysteines $\text{C}\alpha$ and $\text{C}\beta$ carbon atoms were also measured and converted into 5 additional PRE based restraints. Overall, the number of relaxation based restraints used into structure calculation was 175, as summarized in **Table S5**.

Definition of the cluster for structure calculations. The Iron sulfur cluster was inserted into structure calculation according to the procedure originally described [10]. A special residue, named CFS, was added to the CYANA library. The artificial residue, denoted CFS, consists of a cysteinyl residue in which the thiol hydrogen (H_γ) was replaced by an iron atom ($\text{Fe}\delta$) at the proper distance and by adding to the latter, through another covalent bond, the sulfur atom ($\text{S}\epsilon$) constituting the inorganic sulfide of the cluster. Bond lengths and angles used in this construction were taken from previously reported structures

[10,15,16]. Eight additional covalent bonds were added as link statements to the end of the sequence file between each iron atoms (Fe) and the two bonded sulfur atoms (S ϵ). This removes the van der Waals interactions between the Fe and the other ligands. Then, upper and lower distance limits are imposed along the eight edges of the cubane (the remaining four are defined within the four CFS residues), along the six edges of the tetrahedron described by the four iron atoms, six others along the edges of the tetrahedron described by the four inorganic sulfur atoms (S ϵ) and finally, among the six edges of the tetrahedron formed by cysteine S γ atoms. A total of 26 upper and lower distances limits was used in the CYANA calculations. This construction allows us to define a rigid cluster while leaving undefined the chirality of the peptide folding around it. The summary of conformationally restricting constraints is reported in **Table S5**.

Structure Calculation and refinement. Structure calculations were performed with the program CYANA 2.1 [17,18]. A total of 2000 random conformers were subjected to 65000 steps of a simulated annealing process. The 20 conformers with the lowest target function constituted the final family. Each member of the family was subsequently submitted to refinement in explicit solvent with the Amber-16 package [19]. The force field parameters for the 4Fe-4S cluster were taken as in similar systems [20]. A value of 50 kcal mol⁻¹ Å⁻² was used as force constant for the NOE and paramagnetic NMR restraints whereas a values of 32 kcal mol⁻¹rad⁻² was used for torsion angle restraints. The quality of the structure was evaluated in terms of deviations from ideal bond lengths and bond angles and through Ramachandran plots obtained using the programs with PSVS 1.5 program [21].

Table S5. Summary of conformationally-restricting constraints and structure quality factors of the AMBER energy minimized family of conformers of PioC

	PioC^a (20 Conformers)
Total number of meaningful NOE upper distance constraints^b :	344
Intra-residue	136
Inter-residue	
Sequential ($ i-j = 1$)	103
Medium-range ($ i-j < 4$)	56
Long-range ($ i-j > 5$)	49
Total meaningful dihedral angle restraints :	51
Phi	26
Psi	25
Total number of paramagnetic NMR Restraints:	189
Upper Distances constraints derived from $R_{1,2 \text{ para}}$	175
Cys βCH_2 dihedral angle constraints χ_2	4
1D NOEs between Cys βCH_2 and neighboring residues	7
H-bonds linking HN donor atoms to $S\gamma$ of the preceding cluster-bound Cys	3
Cluster	
Geometrical parameters defining the cluster (upl)	26
Geometrical parameters defining the cluster (lol)	26
Residual NOEs constraint violations^c :	
Distance violations / structure	
Between 0.1 -0.2 Å	3.2
Between 0.2 -0.5 Å	0.3
> 0.5 Å	0

RMS of Distances violations per meaningful distance constraint (Å):	0.03
Maximum distance violation ^d	0.33 Å
Residual PRE violations ^e	
Distance violations / structure	
Between 0.1 -0.2 Å	7.6
Between 0.2 -0.5 Å	2.0
> 0.5 Å	0
RMS of Distances violations per meaningful distance constraint (Å):	0.05
Dihedral angle violations / structure:	
1-10 °	3.3
> 10°	0
RMS violations per meaningful dihedral angle constraints (°):	0.82 °
Maximum dihedral angle violation ^d	6.90 °
Average RMSD to the mean (Å):	
Residue range 5-50 (backbone atoms)	0.62±0.11Å
Residue range 5-50 (all heavy atoms)	1.14±0.13Å
residual CYANA Target Function (Å ²)	1.92±0.13
Structure Quality Factors - overall statistics:	Z-score^g
Procheck G-factor e (phi / psi only) ^f	-3.30
Procheck G-factor e (all dihedral angles) ^f	-4.91
Verify3D	-4.01
ProsaII (-ve)	1.24
MolProbity clashscore	1.29
Ramachandran Plot Summary from Procheck ^f	
Most favoured regions	74.2%
Additionally allowed regions	20.9%
Generously allowed regions	3.40%
Disallowed regions	1.40%
Ramachandran Plot Statistics from Richardson's lab ^f	
Most favoured regions	84.6%

Allowed regions	12.4%
Disallowed regions	3.0%
^a The data are calculated over the 20 conformers representing the NMR structure. The mean value and the standard deviation are given ^b Number of meaningful constraints for each class. ^c Analyzed for residues 1 to 54, The analysis has been performed with PSVS 1.5 program considering the 344 meaningful NOEs ^d Largest distance or dihedral angle constraint violation among all the reported structures ^e Analyzed for residues 1 to 54, The analysis has been performed with PSVS 1.5 program considering the 175 Upper Distances constraints derived from R _{1,2} para ^f Selected residue ranges: 5-50 ^g With respect to mean and standard deviation for a set of 252 X-ray structures < 500 residues, of resolution ≤ 1.80 Å, R-factor ≤ 0.25 and R-free ≤ 0.28; a positive value indicates a 'better' score. Z-score generated using PSVS 1.5	

Comparison among different family of structures. Backbone and all heavy atoms RMSD obtained for each family of structure are summarized in **Table S6**. As discussed in text, it appears that: i) the highest precision is obtained when all available structural constraints are used, ii) the NOE-only and the PRE-only families of structures have very similar precision; iii) the contribution of the constraints arising from the cluster-bound residues, i.e. involving the first coordination sphere of the cluster, is extremely important. Indeed, the addition of 14 constraints (less than 5% of the total number of diamagnetic NOEs) provides a 20% improvement in both backbone and all atoms RMSD (see the comparison between the first two columns of **Table S6**).

