J Mol Model (2011) 17:2919–2925
DOI 10.1007/s00894-011-0993-8

ORIGINAL PAPER

# Design and dynamic simulation of minimal metallo-proteins

**Nicolò Mazzucco · Stefano Zanconato · Davide De Lucrezia · Emanuele Argese · Irene Poli · Giovanni Minervini**

Received: 28 July 2010 / Accepted: 25 January 2011 / Published online: 12 February 2011


**Abstract** Ab initio in silico design of proteins and enzymes has emerged as a powerful tool to design application-tailored proteins and catalysts for a wide range of applications. Several enzymes exploit the unique features of metal cofactors to achieve catalytic activity otherwise unattainable through the use of only natural amino acid residues. One of the major bottlenecks in ab initio design of novel proteins relies on long-range and epistatic effects that severely limit the possibility of a rational design. Within this framework there is an ongoing effort to reduce protein length and complexity to unlock the full potential of in silico protein design. In this work we specifically address this problem designing and investigating the dynamic features of 10 in silico designed minimal metallo-proteins. In particular, in this paper we investigate whether and to what extent it is possible to design a minimal metallo-enzyme made of only residues involved in metal binding. In this research we address these questions by investigating the ability of 10 different "mini-proteins" with a length shorter than 15 residues. Molecular dynamics studies clearly show that it is possible to design a minimal protein able to bind a metal atom with the correct geometry. It is noteworthy that designed mini-proteins cannot achieve the formation of a canonical hydrophobic core, rather the metal ion provides a "metal core" around which the entire protein is organized. This opens the possibility of designing synthetic enzymes composed of only functional residues organized around a "metal core" which acts as both structural and functional determinat.

**Keywords** Copper · In silico design · Minimal metallo-proteins · Structural bioinformatics

**Electronic supplementary material** The online version of this article (doi:10.1007/s00894-011-0993-8) contains supplementary material, which is available to authorized users.

N. Mazzucco · S. Zanconato · D. De Lucrezia · I. Poli · G. Minervini (✉)
European Centre for Living Technology,
Ca' Foscari University of Venice,
IT-30124 Venice, Italy
e-mail: giovanni.minervini@ecltech.org

N. Mazzucco · E. Argese
Department of Environmental Sciences,
Ca' Foscari University of Venice,
IT-30124 Venice, Italy

I. Poli
Department of Statistics, Ca' Foscari University of Venice,
IT-30124 Venice, Italy

## Introduction

The design of new enzymes through an "in silico" approach is nowadays regarded as one of the most promising approaches to produce new catalysts of interest to biotechnology. Enzymes can be rightly considered as the best existing catalysts. Their unmatched versatility combined with high specificity, stereoselectivity and catalytic efficiency allow their exploitation in a broad range of fields where high efficiency and accuracy is demanded, from APIs (active pharmaceutical ingredients) production to industrial biotechnology for sustainable chemistry. In the early days of computational biochemistry, computers were used primarily for structure-based drug design [1]. At that time the modest computing power offered by first personal computers did not allow a complete simulation of complex systems, but they immediately gave a clear idea about the potential of such an approach. More recently with the advent of more powerful hardware and thanks to the parallel progress of simulation software, computational

