Prediction of the interaction of metallic moieties with proteins: an update for protein-ligand docking techniques

Giuseppe Sciortino,[§,#] Jaime Rodríguez-Guerra,[§] Agustí Lledós,[§] Eugenio Garribba,[#] and Jean-Didier Maréchal.*[,§]

[§] Departament de Química, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallés, Barcelona, Spain

[#] Dipartimento di Chimica e Farmacia, Università di Sassari, Via Vienna 2, I-07100 Sassari, Italy

Corresponding author. E-mail: jeandidier.marechal@uab.cat (J-D.M.).

ABSTRACT

In this paper, we present a new approach to expand the range of application of protein-ligand docking methods in the prediction of the interaction of coordination complexes (i.e. metallodrugs, natural and artificial cofactors, etc.) with proteins. To do so, we assume that, from a pure computational point of view, hydrogen bond functions could be an adequate model for the coordination bonds since both share directionality and polarity aspects. In this model, docking of metalloligands can be performed without using any geometrical constraint or energy restraint. The hard work consists in generating the convenient atom types and scoring functions. To test this approach, we applied our model to 39 high-quality X-ray structures with transition and main group metal complexes bound via a unique coordination bond to a protein. This concept was implemented in the protein-ligand docking program GOLD. The results are in very good agreement with the experimental structures: the percentage in which the RMSD of the simulated pose is smaller than the X-ray spectra resolution is 92.3% and the mean value of RMSD is < 1.0. Such results also show the viability of the method to predict metal complexes–proteins interactions when the X-ray structure is not available. This work could be the first step for novel applicability of docking techniques in medicinal and bioinorganic chemistry and appears generalizable enough to be implemented in most protein-ligand docking programs nowadays available.

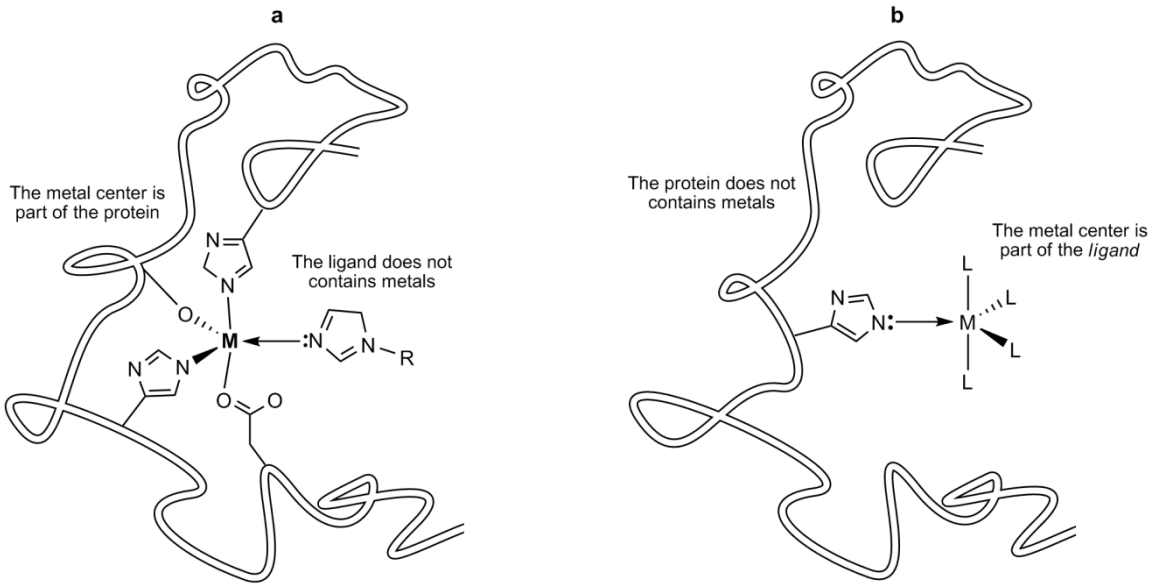
Introduction

Since the appearance of *cis*-platinum in the pharmaceutical landscape in the 1970's,[1] the interest in designing drugs based on coordination complexes has drastically increased. Over the last years, metal salts and/or metal complexes have been proposed in the therapy and diagnosis of many diseases and a large series of textbooks and reviews have been published on medicinal inorganic chemistry.[2-9] Gold, iridium and rhodium complexes are now only few examples of central atoms in drug design.[10] Novel breakthroughs in the development of novel metallodrugs are surely awaiting.

To highlight the specificity of metallodrugs in the drug design landscape, at least three aspects related to their molecular properties need to be highlighted. First, transition metals provide geometric and isomeric complexity absent in organic systems, hence allowing regio- and enantiospecific molecular interactions. Second, the binding of the metallodrugs to biological targets can involve different degrees of chemical changes in the first coordination sphere of the metal, ranging from no exchange at all to multiple ligand exchanges (here the word ligand has the usual meaning in the coordination chemistry and it refers to the chemical species that binds the metal throughout a direct coordination bond). The third point to be mentioned is the difficulty to reach experimental resolution of structures of metallodrugs with X-ray (i.e. lability of the drug-protein interaction under crystal conditions and electron beaming) and NMR (i.e. open shell systems) approaches. Therefore, despite their potential, metallodrugs still represent a narrow field of research when compared with the amazing amount of projects devoted by academia and companies to identify drug candidates based on organic species.

Computation has become a major asset in drug design projects. Either based on combinatorial, pseudo-rational or rational approaches, the use of molecular modelling is now a fundamental tool in the drug design pipeline. Protein-*ligand* dockings are generally the approach used at its early stage so to reach fast and accurate enough predictions of binding energies and geometries of drug-receptor complexes (here the italic writing of *ligand* responds to the docking terminology and corresponds to any molecular species that interact with a protein). Unfortunately, valid predictions

of the binding involving metal-containing *ligands* or changes in the chemical state of the *ligand* are still an open battle. In the case of metallodrugs, this represents the nexus of a major computational challenge. Indeed, how to deal with the formation of coordination bond through the metal and a donor of an amino acid side chain is a very diffuse situation.[11-16] When dealing with a metalloprotein and the possibility to interact with a coordination bond between the metal and an organic *ligand* (Scheme 1a), most of the programs nowadays available offer some solutions.[17-19] However, currently, systems in which a metal-containing *ligand* binds covalently to proteins (Scheme 1b) can only be simulated through *covalent docking* approaches which many programs have now implemented, such as GOLD[19] or Autodock,[20] or source code modification of commercial software, for example CovalentDock[21] or Docktite.[22]



Scheme 1. The two possibilities for the interaction between a metal center and a *ligand*: a) a coordination bond between the metal centre of a metalloprotein and an organic *ligand;* b) coordination bond between a metal-containing *ligand* and a protein side chain. His was used as representative amino acid and can replaced by any other coordinating side chain of Asp, Glu, Tyr, Ser, Thr, Cys, etc.