Table S6

Constraints used in Structure Calculation	<i>NOE-only*</i>	<i>NOE-only**</i>	<i>full-set</i>	<i>PRE-only</i>
Backbone RMSD [§] (residues 5-50)	1.04 ±0.29 Å	1.27±0.19 Å	0.62±0.11Å	1.31±0.27Å
All heavy at. RMSD [§] (residues 5-50)	1.81±0.30 Å	1.95±0.22 Å	1.14±0.13Å	2.00±0.32Å

[§] Data are related to the family of structure obtained upon CYANA calculation followed by 30 ps (15 000 steps with a time step of 2.0 fs) of restrained molecular dynamics at constant temperature and constant pressure (1.0 bar) using AMBER 16;

*NOE-only** 344 meaningful NOEs + 26 geometrical parameters of the cluster + 14 constraints arising from cluster-bound residues (4 χ_2 dihedral angle constraints -7 NOEs + 3 Hbonds)

*NOE-only*** 344 meaningful NOEs + 26 geometrical parameters of the cluster without constraints arising from cluster bound residues.

full-set: 344 meaningful NOEs + 26 geometrical parameters of the cluster + 14 constraints arising from cluster-bound residues +178 PREs

PRE-only: 26 geometrical parameters of the cluster + 14 constraints arising from cluster-bound residues +178 PREs

6. ^1H R_1 and R_2 relaxation measurements and PRE constraints

Relaxation rates ^1H R_1 and ^1H R_2 were measured using a 11.7 T Bruker AVANCE 500 equipped with a triple resonance, inverse detection, cryoprobe (TXI) or a 14.0 T Bruker AVANCE NEO 600, equipped with a room temperature triple resonance inverse detection probe. For the ^1H R_1 rates, two series of experiments were used to measure longitudinal relaxation rates. For slow relaxing signals, a standard ^{15}N -HSQC was edited with a non selective ^1H inversion recovery building block. Fifteen experiments were collected (See Table S1), using a recycle delay of 4 s and inversion recovery delays of 20 ms, 30 ms, 40 ms, 50 ms, 60 ms, 80 ms, 100 ms, 120 ms, 160 ms, 200 ms, 300 ms, 400 ms, 600 ms, 800 ms, 1s. In order to measure relaxation rates of signals severely affected by the hyperfine relaxation, R_1 ^1H

^1H rates were measured also with an IR- ^{15}N -HSQC-AP experiment [22]. Fourteen experiments were collected using a recycle delay of 150 ms, an INEPT transfer delay (formally $1/(4J)$) of 710 μs and inversion recovery delays of 2.0 ms, 4.0 ms, 6.0 ms, 10 ms, 15 ms, 20 ms, 25 ms, 30 ms, 40 ms, 50 ms, 60 ms, 80 ms, 120 ms, 200 ms. In both series of experiments, the intensities of the ^{15}N HSQC spectra fitted according to three parameter fitting $I(t) = I(0) * [1 - 2\exp(-t * R_1)] + C$.

Transverse relaxation rate ^1H - R_2 , were also measured with two different approaches. For signals relatively far from the paramagnetic center, relaxation rates were measured using an experiment where a variable delay is inserted during the INEPT transfer of a ^{15}N HSQC experiment [23]. Fourteen experiments were recorded, using a 4s recycle delay, a 1200 μs selective ^1H inversion pulse for $^3\text{J}_{\text{H}_\text{N}\text{H}_\alpha}$ decoupling, relaxation delays of 8.0 ms, 12 ms, 16 ms, 28 ms, 40 ms, 52 ms, 64 ms, 76 ms, 88 ms, 112 ms, 136 ms, 160 ms, 200 ms and 240 ms. To measure ^1H R_2 rates of signals that are strongly affected by the hyperfine interaction, a new experiment, termed R_2 -weighted ^{15}N -HSQC-AP has been developed, as reported in **Figure S5**. ^1H - R_2 measurements were obtained from a series of sixteen R_2 -weighted ^{15}N -HSQC-AP experiments recorded using recycle delays of 150 ms and INEPT transfer periods of 0.1 ms, 0.2 ms, 0.3 ms, 0.4 ms, 0.6ms, 0.8 ms, 1.0 ms, 1.2 ms, 1.4 ms, 1.6 ms, 2.0 ms, 2.4 ms, 2.8 ms, 3.2 ms, 4.0 ms, 5.0 ms. All relaxation data were analyzed using the Bruker Topspin Dynamics Center.

The R_1 and R_2 rates obtained with the different methods were compared and, for each residue, the value with the lower standard deviation in the fitting was considered. As expected, slow relaxing ^1H resonances were better fitted using an in-phase ^{15}N HSQC as editing spectrum and experimental conditions typical of diamagnetic systems, i.e. a long recycle delay, and longer relaxation periods. Conversely, fast relaxing signals were better fitted using the IR-HSQC-AP and the R_2 weighted HSQC-AP sequences. We found that all signals having R_1 values faster than 30 s^{-1} (13 out of 48 total R_1 measurements) and R_2

values faster than 45 s^{-1} (18 out of 49 total R_2 measurements) were better fitted using the tailored sequences. The results are summarized in **Table S7**.

For ^1H signals observed in ^{13}C HSQC experiments, R_1 ^1H rates were measured by inserting a non-selective ^1H inversion recovery filter prior to a standard ^{13}C HSQC experiments and fitting the intensities of the ^{13}C HSQC spectra as described above. Seventeen experiments were collected using a 2s relaxation delay and inversion recovery delays of 1.0 ms, 5.0 ms, 10 ms, 20 ms, 30 ms, 40 ms, 60 ms, 70 ms, 90 ms, 120 ms, 200 ms, 300 ms, 400 ms, 600 ms, 800 ms, 1.2 s, 2.0 s. The results are also reported in **Table S7**. ^1H relaxation rates of Cysteines $\text{H}\beta$ and $\text{H}\alpha$ protons were measured from an IR- ^{13}C HSQC-AP experiment, shown in **Figure S6**. The experiments were performed using acquisition and recycle delays of 17 ms and 65 ms, respectively. An INEPT transfer delay of 600 μs was used throughout the series. Ten experiments were performed using inversion recovery delays of 500 μs , 2.0 ms, 4.0 ms, 6.0 ms, 8.0 ms, 10 ms, 15 ms, 30 ms, 50 ms, 80 ms.