biochemistry has gone through a phase of "de novo" design of enzymes. In 2004 Hellinga and coworkers developed a structure-based computational design technique to predict favorable mutations to redesign an extant ribose-binding protein -that normally lacks any enzyme activity- to display triose phosphate isomerase activity. Researchers achieved a rate enhancement of about $10^5$ to $10^6$ by introducing a limited number of mutations (18 to 22) [2]. On the other extreme, Baker and coworkers combined computation and directed evolution to engineer a synthetic protein that catalyzes a Kemp elimination, positioning key residues in a partially dehydrated pocket, based on two alternative catalytic mechanisms obtaining a rate enhancement of up to $10^5$ and multiple turnovers [3, 4]. In many enzymes the catalytic properties are due to the presence of cofactors or metals. In particular, the presence of metal ions assumes a very important role for the catalytic activity since the presence of metals gives the protein a peculiar redox potential otherwise unattainable through the use of only natural amino acid residues. An outstanding example is the de novo design of a four-helical bundle binding two iron ions that displays phenol oxidase activity [5]. The activities provided by the binding of metals are used for a wide range of functions, including: i) active catalysis such as in *carbonic anhydrase* enzyme [6], ii) electronic transfer as in *cytochromes* proteins [7], iii) stabilization of protein structures as in *zinc fingers* proteins [8]. Metal binding is relatively simple for a protein to obtain: a few amino acids are sufficient for a stable metal binding in a catalytic site. Despite recent successes, de novo enzyme design remains a challenging task due to the non-linear interactions among designing variables such as the relationship between protein stability and evolution of new enzymatic specificities as non-obvious trade-off between the acquisition of new enzymatic functions and stability [9], long-range and epistatics effects where the effect of multiple point mutations on enzyme fitness are larger than that expected from the multiplication of their individual effects [10], and finally incomplete knowledge of the underpinning rules that link sequence to structure to function. A typical example is the aforementioned paper by Faiella et al. where the introduction of the cofactor-binding and phenol binding sites was detrimental to the free energy of folding of the protein. In order to rescue the protein's fold, it was necessary to modify the sequence of a loop distant from the active site [5]. In this regard, there is an ongoing effort to reduce protein complexity and length in order to design mini-enzymes. Within this framework, in this paper we address the problem of designing a mini-protein adopting a stable fold exploiting a metal coordination rather than a hydrophobic core. In particular we address the following questions: What happens when the entire peptide is composed of only those residues that bind the metal? What is the minimal number of residues required for a peptide to display a stable backbone conformation and bind at the same time a metal? Finally, is it possible to design a minimal metal-enzyme made of only residues involved in metal binding? In this study we address these questions by investigating the ability of 10 different "mini-proteins" with global length shorter than 15 residues to bind a metal atom. These de novo mini-proteins were designed using the naturally occurring motif "HGHGH", this short sequence is very common in many natural metal binding proteins such as electron transporters [11] and enzymes [12]; however in all extant proteins this motif is always embedded in larger and more complex structures and it has never been reported to occur as a stand-alone module.

## Methods

### Modeling of mini proteins and MiniPro-copper complexes

The three-dimensional model structures of mini enzymes were predicted using the Rosetta ab initio package [13]. For each sequence 10,000 decoys were predicted. The decoys were clustered using the Rosetta *clustering* integrated module. Only the first model proposed for each sequence was taken into consideration. The Rosetta software is not yet able to predict the interaction with metal cofactor using the ab initio method, to avoid artifacts due to the utilization of the *idealize* routine the side-chain conformation of the histidine residues was manually adjusted to obtain a metal-binding site with regular tetragonal geometry, taking as a reference the copper-binding site of bovine superoxide dismutase (PDB code 2SOD) [14]. Due to the small space available for the putative binding site, every model was carefully evaluated to obtain the best combination of protonation states for the histidine residues to assure the lower torsional disorder. Where possible has been preferred the $N^{\varepsilon}$. Stereochemical compatibility test for the metal-binding site formation was carried out by energy minimization in explicit solvent using NAMD [15].

### Molecular dynamics simulations

The complexes were equilibrated in water by molecular dynamics simulations in explicit solvent using the NAMD parallel molecular dynamics code [15] and the CHARMM27 parameters and force field [16]. For water molecules the three-site TIP3p model [17] was used. The minimized complexes were placed in a cubic hexahedron, constructed from a cubic volume of water molecules of dimension $20.0 \times 20.0 \times 20.0$ Å. Periodic boundary conditions were applied to all dimensions. Cutoff was set to 2.8 Å to remove the water molecules overlapping with protein-Cu atoms.

Solvated structures, containing approximately ~1000 solvent molecules, were energy-minimized by applying a harmonic force of 10 kcal $mol^{-1}$ to non-hydrogen atoms of the complexes to allow reorganization of the solvent. The equilibrium was reached when the root mean square deviation (RMSD) of individual atoms (excluding hydrogen atoms) reaches a limiting value of $\Delta RMSD<0.05$ Å with respect to their initial position in the structure at $t_0$. Minimized solvated structures were then subjected to a molecular dynamics simulation with physical parameter set according to standard conditions, after a heating run of 100 ps, during which the temperature was gradually increased from 0 to 298 K. $Na^+$ and $Cl^-$ ions were added into the box to obtain an electro-neutral system. The simulation time step was set to 0.002 ps. The total time of simulation was 4 ns for each model. The same simulation protocol was applied to each model, both with and without the presence of Cu atom. A partial charge of +2*e* were used to model the non-bonded interaction involving copper ions. The CHARMM force field, including van der Waals (vdW) parameters for the $Cu^{2+}$ ion, were used as described by Ungar et al. [18], Ullmann et al. [19], for the interaction between Cu and His residues. Additional parameter for the interaction between Glu residue and Cu were used as described by Wiesemann et al. [20]. All computations were performed on Intel quad-core x86 computers in Scientific Linux 5.5 environment [21].