All these approaches have a critical limitation: the user needs to define *a priori* the specific atom pair involved in the protein-*ligand* bond and force the docking with energy restraints and/or geometrical constraints. Therefore, the applicability of these methods is limited to the systems in which the specific residue involved in the bond is already known, which significantly limits the application of docking as a predictive method. A further limitation concerns the applicability to metal-containing *ligands* for which parametrization of coordination bonds is totally absent in the scoring functions. As such, no computational docking tool exists to reproduce accurately the structures of metallodrugs bound to proteins without using strong geometrical constraints or energy restraints.

In this paper we present an extensive study to include coordination scoring parameters into the docking program GOLD for 39 transition and main group metal-containing *ligands*[23-51] and taking advantage of the similarity, at least from the computational point of view, between coordination bonds and standard polar interactions (i.e. hydrogen bonds or *hbonds*) available in protein-*ligand* docking software. The results suggest that docking methods could represent a new generalizable tool to predict metal complexes–proteins interactions and could have a general applicability not only in medicinal chemistry but also in the entire field of bioinorganic chemistry.

Materials and methods

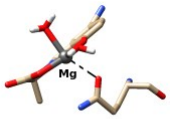
Dataset. To assess our model, first we collected a dataset of 39 high quality X-ray structures in which the protein interacts with a metal complex through a single donor forming one coordination bond. All structures (reported in Table S1 of Supporting Information and represented in Figure 1) were obtained from the Protein Data Bank (PDB) and include a wide range of metals, coordination geometries and donor types. The cross-validation of docking results on the dataset has allowed us to

develop adequate force fields for each metal and a general method able to predict the binding site of a various kinds of metallo-compounds. The distribution of different coordination geometries and metals present in the dataset is summarized in Table 1. The selected X-ray structures[23-51] and the donor set of the first coordination sphere are reported and explicitly described in Table S1.

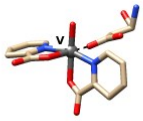
Table 1. Metals and geometry distribution in the dataset.[a]

	lin	td	sp	tbp	spy	oct	Tot
Mg						1	1
V				3		2	5
Cr						1	1
Mn						2	2
Fe					3		3
Co					2	1	3
Ni					1		1
Cu			3		2		5
Zn					2		2
Ru		2				3	5
Rh						1	1
Re						2	2
Os						1	1
Pt			6				6
Au	1						1
Total	1	2	9	3	10	14	39

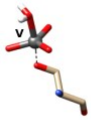
[a] lin = linear; td = tetrahedral; sp = square planar; tbp = trigonal bipyramidal; spy = square pyramidal; oct = octahedral.



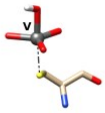
1g3b



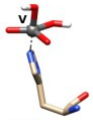
4c3w



1b8i

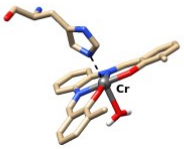


1rpt

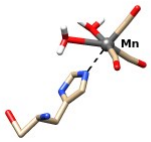


3i80

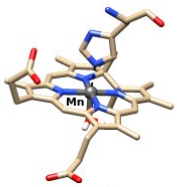
2p8o



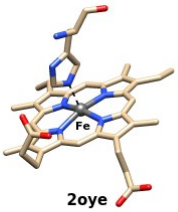
1j3f



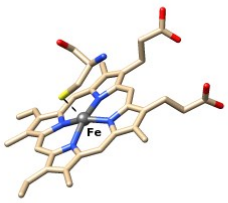
2q0m



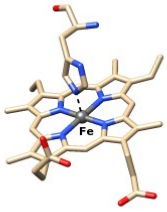
3wi8



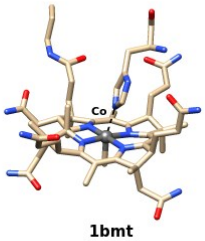
2oye



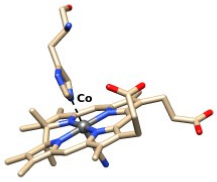
3wrh



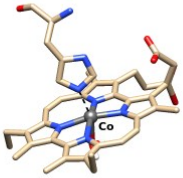
5d5r



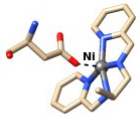
1bmt



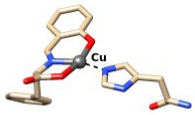
5azq



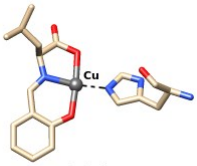
5azr



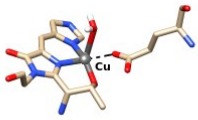
4kxi



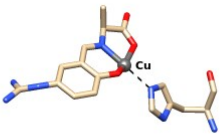
2eb8



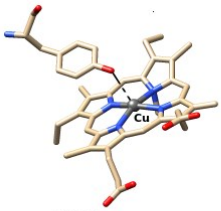
2eb9



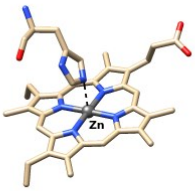
1kyr



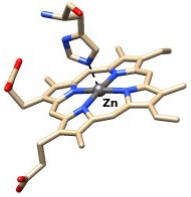
3aas



1c9e

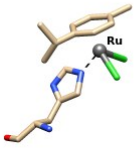


4mec

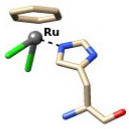


4mxl

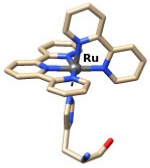
1jzf



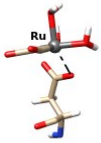
1t3p



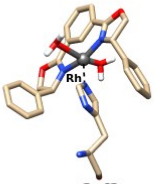
3w6a



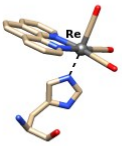
1jzh



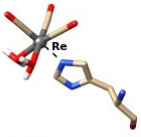
2xjw



2ef2

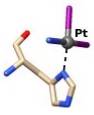


1jzi

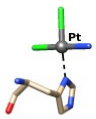


3kam

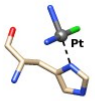
1jzj



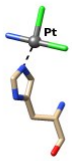
4mr1



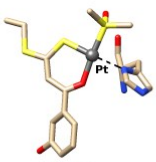
4n10



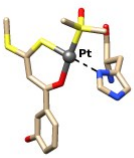
4s1y



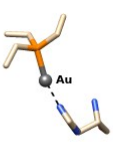
4yeo



4z3m



4z41



1e3b

Figure 1. Graphical representation of the dataset. Coordination complex formed by the 39 metallo-containing *ligands*.

Method. All crystallographic water and small molecules were removed and hydrogen atoms were added with UCSF Chimera Software.[52] The metal complexes co-crystallized with the proteins were removed from the active site and the coordinates saved in a new MOL2 file. The hydrogen atoms of the metal complexes and proteins were added with UCSF Chimera software.

Docking protocol. Docking calculations were performed with the program GOLD. As a first attempt to test our hypothesis, all calculations were performed with the GoldScore scoring function. This choice comes from its high flexibility for parameter modification as well as its robustness for posing prediction. All modifications to the GOLD force-field were made by adapting the force-field table and parameter files, without any source code modifications.