Conversion of Relaxation rates into distance constraints. Each set of relaxation rates, i.e. $R_1\text{ H}_\text{N}$, $R_2\text{ H}_\text{N}$, $R_1\text{ H}_\alpha$, R_1 ^1H side chains, R_1 ^1H Cysteine resonances, were analyzed independently. For each set of data, the diamagnetic contribution to the observed rate was estimated by taking the average value of those residues that are not affected by the hyperfine interaction. For this analysis were taken into consideration only those residues that, according to the internal dynamics measured with ^{15}N relaxation shown **Figure S7**, do not exhibit local internal motions. The $R_{1,2\text{ para}}$ contribution is then calculated according to $R_{1\text{obs}} = R_{1\text{dia}} + R_{1\text{para}}$.

Then, each set of $R_{1,2\text{ para}}$ values is converted into a distance (d) according to an equation of the form $d = (A/R_{1,2})^{1/6}$, where the constant A was calibrated empirically. The calculation of the A factor from the Solomon equation describing the dipolar coupling between electron spin and nuclear spin (ref, Bertini libro) can not be reliably performed for a number of reasons: i) the electronic correlation time describing the electron relaxation of the $\text{Fe}^{+2.5}$ ions is dependent on the magnetic coupling operative within the cluster; therefore the effective τ_e is expected to be significantly different from what observed in isolated, high spin Fe(III) and Fe(II) ions. Its value can only be estimated based on R_1 and R_2 values of atoms at fixed distance from the cluster and from previously studied HiPIPs. ii) according to the metal centered approximation, which is mandatory for the Solomon equation (ref x), the unpaired electron spin density should be considered as fully localized onto the metal ions. In the case of a cluster, this is clearly an oversimplification of the problem, because the electron spin is also partly delocalized among all cluster

atoms. Additionally, it is also known that the unpaired spin density is partly delocalized also onto the ligands, via both spin delocalization and spin polarization mechanisms. iii) paramagnetism arises from the population of the excited states of the electron spin energy ladder, which in turn depends from the antiferromagnetic coupling among the iron ions of the cluster (ref), which cannot be safely predicted. Provided all the above consideration, the use of an empirical calibration coefficient is recommended.

The conversion from $R_{1,2}$ values to distance has been previously done made according to an equation of the form:

$$R_{1,2\text{para}} = A/(\sum d_i^6) \quad (1) \text{ where } d_i \text{ are the distances from the nuclear spin to the four iron ions.}$$

Eq(1) accounts for atoms being at similar distances from two different Fe ions; however eq(1) still assumes that the electron spin density is fully localized onto the Iron ions. Therefore, to properly convert $R_{1,2\text{para}}$ values into upl we prefer to replace the $\sum d_i^6$ term with the power sixth distance from the center of mass of the cluster. In order to perform this, a special linker made of 100 pseudo-residues called LL2 was added at the end of the protein sequence. The "atoms" of these linker "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 linker is an ION residue (cyana library) which has been subsequently linked at fixed distances with the four Iron 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. For the AMBER refinement, the linker and the ION residue have been removed and, for each PRE constraint, the center of mass of the cubane has been replaced with the closest iron ion of the cluster (or with the two closest iron ions when ambiguous metal-to-proton distances occur), and the upper limit distance reduced accordingly.

Table S7					
n protein		HN R₂	std	HN R₁	std
VAL	1				
THR	2	11,1	0,4	12,9	0,5
LYS	3	15,4	1,3	15,9	0,8
LYS	4	19,1	1,5	3,83	0,1
ALA	5	23,4	0,8	5,41	0,17
SER	6	17,8	0,6	2,21	0,08
HIS	7	21,3	0,8	4,90	0,08
LYS	8	32,8	2,2	14,8	0,5
ASP	9	16,9	0,7	2,74	0,13
ALA	10	19,6	0,8	7,28	0,25
GLY	11	20,0	1,4	4,77	0,12

TYR	12	29,9	1,5	11,1	0,5	
GLN	13	53,2	4,1	16,6	0,6	
GLU	14	19,4	1,1	3,18	0,11	
SER	15	21,6	1,0	3,19	0,18	
PRO	16					
ASN	17	22,8	1,1	4,69	0,19	
GLY	18	25,1	1,5	7,50	0,36	
ALA	19	44,4	2,8	22,5	1,3	
LYS	20	17,8	0,8	2,79	0,19	
ARG	21	44,4	3,4	12,0	0,6	
CYS	22	311	18	116	32	
GLY	23	60,4	6,8	19,0	0,6	
THR	24	42,0	2,8	14,6	0,3	
CYS	25	60,1	2,8	37,6	3,8	
ARG	26	43,6	3,6	32,0	8,5	
GLN	27	246	18	160	13	
PHE	28	146	9	89,1	5,6	
ARG	29	49,3	3,0	25,0	1,0	
PRO	30					
PRO	31					
SER	32	25,1	1,5	7,50	0,36	
SER	33	24,2	0,8	5,75	0,27	
CYS	34	42,5	4,0	17,6	0,5	
ILE	35	81,4	5,7	48,0	5,1	
THR	36	107	9	100	7	
VAL	37	156	16	129	12	
GLU	38	24,2	0,8	8,71	0,17	
SER	39	19,7	1,0	3,89	0,10	
PRO	40					
ILE	41	27,7	1,3	5,61	0,12	
SER	42	42,8	2,4	13,9	0,4	
GLU	43	19,2	0,5	5,18	0,06	
ASN	44	27,4	1,5	5,62	0,11	
GLY	45	59,3	4,3	24,7	1,2	
TRP	46	69,5	5,2	42,7	2,3	
CYS	47	194	10	134	29	
ARG	48	53,5	4,2	35,4	2,1	
LEU	49	167	15	129	7	
			7,8			
TYR	50	47,7		39,8	10,5	
ALA	51	240	36			
GLY	52	23,0	1,3	6,38	0,19	
LYS	53	18,0	0,9	6,42	0,18	
ALA	54	7,7	0,3	2,15	0,07	
Trp sc	46	33,9	2,1			
		diamag- netic ex- periments		para- magnetic experi- ments		