## Results

The amino acid sequences of de novo designed mini-proteins are reported in Table 1. Molecular dynamic simulations show that the interaction with the copper ion is always maintained in all candidate small proteins except for protein 002; where the major differences observed are in the number of residues involved in binding of the metal (online supplementary material Fig. S1). Binding was monitored for 4ns and in no case the copper ion was coordinated by less than three residues at any time. Detailed analysis of molecular dynamic's trajectories revealed that the most favored interaction involved histidine residues. This is not surprising and it is perfectly in line with what is observed in natural proteins. A case worthy of notice is the behavior of protein 005, the ab initio three-dimensional structure prediction conducted on this particular protein clearly shows that to design a minimal site able to bind a metal atom with the correct geometry four polarizable residues are not always sufficient. In this case, the insertion of a disulfide bridge forces the backbone to take a geometry that does not allow proper positioning of the coordinating residues. This result, in apparent contrast with that obtained for the other nine proposals proteins, could be interpreted as the result of a non-optimal three-dimensional structure prediction by Rosetta. In the large majority of the investigated proteins the disulfide bridge plays a key role in the stabilization of the ion binding. In general the disulfide bridge guides the backbone of the mini proteins to assume a "flat" conformation in which all polarizable residues are oriented in the same direction. A general guideline for a proper design of minimal enzymes should not neglect this evidence. A similar result was also obtained in model 002, in this case the distortion induced by the contemporary presence of a "restricted" disulfide bridge and proline residues, completely inhibits the formation of a small binding site (number of AA involved below 2). Results suggest that two cysteines can be successfully used as "clamp button" to stabilize short polypeptides. In addition, if you want to guarantee a certain flatness to the design, you must foresee the presence of at least six to seven residues between the cysteines involved in the disulfide bridge. This observation is confirmed by the molecular dynamics simulation of the protein 010. After an initial phase of relaxation of the structure (first 150 ps), the system seems to reach an equilibrium already around 500 ps (Fig. 1a). Suddenly at ~950 ps we observe a dramatic change. The movement induced by the variation in the backbone and side chain of glutamate 6 brings the copper ion closer to the glutamate (Fig. 1b). This approach, however, is fatal for the interaction between the ion and histidine 8 which, not held in right position by any interaction with the metal core, abruptly leaves the copper interaction sphere to interact with solvent molecules (Fig. 1c). In Fig. 2, we report the relative distance between Cu ion, glutamate 6 and histidine 8 as function of time. We observe that as the distance between the charged moieties of glutamate 6 approaches the threshold value of <2 Å, histidine 8 abruptly leaves the interaction sphere of copper. In order to understand if the result was only due to an artifact introduced by the simulation, an additional simulations was conducted. Even in that the second independent simulation the interaction between the His residue and

**Table 1** Amino acid sequence of designed mini metallo-proteins

| | |
|---|---|
| CGVGHGHGDGCGHGY | Protein 001 |
| RGHPHCGEGHC | Protein 002 |
| CHGHGHGHC | Protein 003 |
| CGRGHGHGHGC | Protein 004 |
| HCEGHCHPY | Protein 005 |
| CGRGHGHGHGCGH | Protein 006 |
| CHGEGHCPH | Protein 007 |
| RPCGHGHGEGHGHGC | Protein 008 |
| CGDGHGHGHGC | Protein 009 |
| HPHCGEGHC | Protein 010 |