The GoldScore parameter file was modified to include parameters of atom types not included in GOLD's database, such as metals and possible coordinating amino acid groups and, in particular, sp^2 oxygens (e.g. those of carboxylate group), sp^3 oxygen of water and sp^3 negatively charged oxygens of deprotonated serine, threonine and tyrosine residues. The metal atom types (M.Metal) were built with the values reported in the literature, the keto (O.pl3) and the water oxygen atoms (O.H2O) and the sp^3 negatively charged oxygens were fixed considering the geometry of the electron pairs derived from the VSEPR theory.

Genetic algorithm (GA) parameters were set with a number of GA runs equal to 50 and a minimum of 100,000 operations. The rest of parameters, including pressure, number of islands, niche size, crossover, mutation and migration were left to default.

The metal complexes, separated from the protein, were blindly re-docked to reproduce the crystallographic structure; to prevent any bias, the coordinates of the resulting MOL2 file of the metallo*ligand* were randomized (with both rotational and translational transformations). In all the calculations, the metal-containing *ligands* were treated as a flexible structure with the algorithm

implemented in GOLD, the protein residues were considered rigid as the crystallographic structure is already in the ideal conformation to bind the metal complex; this is considered enough for the first series of tests on the validity of this approach. The radius of the evaluation sphere was set to 20 Å and centered on the binding site.

The docking calculation was carried out without any geometrical constraints or energy restrains. The efficacy of the method can be evaluated on the site discrimination capability (see the movie BlindDocking_2eb9.avi in the Supporting Information to illustrate the approach). The docking solutions were analyzed by means of GaudiView, an *in-house* GUI (graphical user interface) tool built as an extension for UCSF Chimera.[53]

Coordination docking implementation. Our model consists of associating hydrogen bond interactions with a coordination bond. In fact, as shown in Figure 2, the two kinds of interactions are quite similar from a chemical point of view: both forms a Lewis adduct in which the protein side chain represents the electron donor and the metal or the proton play the role of the electron acceptor species. In our model, the acceptor has been translated from the metal to fictitious hydrogen located at the metal-vacant bond axis to preserve the coordination directionality.

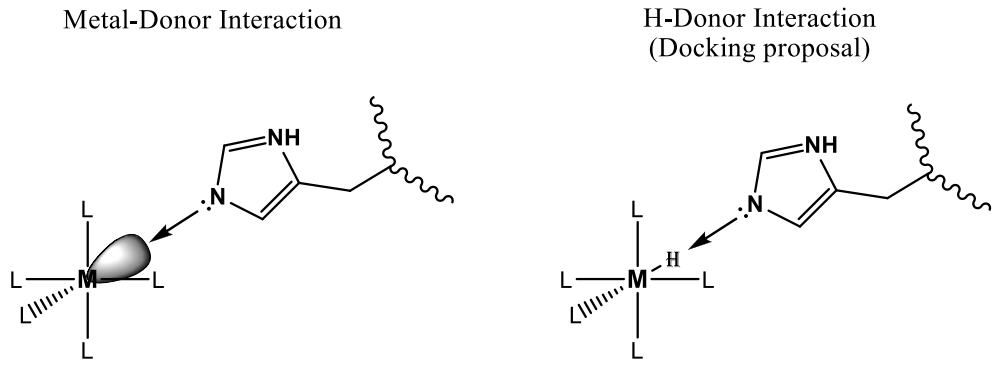


Figure 2. Comparison between the M–donor coordination bond and the GOLD model as a hydrogen bond interaction. His was used as representative amino acid and can replaced by any other coordinating amino acid such as Asp, Glu, Tyr, Ser, Thr, Cys, etc.

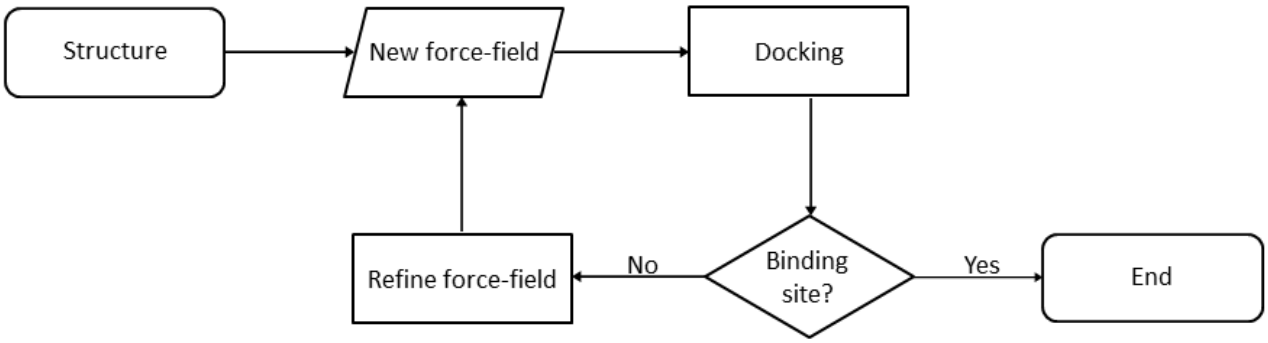
In our case, the scoring function *hbond* parameter of metals (see Supporting Information) was modified so that the software did recognize the metal as h-bond donor liable of a *pseudocovalent* interaction with *hbond* acceptors present in protein[16] and the number of interactions was set to the maximum value possible for the specific geometry of all complexes. In this way, it is possible to evaluate metal interactions in the *hbond intermolecular term* (S_{hbond}) of the scoring function.

This specific term takes into account directionality and distance equilibrium of the bond by the multiplication of its potential for a weight block function ($wt = distance_{wt} + angle_{wt}$), as reported in eq. 1:[54]

$$x_{wt}(x, x_{ideal}, x_{max}) = \begin{cases} 1, & if\ x \leq x_{ideal} \\ 1.0 - \frac{x - x_{ideal}}{x_{max} - x_{ideal}}, & if\ x_{ideal} \leq x \leq x_{max} \\ 0, & if\ x > x_{max} \end{cases} \qquad \textbf{eq. 1}$$

For this approach to work, a dummy hydrogen atom must be added to the metal in the free coordination position. It is important to understand that this dummy hydrogen atom serves only to allow the metal to form a standard hydrogen bond with accepting atoms of the protein residues and to 'activate' this interaction. Moreover, this atom is a pure artifact and so, after a series of optimization tests, its best distance to the metal was set to 0.75 Å. To obtain a good agreement with the experimental coordination bond distances, the H_BOND_LEN parameter of GoldScore, which represents the average length of a hydrogen bond, was modified to 2.0 Å as an optimized value. The use of dummy atoms and, in particular, dummy charges to simulate metals and in general Lewis acid-bases interactions, it is widely-known in the literature and applied in non-bonded molecular dynamics technique.[55-56]

The strength of *pseudocovalent* interactions (contribution to the scoring *Fitness*, see Table S2 of Supporting Information) with all acceptor types was defined in the parameter file of GoldScore. As a first approximation, the force constants were defined for each metal on the basis of the HSAB (Hard and Soft (Lewis) Acid and Bases) theory, considering the hardness/softness acid properties of the specific metals and defining a relative order of affinity toward the amino acid possible donors. The order of magnitude of the starting strength was chosen in coherence with the standard force constants implemented in GoldScore for the evaluation of metals-ligands interactions. The approximated relative strength values were then fine-balanced, for each metal, analyzing for each docking the binding site, spatial orientation, RMSD value and bond lengths, until reaching the best solution for all the structures of the same metal. This iterative process is reported in Scheme 2. The force-field developed for each metal was reported in the Table S2 of Supporting Information.