		R ₁ H α	R ₁ H β	R ₁ H γ	R ₁ H δ	R ₁ H ϵ	R other
VAL	1						
THR	2	2,50	2,66	2,73			
LYS	3	2,34	3,60	2,80	2,56	2,14	
LYS	4	3,79	3,93/4, 41	4,40/6,36		3,94/4, 25	
ALA	5	2,93	7,20				
SER	6	2,92	2,23/2, 47				
HIS	7	10,2	4,06/4, 13				
LYS	8	2,39	3,16/3, 36	3,17/3,15	2,77/2, 69	2,44	
ASP	9	1,96	2,72/2, 80				
ALA	10	8,16	5,58				
GLY	11	5,01/3, 09					
TYR	12	16,7	24,6/15, 8		56,7		
GLN	13	3,22	2,78/3, 41	5,03/5,09			
GLU	14	4,79	3,00/2, 79	4,02			
SER	15	2,30	2,94/2, 96				
PRO	16	4,99	4,04/3, 61	3,10/6,02	2,46/2, 58		
ASN	17	2,24	4,48/4, 23				
GLY	18	1,37/2, 14					
ALA	19	1,55	2,31				
LYS	20	5,78	6,90/8, 35	4,60	3,67/4, 27	2,35/2, 32	
ARG	21	29,8	8,09/3, 03	5,26/7,89	4,08/4, 01		
CYS	22	53,6	368/13 0				¹³ C β 67
GLY	23	7,81/6, 15					
THR	24		4,47	3,77			
CYS	25	59,2	313/28 2				¹³ C β 59 ¹³ C α 13
ARG	26	10,2	8,02/4, 87	8,65/6,67	3,65/4, 54		
GLN	27	29,6	1,61	1,76/2,03			
PHE	28	2,34	45,1		44,5		

ARG	29	4,94	5,32/6, 11	5,96/5,40	4,22/4, 39		
PRO	30	3,36	2,43/2, 50	2,31/2,48	3,28/4, 38		
PRO	31	2,15	1,64/1, 97	4,79/2,02	1,77/1, 42		
SER	32	4,67	3,45/3, 44				
SER	33		5,89/5, 72				
CYS	34	69,3	164/58 8				¹³ Cβ 65
ILE	35	8,32	10,9	6,84/10,0/7 ,56	4,95		
THR	36	3,00	2,49	2,98			
VAL	37	21,6		59,9			
GLU	38	4,41	4,50/4, 53	3,66/3,58			
SER	39	6,34	3,08/2, 69				
PRO	40	3,62	2,12/2, 79	4,69/4,24	1,88/1, 87		
ILE	41	4,42	34,1	31,9/37,5/9 5,0	2,71		
SER	42	5,17	5,72/8, 04				
GLU	43	16,2	4,91/4, 19	4,03/3,26			
ASN	44	5,53	3,35/3, 62				
GLY	45	14,9/1 0,9					
TRP	46		16,7/22 ,7		36,4	84,3	
CYS	47		130/45 0				¹³ Cβ 67
ARG	48	13,8	4,23	7,25/6,63	3,10/4, 39		
LEU	49	12,7		4,44	5,37		
TYR	50	2,70			97,0		
ALA	51	10,7	13,1				
GLY	52	4,12/7, 56					
LYS	53	3,64	4,31	4,75	3,55		
ALA	54	2,56	2,08				
		diamag- netic exp.		paramag- netic exp.			

REFERENCES

1. Bird LJ, Saraiva IH, Park S, Calcada EO, Salgueiro CA, Nitschke W, et al. Nonredundant roles for cytochrome c2 and two high-potential iron-sulfur proteins in the photoferrotroph *Rhodospseudomonas palustris* TIE-1. *J Bacteriol.* 2014;196(4):850-8. doi: 10.1128/jb.00843-13
2. Bertini I, Jiménez B, Piccioli M. ^{13}C direct detected experiments: optimisation to paramagnetic signals. *J Magn Reson.* 2005;174(1):125-32. doi:
3. Balayssac S, Jiménez B, Piccioli M. Assignment Strategy for Fast Relaxing Signals: Complete Aminoacid Identification in Thulium Substituted Calbindin D_{9k}. *J Biomol NMR.* 2006;34(2):63-73. doi:
4. Ciofi-Baffoni S, Gallo A, Muzzioli R, Piccioli M. The IR- ^{15}N -HSQC-AP experiment: a new tool for NMR spectroscopy of paramagnetic molecules. *J Biomol NMR.* 2014;58:123-8. doi:
5. Bertini I, Capozzi F, Luchinat C, Piccioli M, Vicens Oliver M. NMR is a unique and necessary step in the investigation of iron- sulfur proteins: the HiPIP from *R. gelatinosus* as an example. *Inorg Chim Acta.* 1992;198-200:483-91. doi:
6. Bertini I, Jiménez B, Piccioli M, Poggi L. Asymmetry in ^{13}C - ^{13}C COSY spectra identifies geometry in paramagnetic proteins. *J Am Chem Soc.* 2005;127(35):12216-7. doi:
7. Hirano Y, Takeda K, Miki K. Charge-density analysis of an iron-sulfur protein at an ultra-high resolution of 0.48 Å. *Nature.* 2016;534(7606):281-4. doi: 10.1038/nature18001
8. Ohno H, Takeda K, Niwa S, Tsujinaka T, Hanazono Y, Hirano Y, et al. Crystallographic characterization of the high-potential iron-sulfur protein in the oxidized state at 0.8 Å resolution. *PLoS One.* 2017;12(5):e0178183. doi: 10.1371/journal.pone.0178183
9. Stelter M, Melo AM, Hreggvidsson S, Saraiva LM, Teixeira M, Archer M. Structure at 1.0 Å resolution of a high-potential iron-sulfur protein involved in the aerobic respiratory chain of *Rhodothermus marinus*. *J Biol Inorg Chem.* 2010;15(3):303-13. doi: 10.1007/s00775-009-0603-8
10. Banci L, Bertini I, Eltis LD, Felli IC, Kastrau DHW, Luchinat C, et al. The three dimensional structure in solution of the paramagnetic protein high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* through nuclear magnetic resonance. *Eur J Biochem.* 1994;225(2):715-25. doi:
11. Rayment I, Wesenberg G, Meyer TE, Cusanovich MA, Holden HM. Three-dimensional structure of the high-potential iron-sulfur protein isolated from the purple phototrophic bacterium *Rhodocyclus tenuis* determined and refined at 1.5 Å resolution. *J Mol Biol.* 1992;228(2):672-86. doi: 10.1016/0022-2836(92)90849-f
12. Bertini I, Capozzi F, Luchinat C, Piccioli M, Vila AJ. The Fe₄S₄ centers in ferredoxins studied through proton and carbon hyperfine coupling. Sequence specific assignments of cysteines in ferredoxins from *Clostridium acidi urici* and *Clostridium pasteurianum*. *J Am Chem Soc.* 1994;116:651-60. doi:
13. Banci L, Camponeschi F, Ciofi-Baffoni S, Piccioli M. The NMR contribution to protein-protein networking in Fe-S protein maturation (vol 23, pg 665, 2018). *Journal of Biological Inorganic Chemistry.* 2018;23(4):687-. doi: 10.1007/s00775-018-1573-5
14. Banci L, Bertini I, Luchinat C, Piccioli M, Scozzafava A, Turano P. ^1H NOE studies on dicopper(II) dicobalt(II) superoxide dismutase. *Inorg Chem.* 1989;28:4650-6. doi:
15. Bertini I, Donaire A, Feinberg BA, Luchinat C, Piccioli M, Yuan H. Solution structure of the oxidized 2[Fe₄S₄] ferredoxin from *Clostridium pasteurianum*. *Eur J Biochem.* 1995;232:192-205. doi:
16. Bertini I, Eltis LD, Felli IC, Kastrau DHW, Luchinat C, Piccioli M. The solution structure of oxidized HiPIP I from *Ectothiorhodospira halophila*, can NMR probe rearrangements associated to electron transfer processes? *Chemistry - A European Journal.* 1995;1:598-607. doi:
17. Güntert P. Automated NMR structure calculation with CYANA. *Methods Mol Biol.* 2004;278:353-78. doi:
18. Güntert P, Buchner L. Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR.* 2015;62(4):453-71. doi: 10.1007/s10858-015-9924-9
19. D.A. Case IYB-S, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. AMBER 2018. 2018. doi:

20. Carvalho AT, Swart M. Electronic structure investigation and parametrization of biologically relevant iron-sulfur clusters. *Journal of chemical information and modeling*. 2014;54(2):613-20. doi: 10.1021/ci400718m
 21. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. *Proteins: Structure, Function, and Bioinformatics*. 2007;66(4):778-95. doi:
 22. Ciofi-Baffoni S, Gallo A, Muzzioli R, Piccoli M. The IR-N-15-HSQC-AP experiment: a new tool for NMR spectroscopy of paramagnetic molecules. *Journal of Biomolecular Nmr*. 2014;58(2):123-8. doi: 10.1007/s10858-013-9810-2
 23. Donaldson LW, Skrynnikov NR, Choy W-Y, Muhandiram DR, Sarkar B, Forman-Kay JD, et al. Structural Characterization of Proteins with an Attached ATCUN Motif by Paramagnetic Relaxation Enhancement NMR Spectroscopy. *J Am Chem Soc*. 2001;123:9843-7. doi:
-

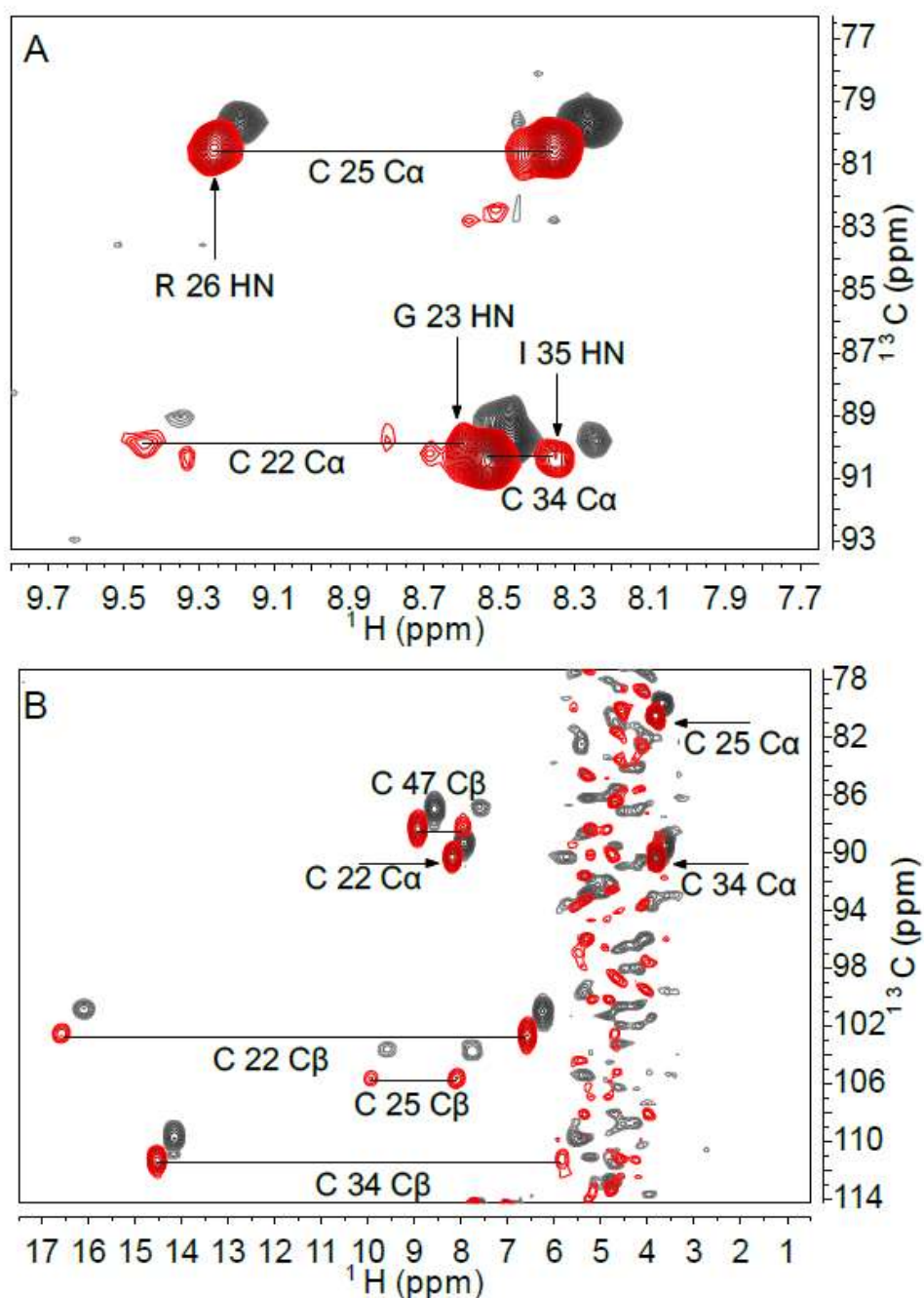


Figure S1. 500 MHz, H(N)CA (A) and ^{13}C HSQC (B) experiments at 288K (black) and 298K (red) optimized for peaks involving fast relaxing resonances. Peculiar ^{13}C shift values allows the sequential assignment of three out of the four cluster-bound Cysteines in the HNCA, while the CH_2 pairs are unambiguously identified via ^{13}C HSQC.

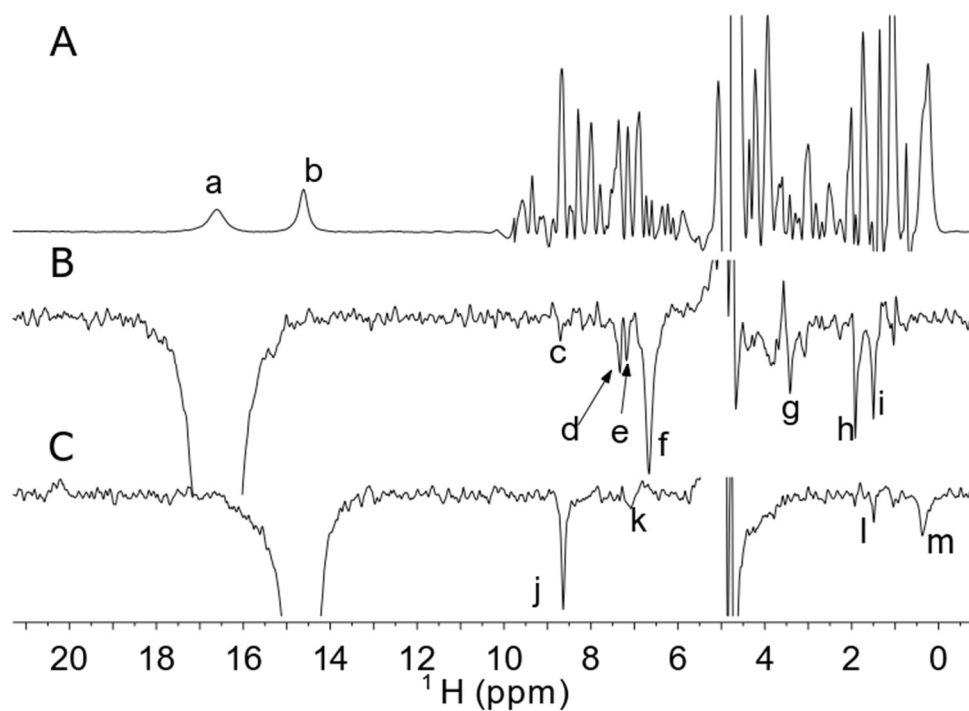


Figure S2. A 1D ^1H NMR spectrum of PioC, optimized to observe hyperfine shifted and fast relaxing resonances. Two isolated peaks (labeled a-b) are observed. B-C 1D NOE difference spectra obtained upon selective saturation of signals a-b. The NOE peaks observed in the difference spectra, labeled c-m, are used to perform the assignment of signals a and b. Experiment were performed at 400 MHz and 298K. About 400.000 scans were acquired for each NOE, using the inversion recovery sequence experiments and collecting the difference between on and off resonance selective irradiation