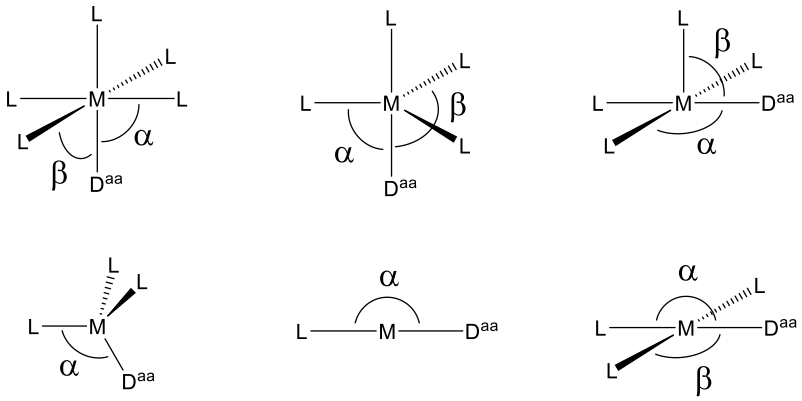
Scheme 2. Force-field refinement workflow used in this work.

Results and discussion

Analysis of 39 metalloproteins. To test our approach, this study consists of redocking (without any geometrical constraints or energy restraints) the 39 metal complexes examined after separating them from the protein at a distance of non-interaction, with the aim to reproduce their crystallographic

structures. The results were always successful and, in all cases, the pose with highest affinity is the one suggested by the X-ray analysis.

The predictive capabilities of the implemented method were evaluated through three different criteria: i) the RMSD value of each docked structure calculated on the heavy atom position; ii) the absolute percent deviation (APD) of the metal–(protein donor) bond length and iii) the APD of the bond angles. These last values were calculated considering the absolute deviation calculated on two selected angles (these angles, named α and β, are reported in Scheme 3).



Scheme 3. Selected bond angles for each geometry.

The mean absolute percent deviation (MAPD) of these selected parameters, calculated according to the eq. 2, and the mean RMSD values are reported in Table 2.

$$MAPD = \frac{1}{N}\sum_{J=1}^{N}\left|\frac{x^{calcd}(j)-x^{exptl}(j)}{x^{exptl}(j)} \times 100\right| \qquad \textbf{eq. 2}$$

where x = is the bond length or bond angle and N is the number of structure of the dataset (39 PDB structures). As pointed out in the literature, the value of MAPD can be chosen as a criterion of quality.[57]

Table 2. Values of RMSD mean, and MAPD for the bond lengths and bond angles obtained with the standard GoldScore and the modification implemented in this work.			
Scoring modification implemented in this work			
	RMSD mean[a,b]	MAPD (bond lengths)[c]	MAPD (bond angles)[c]
	0.958	6.33	6.08
Std. Dev.[d]	0.829	5.38	4.15
Standard GoldScore scoring function			
	RMSD mean[a,b]	MAPD (bond lengths)[c]	MAPD (bond angles)[c]
	3.451	92.49	24.07
Std. Dev.[d]	3.903	151.11	20.20
[a] Value reported in Å. [b] RMSD computed with UCSF Chimera. [c] Value given in percentages. [d] Standard deviation (SD).			

It can be noticed that the mean value of RMSD is very small (0.958 Å with the modified scoring function). Moreover, for 92.3% of the cases it is smaller than the X-ray diffraction analysis resolution. The values of the standard deviation highlight that the error distribution is very close to the mean of the set. Thus, the method appears solid and very effective in the prediction of the correct binding sites. Analogous considerations on the bond lengths and angles show that the proposed method is able to reproduce with great accuracy the structures of the metal complexes bound to a protein. The value of MAPD is 6.33% for the distances and 6.08% for the angles. Overall, the results can be represented graphically fitting them with a Gaussian function as shown in Figure 3.

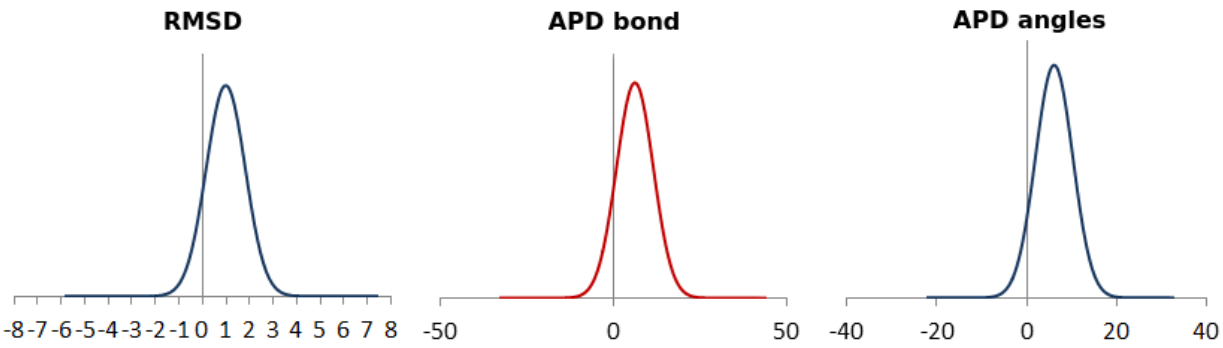


Figure 3. Gaussian fitting of the distribution of RMSD, bond length and bond angles absolute percent deviation.

An accurate analysis of the docking calculations shows that the crystallographic structure is reproduced with a success rate of 100% with a RMSD < 2.5 Å and of 90% with a RMSD < 1.5 Å. Furthermore, an energy and population analysis shows that the best docking solutions in the 90% of the cases have the best score and in the 87% of the cases they are placed in the most populated cluster. The discussed data are reported in Table S3 of Supporting Information and the predicted structures of the entire dataset are shown in Figure 4.