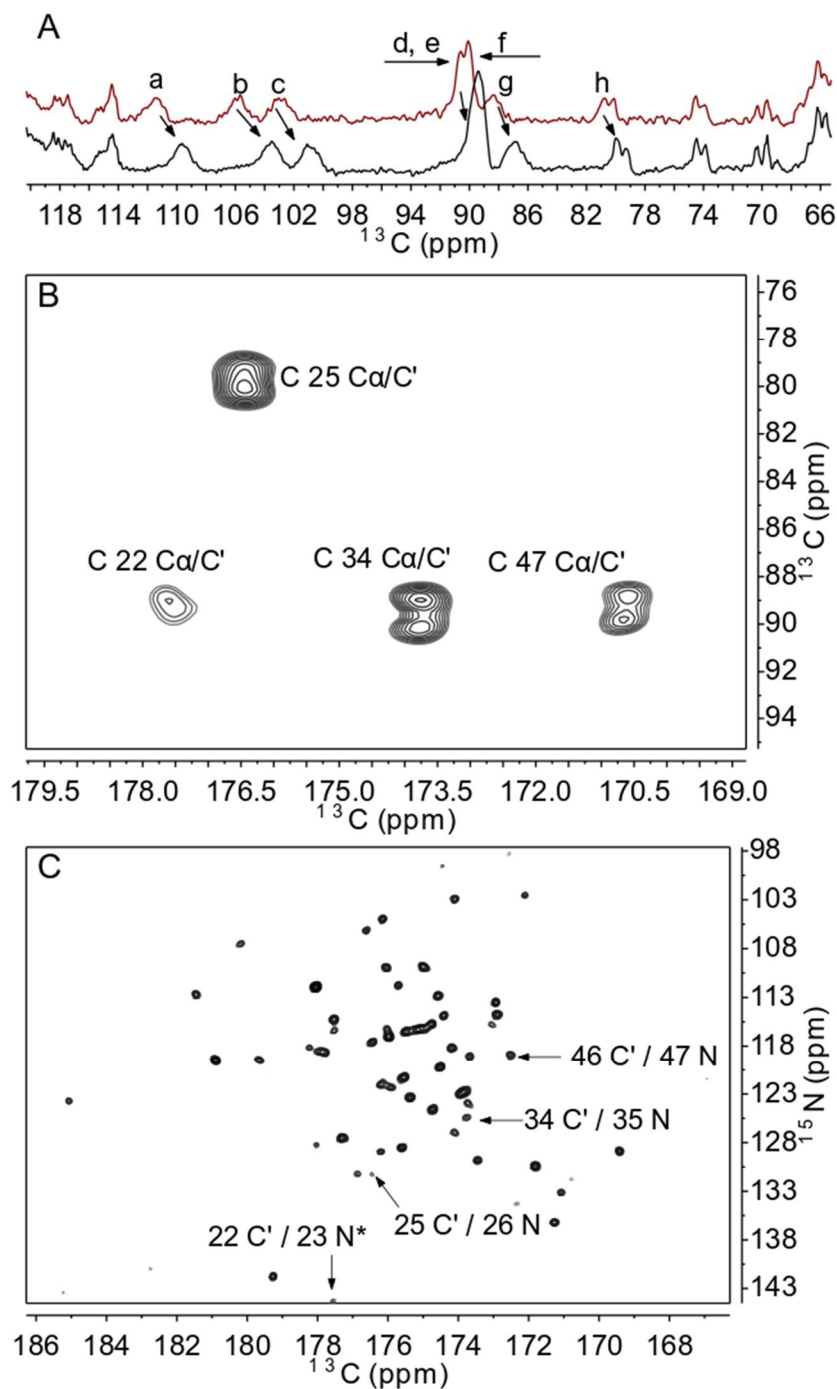


Figure S3. A 1D ^{13}C NMR spectra of PioC, optimized to observe fast relaxing resonances. Figure shows the spectral region where we expect to observe $^{13}\text{C}\alpha$ and $^{13}\text{C}'$ of cluster-bound residues. Spectrum in black is recorded at 288 K, in red at 298K. B ^{13}C - ^{13}C COSY spectrum (upper diagonal part) showing the connectivities between C' and $\text{C}\alpha$ signals. Acquisition and processing parameters are optimized to identify connectivities among fast relaxing resonances. In both dimensions no ^1H decoupling has been used. C ^{13}C - ^{15}N CON experiment. Signals involving C' spins from cluster-bound Cysteines are labeled in figure. All the above experiments were performed at 175 MHz, using ^{13}C direct detection

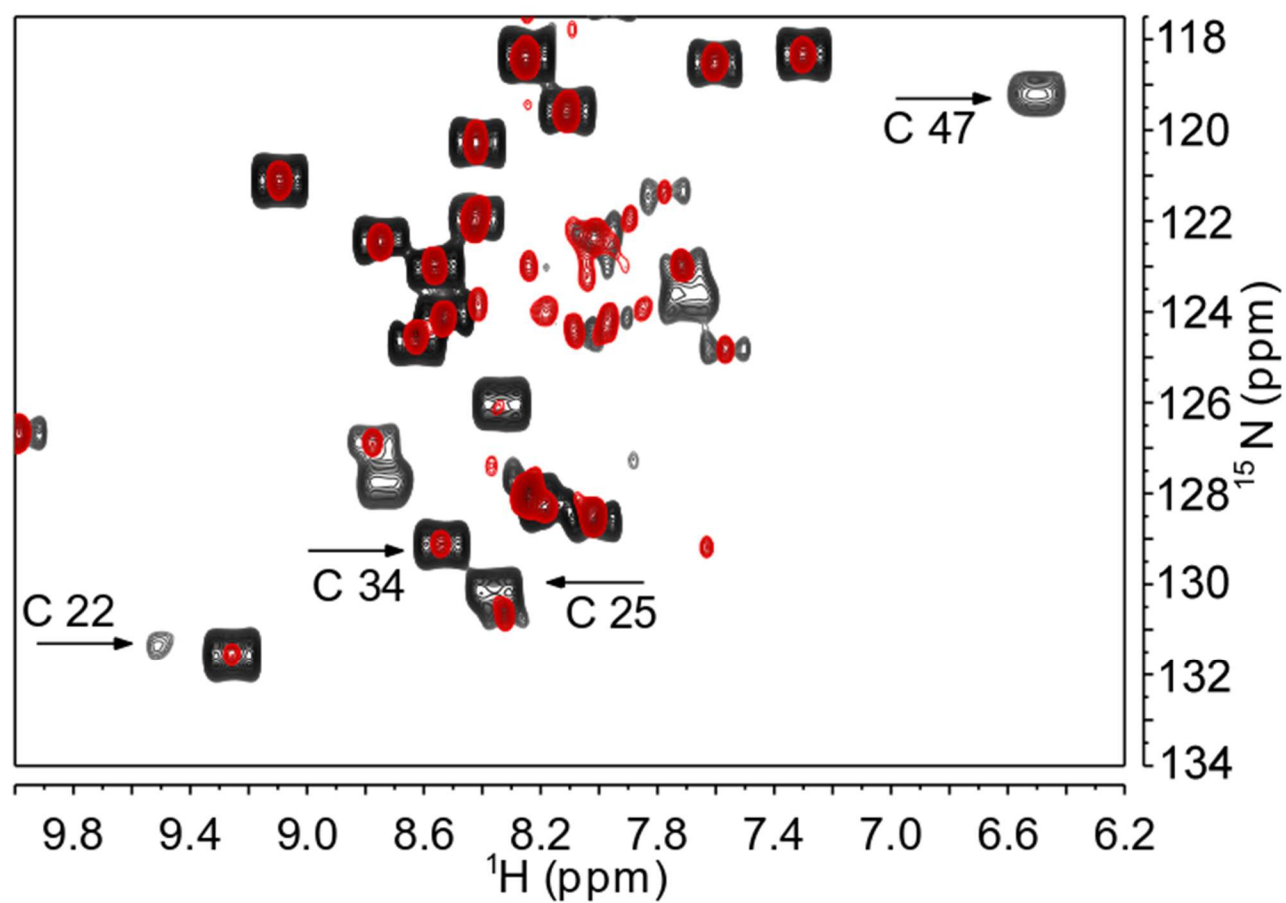


Figure S4. Expanded region of a ^{15}N HSQC-AP experiment where the four cluster-bound cysteines are labeled. Spectrum has been recorded at 500 MHz, 298K, and a standard in-phase HSQC experiment is overlaid in red for comparison purposes.