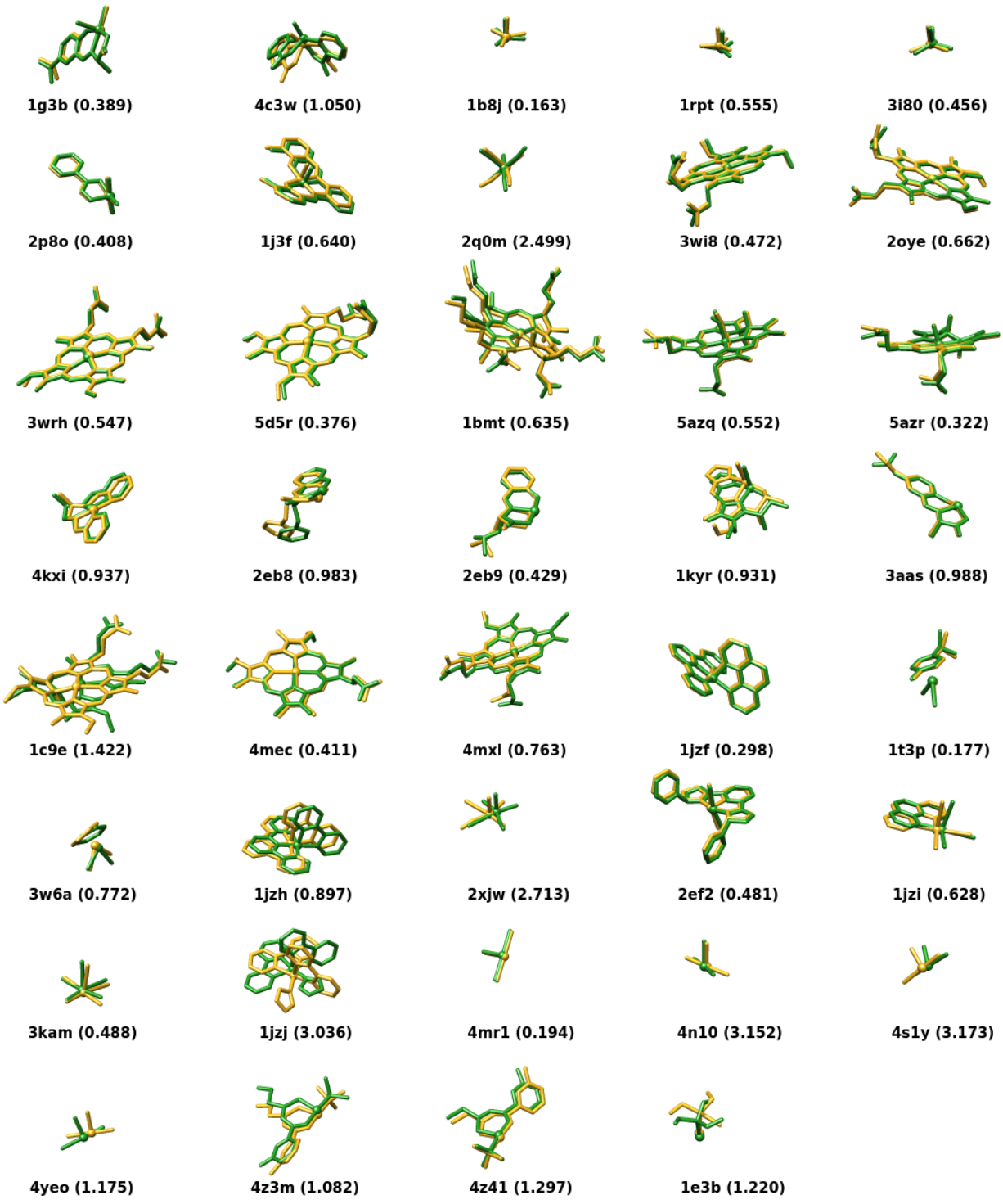


Figure 4. Overlap of the better docking pose (in yellow) and the X-ray structure (in green) of the dataset. In parenthesis the RMSD value for each structure is reported.

In Table 2 the efficiency of our set of modified parameters was compared with the standard GoldScore (GS) scoring function. In this case, the standard GoldScore performed worse and is unable to reproduce the totality of the structures. The success percentage obtained is 53.8% and the mean value of RMSD of the dataset is 3.451 Å versus 0.958 Å with the modified scoring function. Concerning the mean absolute percent deviation from the experimental bond length and angles, the standard GS reach respectively 92.49% and 24.07% with respect to 6.33% and 6.08% achieved by the modification presented in this work and summarized in Table 2. The effect of our modification is shown in Figure 5a in which a comparison between the values of RMSD obtained with the standard GS scoring function (markers in green) versus the values obtained with the implementation of the metal parameters (markers in blue) is reported. In Figure 5 it can be clearly visualized that the formation of the coordination bond improves significantly the accuracy of the prediction of the crystallographic structure (success rate of 100% versus 53.8% of the standard GS, considering RMSD < 2.5 Å).

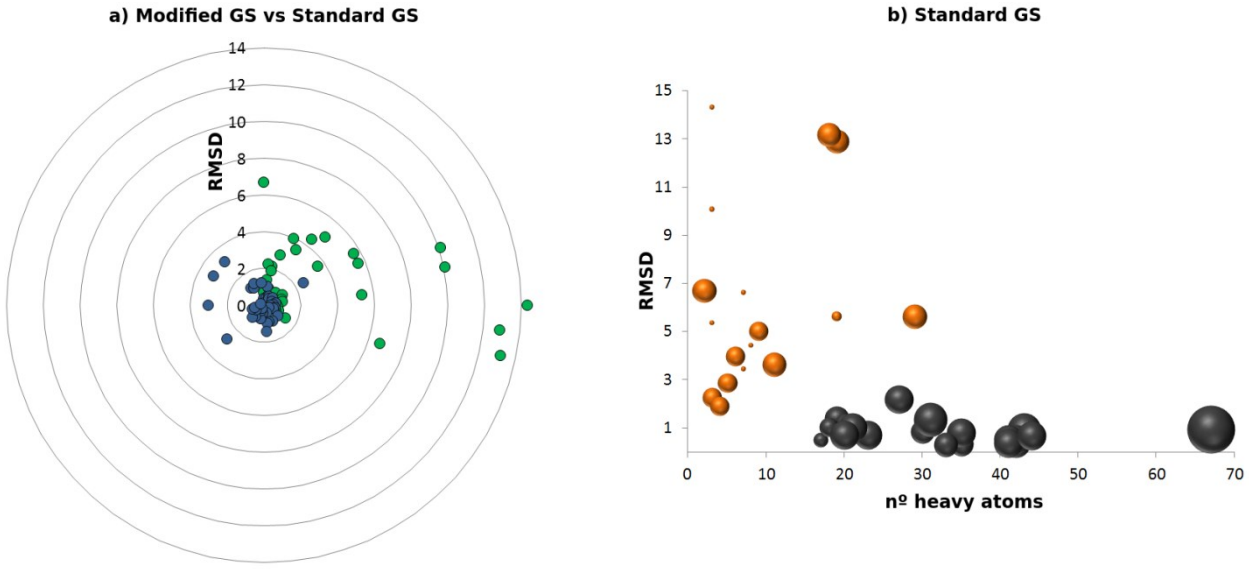


Figure 5. a) RMSD distribution of the docking simulations using the standard (green markers) and modified GoldScore scoring function developed in this work (blue markers); the concentric circles represent the limit of the RMSD value reported in the axis. b) RMSD value of GoldScore standard simulations in function of the number of atoms in the metal-containing *ligands*; the radius of the

markers is proportional to the number of potential interacting groups. The markers in orange represent the structures with an orientation incompatible with the formation of the coordination bond ($x_{wt} \approx 0$, see eq. 1).