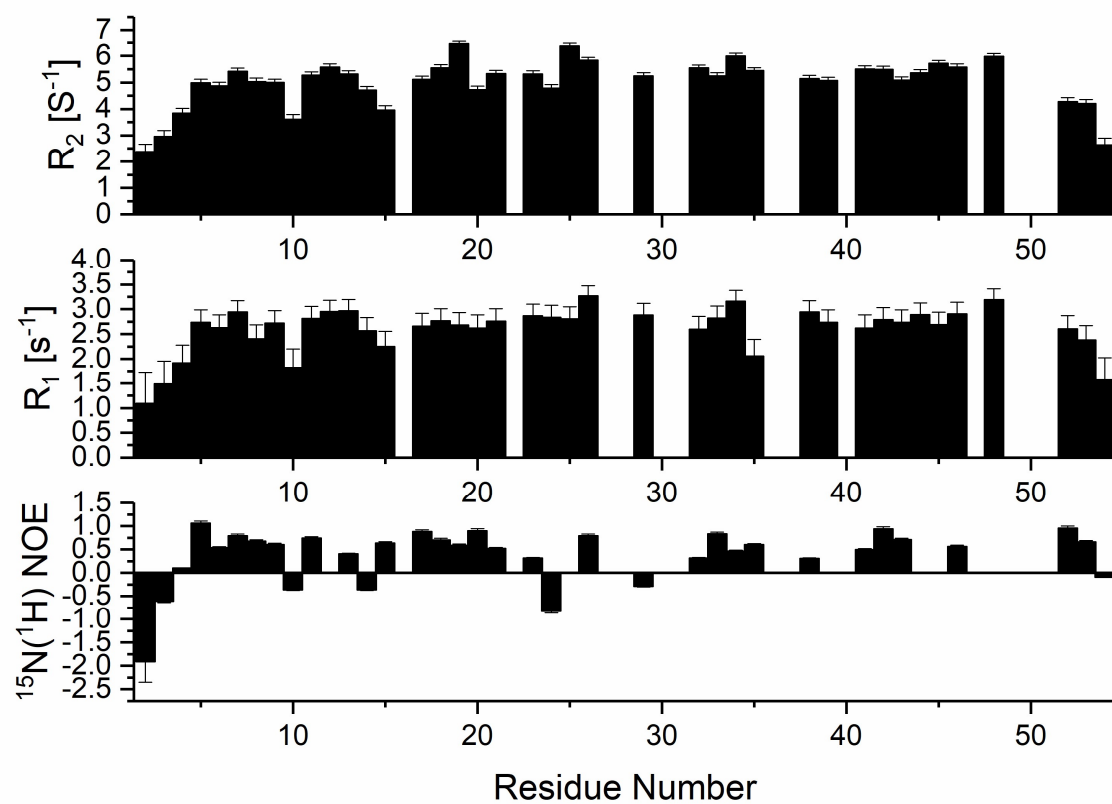
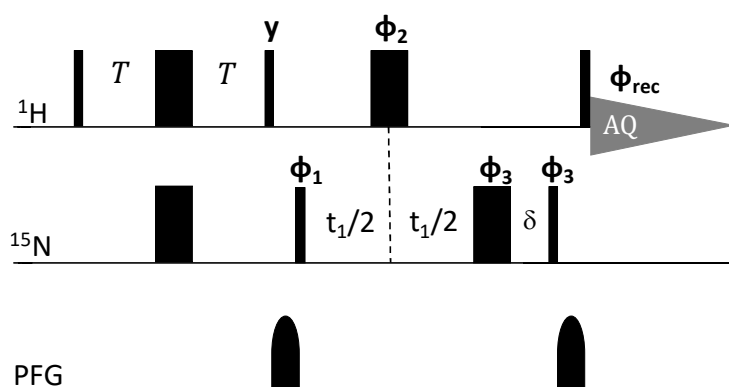


Figure S5. Experimental ^{15}N R_1 , R_2 rates and heteronuclear NOEs as obtained from the ^{15}N relaxation experiments performed at 500 MHz, 298K on a ^{15}N labelled PioC sample



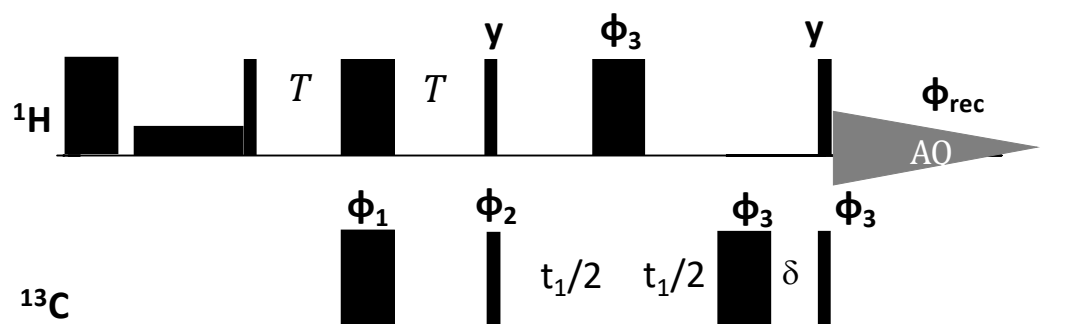
$$\phi_1 = x, -x$$

$$\phi_2 = 2(x), 2(-x)$$

$$\phi_3 = 4(x), 4(-x)$$

$$\phi_{rec} = x, -x, x, -x, -x, x, -x, x$$

Figure S6. Pulse sequence used for the measurement of $^1\text{H}_\text{N}$ R_2 relaxation rates. The experiment is an ^{15}N -HSQC experiments with detection in antiphase. For each HN peak, the intensity of the doublet as a function of T provides the values of ^1H transverse relaxation rate



$$\phi_1 = 8(x), 8(-x)$$

$$\phi_2 = x, -x$$

$$\phi_3 = 2(x), 2(-x)$$

$$\phi_{rec} = x, -x, x, -x, -x, x, -x, x$$

Figure S7. Pulse sequence used for the measurement of $^1\text{H}_\text{C}$ R_1 relaxation rates. The experiment is a ^{13}C -IR-HSQC experiment with detection in antiphase. For each HC peak, the intensity of the doublet as a function of the inversion recovery delay \square provides the values of ^1H longitudinal relaxation rates