It can be noticed that the structures predicted using the standard GoldScore function present lower scoring values compared with those obtained with the implemented methodology and the new set of parameters. In general, using the standard approach, the prediction of the correct binding site is obtained with high molecular mass *ligands* with potential interacting groups such as *hbond* donor or acceptors, polar groups or aromatic rings that compensate, in the recognition process, the absence of effective coordination bonds into the simulated structure and so stabilize the complex throughout second sphere interactions. None of the small structures can be reproduced without the coordination bond stabilization (Figure 5b). In contrast, our approach is able to predict the interaction of a metal complex, independently of the molecular weight of the *ligands* and of the presence of polar groups.

Case study. An interesting case study is represented by four copper(II) structures in which a Cu^{II}-complex is bound to apo-myoglobin, β-trypsin, ferrochelatase and mutant BFPms1 with three different donor types of amino acid side chains. In these structures there is a direct coordination bond between the metal center and a His-N (PDB: 2eb9[36] and 3aas[34]), Tyr-O^- (PDB: 1c9e[33]) and Glu-COO^- (PDB: 1kyr[35]).

A comparison between the experimental X-ray determination and the calculated structures is shown in Figure 6.

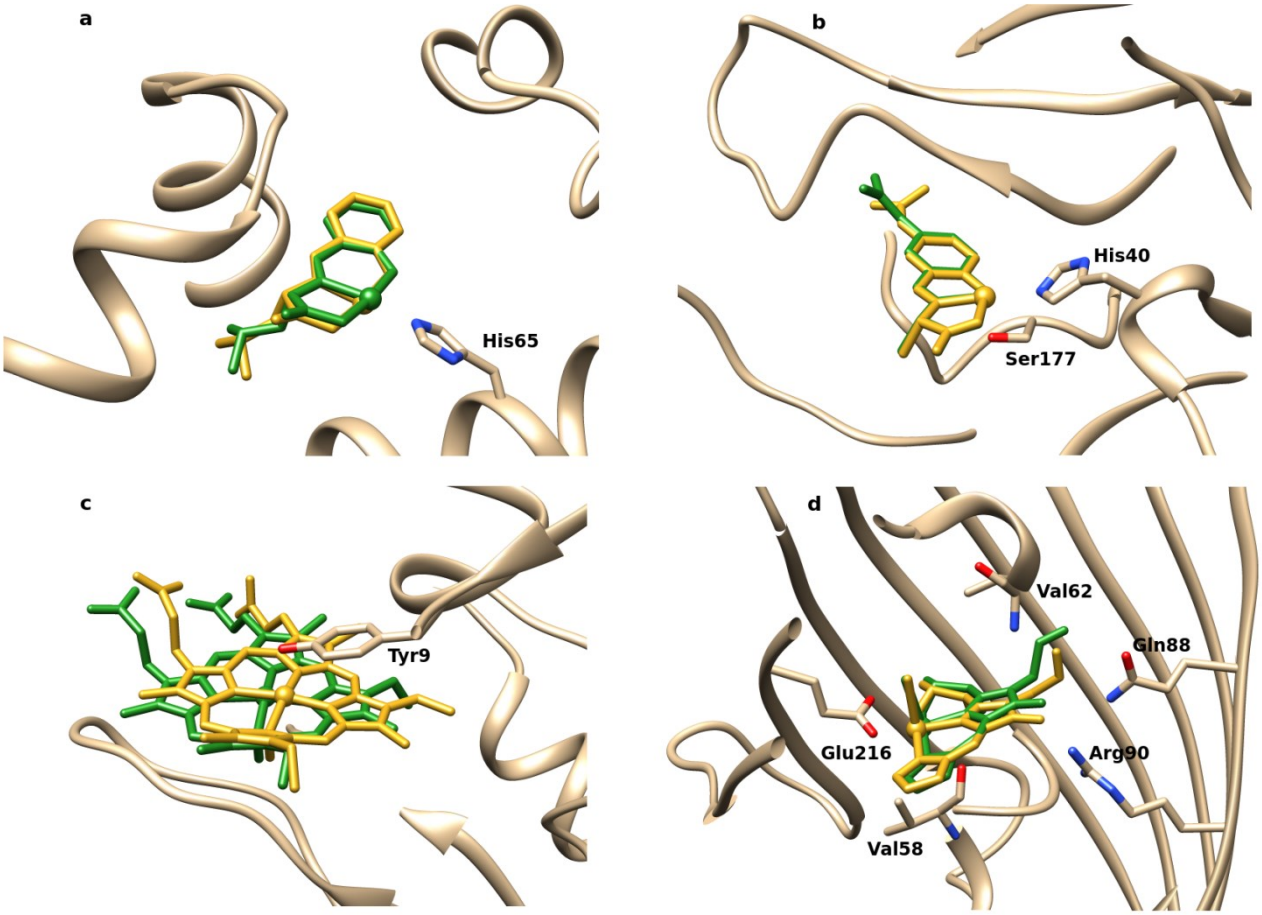


Figure 6. Superimposition of simulated (in yellow) and X-ray (in green) structure of: a) [Cu^{II}(Sal-Leu)(apo-myoglobin)] (PDB: 2eb9); b) [Cu^{II}(*m*-guanidino)(β-trypsin)] (PDB: 3aas); c) [Cu^{II}(*N*-methylmesoporphirin)(ferrochelatase)] (PDB: 1c9e) and d) [Cu^{II}(BFP chromophore Y66H)(BFPms1)] (PDB: 1kyr).

As reported in Table S3 of Supporting Information, all the poses with lowest RMSD have the highest fitness and the related clusters are fully populated (50 of 50 individuals are in the correct binding site). The solutions are very close to the X-ray structures since the calculated RMSD is 0.429 Å for 2eb9, 0.988 Å for 3aas, 1.422 Å for 1c9e and 0.931 Å for 1kyr. The calculated absolute deviation for the only coordination bond length is very small for the four structures: 0.214 Å for 2eb9, 0.017 Å for 3aas, 0.485 Å for 1c9e and 0.151 Å for 1kyr, respectively. The structural data of the discussed case study are summarized in Table 3.

Table 3. Comparison between the calculated (calcd) and experimental (exptl) Cu^{II} structure. The value of RMSD, bond lengths and bond angles are reported for each complex.

PDB	RMSD [a]	Bond length^{calcd}	Bond length^{exptl}	Bond angles^{calcd} [b]	Bond angle^{exptl} [b]
2eb9	0.429	2.279	2.065	α = 96.86 β = 172.22	α = 93.40 β = 171.50
1kyr	0.931	2.023	2.174	α = 105.63 β = 97.36	α = 91.39 β = 79.83
3aas	0.988	2.313	2.296	α = 108.98 β = 166.37	α = 101.39 β = 175.78
1c9e	1.422	2.365	2.850	α = 90.13 β = 110.27	α = 110.27 β = 97.15

[a] Value reported in Å. [b] Value in degrees measured as reported in Scheme 3.

It must be highlighted that the docking calculations allowed us to predict three structures, characterized by the monodentate coordination of three different side-chain donors, such as Tyr-O^- (1c9e), His-N (2eb9 and 3aas) and Glu-COO^- (1kyr). This means that the method and the parameters are well balanced and implicitly take into account the difference in the basicity of these donors.

Conclusions

The study of metal–protein systems is of fundamental importance in biology, pharmacy and medicine.[58] In fact, many metal ions are part of the active site of proteins with transport, storage or enzymatic functions. Often these systems cannot be investigated through X-ray diffraction analysis, and other methods are necessary to characterize the structure and active site. The common spectroscopic and analytical techniques, such as NMR, EPR, ESEEM, ENDOR, ESI-MS, CD and UV-Vis spectroscopy, often do not provide information on the region of the protein where the metal species is bound or on the amino acid residues involved in the coordination.

In this context, the computational methods can represent a valuable alternative to the instrumental techniques;[59-61] in particular, docking methods appear to be suitable in these cases and, over the last years, have been applied to predict the non-covalent bonds (such as hydrogen contacts) between

proteins and organic compounds or metal complexes.[52],[62-63] Concerning the systems with covalent interactions, some authors recently tried to simulate the binding of small inorganic molecules to proteins,[64-65] but up to now there are very few examples in which the docking methods have been systematically applied to the simulation of the covalent bonds between a protein and a metal species which, as mentioned above, represent most of situations.

In this paper, we discuss the potential of docking methods in the prediction of metal–protein structures without using any geometrical constraints or energy restrains. In particular, 39 systems with transition and main group metal-containing *ligands* were examined and new coordination scoring parameters were generated. The results were always successful and, for all 39 structures, the pose with highest affinity is the one suggested by the X-ray analysis; moreover, the crystallographic structure is reproduced with a success rate of 100% with a RMSD < 2.5 Å and of 90% with a RMSD < 1.5 Å.

Therefore, the results suggest that the docking methods could represent a new tool to predict metal complex–proteins interactions and could have a general applicability not only in medicinal but also in bioinorganic chemistry.

As a final comment, it must be noticed that major changes in the metal coordination environment are rarely observed in the structure of metallodrugs bound to their target and, therefore, this should not be a critical point in the quality of the docking prediction. In those cases in which the coordination sphere of the metal could eventually adapt into the binding process, a QM/MM refinement could follow the docking calculations to reach a large conformational exploration.[66-67]

Acknowledgments

J-D.M., A.L., J.R-G. and G.S. are thankful for the support given by the Spanish grant CTQ2014-54071-P and the Generalitat de Catalunya grant 2014SGR989. Support of COST Action CM1306 is

kindly acknowledged. G.S. thanks the Universitat Autònoma de Barcelona for its support to his Ph.D. E.G. thank Fondazione di Sardegna (project FDS15Garribba) for the financial support.

References

[1] M. J. Cleare, J. D. Hoeschele, *Bioinorg. Chem.* **1973**, *2*, 187-210.

[2] N. Farrell in *Comprehensive Coordination Chemistry II*, *Vol. 9* (Eds.: J. A. McCleverty, T. J. Meyer), Pergamon, Oxford, **2003**, pp. 809-840.

[3] *Metallotherapeutic Drugs and Metal-Based Diagnostic Agents. The Use of Metals in Medicine*, John Wiley & Sons Ltd, Chichester, **2005**.

[4] C. Jones, J. Thornback, *Medicinal Applications of Coordination Chemistry*, The Royal Society of Chemistry, Cambridge, **2007**.

[5] J. C. Dabrowiak, *Metals in Medicine*, John Wiley & Sons, Ltd, **2009**.

[6] *Bioinorganic Medicinal Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, **2011**.

[7] K. H. Thompson in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, *Vol.* (Eds.: R. B. King), John Wiley & Sons, Ltd, Chichester, **2011**.

[8] *Interrelations between Essential Metal Ions and Human Diseases*, Springer Science+Business Media, Dordrecht, **2013**.

[9] K. D. Mjos, C. Orvig, *Chem. Rev.* **2014**, *114*, 4540-4563.

[10] S. Medici, M. Peana, V. M. Nurchi, J. I. Lachowicz, G. Crisponi, M. A. Zoroddu, *Coord. Chem. Rev.* **2015**, *284*, 329-350.

[11] A. R. Timerbaev, C. G. Hartinger, S. S. Aleksenko, B. K. Keppler, *Chem. Rev.* **2006**, *106*, 2224-2248.

[12] J. B. Vincent, S. Love, *Biochim. Biophys. Acta, Gen. Subj.* **2012**, *1820*, 362-378.

[13] A. Casini, J. Reedijk, *Chem. Sci.* **2012**, *3*, 3135-3144.

[14] D. Sanna, G. Micera, E. Garribba, *Inorg. Chem.* **2013**, *52*, 11975-11985.

[15] J. Costa Pessoa, E. Garribba, M. F. A. Santos, T. Santos-Silva, *Coord. Chem. Rev.* **2015**, *301-302*, 49-86.

[16] V. Muñoz Robles, E. Ortega-Carrasco, L. Alonso-Cotchico, J. Rodriguez-Guerra, A. Lledós, J.-D. Maréchal, *ACS Catalysis* **2015**, *5*, 2469-2480.

[17] R. Farid, T. Day, R. A. Friesner, R. A. Pearlstein, *Bioorg. Med. Chem.* **2006**, *14*, 3160-3173.

[18] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, *J. Mol. Biol.* **1996**, *261*, 470-489.

[19] A. N. Jain, *Curr. Protein Pept. Sci.* **2006**, *7*, 407-420.

[20] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785-2791.

[21] R. C. Glen, *J. Comput.-Aided Mol. Des.* **1994**, *8*, 457-466.

[22] C. Scholz, S. Knorr, K. Hamacher, B. Schmidt, *J. Chem. Inf. Model.* **2015**, *55*, 398-406.

[23] J. Zou, P. Taylor, J. Dornan, S. P. Robinson, M. D. Walkinshaw, P. J. Sadler, *Angew. Chem., Int. Ed.* **2000**, *39*, 2931-2934.

[24] C. Mügge, T. Marzo, L. Massai, J. Hildebrandt, G. Ferraro, P. Rivera-Fuentes, N. Metzler-Nolte, A. Merlino, L. Messori, W. Weigand, *Inorg. Chem.* **2015**, *54*, 8560-8570.

[25] L. Messori, T. Marzo, C. Gabbiani, A. A. Valdes, A. G. Quiroga, A. Merlino, *Inorg. Chem.* **2013**, *52*, 13827-13829.

[26] S. L. Binkley, C. J. Ziegler, R. S. Herrick, R. S. Rowlett, *Chem. Commun.* **2010**, *46*, 1203-1205.

[27] Y. Satake, S. Abe, S. Okazaki, N. Ban, T. Hikage, T. Ueno, H. Nakajima, A. Suzuki, T. Yamane, H. Nishiyama, Y. Watanabe, *Organometallics* **2007**, *26*, 4904-4908.

[28] T. Santos-Silva, A. Mukhopadhyay, J. D. Seixas, G. J. L. Bernardes, C. C. Romão, M. J. Romão, *J. Am. Chem. Soc.* **2011**, *133*, 1192-1195.

[29] I. W. McNae, K. Fishburne, A. Habtemariam, T. M. Hunter, M. Melchart, F. Wang, M. D. Walkinshaw, P. J. Sadler, *Chem. Commun.* **2004**, 1786-1787.

[30] B. R. Crane, A. J. Di Bilio, J. R. Winkler, H. B. Gray, *J. Am. Chem. Soc.* **2001**, *123*, 11623-11631.

[31] S. Chakraborty, J. Reed, M. Ross, M. J. Nilges, I. D. Petrik, S. Ghosh, S. Hammes-Schiffer, J. T. Sage, Y. Zhang, C. E. Schulz, Y. Lu, *Angew. Chem., Int. Ed.* **2014**, *53*, 2417-2421.

[32] E. Harada, M. Sugishima, J. Harada, K. Fukuyama, K. Sugase, *Biochemistry* **2015**, *54*, 340-348.

[33] D. Lecerof, M. Fodje, A. Hansson, M. Hansson, S. Al-Karadaghi, *J. Mol. Biol.* **2000**, *297*, 221-232.

[34] D. Iyaguchi, S. Kawano, K. Takada, E. Toyota, *Bioorg. Med. Chem.* **2010**, *18*, 2076-2080.

[35] D. P. Barondeau, C. J. Kassmann, J. A. Tainer, E. D. Getzoff, *J. Am. Chem. Soc.* **2002**, *124*, 3522-3524.

[36] S. Abe, T. Ueno, P. A. N. Reddy, S. Okazaki, T. Hikage, A. Suzuki, T. Yamane, H. Nakajima, Y. Watanabe, *Inorg. Chem.* **2007**, *46*, 5137-5139.

[37] Y. Morita, K. Oohora, E. Mizohata, A. Sawada, T. Kamachi, K. Yoshizawa, T. Inoue, T. Hayashi, *Inorg. Chem.* **2016**, *55*, 1287-1295.

[38] C. Drennan, S. Huang, J. Drummond, R. Matthews, M. Lidwig, *Science* **1994**, *266*, 1669-1674.

[39] T. R. M. Barends, L. Foucar, A. Ardevol, K. Nass, A. Aquila, S. Botha, R. B. Doak, K. Falahati, E. Hartmann, M. Hilpert, M. Heinz, M. C. Hoffmann, J. Köfinger, J. E. Koglin, G. Kovacsova, M. Liang, D. Milathianaki, H. T. Lemke, J. Reinstein, C. M. Roome, R. L. Shoeman, G. J. Williams, I. Burghardt, G. Hummer, S. Boutet, I. Schlichting, *Science* **2015**, *350*, 445-450.

[40] C. A. Harman, M. V. Turman, K. R. Kozak, L. J. Marnett, W. L. Smith, R. M. Garavito, *J. Biol. Chem.* **2007**, *282*, 28096-28105.

[41] K. Oohora, Y. Kihira, E. Mizohata, T. Inoue, T. Hayashi, *J. Am. Chem. Soc.* **2013**, *135*, 17282-17285.

[42] M. Razavet, V. Artero, C. Cavazza, Y. Oudart, C. Lebrun, J. C. Fontecilla-Camps, M. Fontecave, *Chem. Commun.* **2007**, 2805-2807.

[43] T. Ueno, T. Koshiyama, M. Ohashi, K. Kondo, M. Kono, A. Suzuki, T. Yamane, Y. Watanabe, *J. Am. Chem. Soc.* **2005**, *127*, 6556-6562.

[44] K. M. Holtz, B. Stec, E. R. Kantrowitz, *J. Biol. Chem.* **1999**, *274*, 8351-8354.

[45] E. Toyota, K. K. S. Ng, H. Sekizaki, K. Itoh, K. Tanizawa, M. N. G. James, *J. Mol. Biol.* **2001**, *305*, 471-479.

[46] T. A. S. Brandão, A. C. Hengge, S. J. Johnson, *J. Biol. Chem.* **2010**, *285*, 15874-15883.

[47] Y. Lindqvist, G. Schneider, P. Vihko, *Eur. J. Biochem.* **1994**, *221*, 139-142.

[48] A. Moulin, J. H. Bell, R. F. Pratt, D. Ringe, *Biochemistry* **2007**, *46*, 5982-5990.

[49] M. F. A. Santos, I. Correia, A. R. Oliveira, E. Garribba, J. C. Pessoa, T. Santos-Silva, *Eur. J. Inorg. Chem.* **2014**, 3293-3297.

[50] G. Ferraro, L. Massai, L. Messori, A. Merlino, *Chem. Commun.* **2015**, *51*, 9436-9439.

[51] I. Shabalin, Z. Dauter, M. Jaskolski, W. Minor, A. Wlodawer, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 1965-1979.

[52] E. Yuriev, J. Holien, P. A. Ramsland, *J. Mol. Recognit.* **2015**, *28*, 581-604.

[53] J. Rodríguez-Guerra, *Insilichem/gaudiview: Pre-alpha public releas, Zenodo*, **2017**.

[54] G. Jones, P. Willett, R. C. Glen, *J. Mol. Biol.* **1995**, *245*, 43-53.

[55] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, R. A. Friesner, *J. Chem. Theory Comput.* **2016**, *12*, 281-296.

[56] F. Duarte, P. Bauer, A. Barrozo, B. A. Amrein, M. Purg, J. Åqvist, S. C. L. Kamerlin, *J. Phys. Chem. B* **2014**, *118*, 4351-4362.

[57] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.

[58] D. Rehder, *Bioinorganic Chemistry*, Oxford University Press, Oxford, **2014**.

[59] F. Neese, *Curr. Opin. Chem. Biol.* **2003**, *7*, 125-135.

[60] F. Neese, *J. Biol. Inorg. Chem.* **2006**, *11*, 702-711.

[61] *Computational Inorganic and Bioinorganic Chemistry*, John Wiley & Sons, Ltd, Chichester, **2009**.

[62] G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff, M. S. Head, *J. Med. Chem.* **2006**, *49*, 5912-5931.

[63] I. A. Guedes, C. S. de Magalhães, L. E. Dardenne, *Biophys. Rev.* **2014**, *6*, 75-87.

[64] G. Bianco, S. Forli, D. S. Goodsell, A. J. Olson, *Protein Sci.* **2016**, *25*, 295-301.

[65] X. Ouyang, S. Zhou, C. T. T. Su, Z. Ge, R. Li, C. K. Kwoh, *J. Comput. Chem.* **2013**, *34*, 326-336.

[66] V. M. Robles, E. Ortega-Carrasco, E. G. Fuentes, A. Lledos, J.-D. Marechal, *Faraday Discuss.* **2011**, *148*, 137-159.

[67] E. Ortega-Carrasco, A. Lledós, J.-D. Maréchal, *J. R. Soc. Interface* **2014**, *11*, 20140090.

Graphical Abstract

Text

Thirty-nine X-ray metal complex–protein structures were predicted using docking methods with new coordination scoring parameters without any geometrical constraints or energy restraints. All the results were successful and the pose with highest affinity agreed with the one proposed by the X-ray analysis, suggesting that this approach could represent a new and generalizable tool to predict metal complex–proteins interactions in medicinal and bioinorganic chemistry.

Graphical Abstract Figure