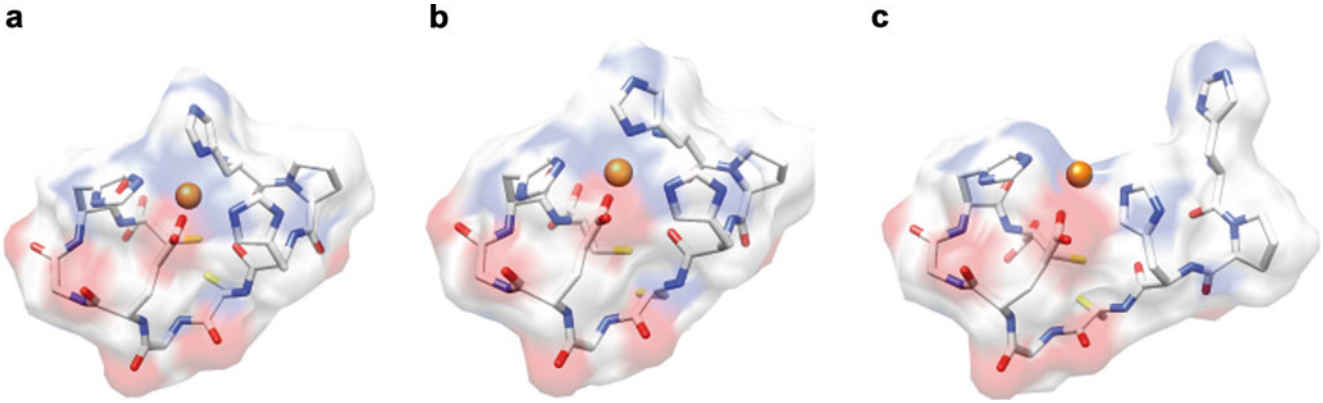


**Fig. 1** The pictures show the conformational change observed for the protein 010 during the molecular dynamics simulation: (**a**) the protein at ~950 ps, four residues binds the ion copper; (**b**) fast spatial reorganization of Glu 6 (between 950 and 965 ps) due to the loss of His 8 coordination (**c**). In orange the ion copper. Figure created with Chimera software [25]

copper is lost at ~800 ps. In Fig. 3a and in Fig. 3b profiles of RMSD obtained by analyzing the trajectories of the two independent simulations are shown. Both simulations show low level of RMSD fluctuations until copper is properly coordinated by four residues. The loss of the interaction between the His1 and copper causes a partial unfolding of the protein 010, measured in a significant increase in RMSD value (RMSD max ~4 Å and ~2 Å, respectively). Notable higher RMSD values were obtained by analyzing the trajectory obtained for the 010 protein in the absence of copper (Fig. 4). This result stresses the importance of the role played by copper in the process of stabilizing the protein fold. Figures 5 and 6 show respectively protein 004 and 006. In these figures the two types of coordination of copper more frequently encountered in this study are highlighted. In particular, protein 4 shows a "3+solvent" coordination in which three histidine residues form a triangular based pyramid with water molecules permanently bound to the other valence available (Fig. 5). Figure 6 shows a rather more conventional square planar coordination where four histidine nitrogen atoms are positioned on the same geometric plane. This geometry was encountered in proteins 006, 007, 008, and 009. Interesting results were obtained by molecular dynamics simulations performed on the same mini-proteins deprived of the metal. For all of them it was observed that the lack of metal produces a "breakup" of the structure. In Fig. 7a, b are shown the RMSD profiles calculated on the trajectory of protein 006 with and without copper ion. They show that in minimal proteins (under 15 residues), where intuitively it is not possible to form a canonical hydrophobic core, metal ions can ensure the formation of a "metal core" around which the entire protein can be organized. This is consistent with what was observed in a much larger copper-protein by Nordlund et al., 2009 [22]. In this study, the de-metallization of the active site of the enzyme superoxide dismutase (SOD1) induces a marked misfolding of the whole three-dimensional structure of the protein. In our case it is therefore clear that these designed mini-proteins must be viewed as a heterogeneous binary system consisting of protein coupled with the metal in which the stabilization of the metal binding pocket is the driving force of the structural organization of the whole protein.

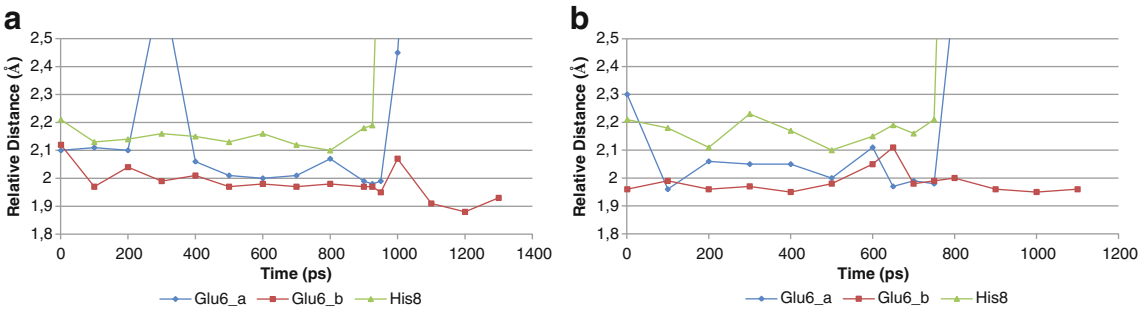


**Fig. 2** The figure shows the relative distance between Glu 6 and His 8 of two independent (**a** and **b** respectively) simulations of protein 010 with copper. The plot represents the distance variation as function of time. Glu6_a and Glu6_b indicate the two tritratable position of the glutamate residue 6. Related distance values are reported in Table S1 available in on-line supplementary materials

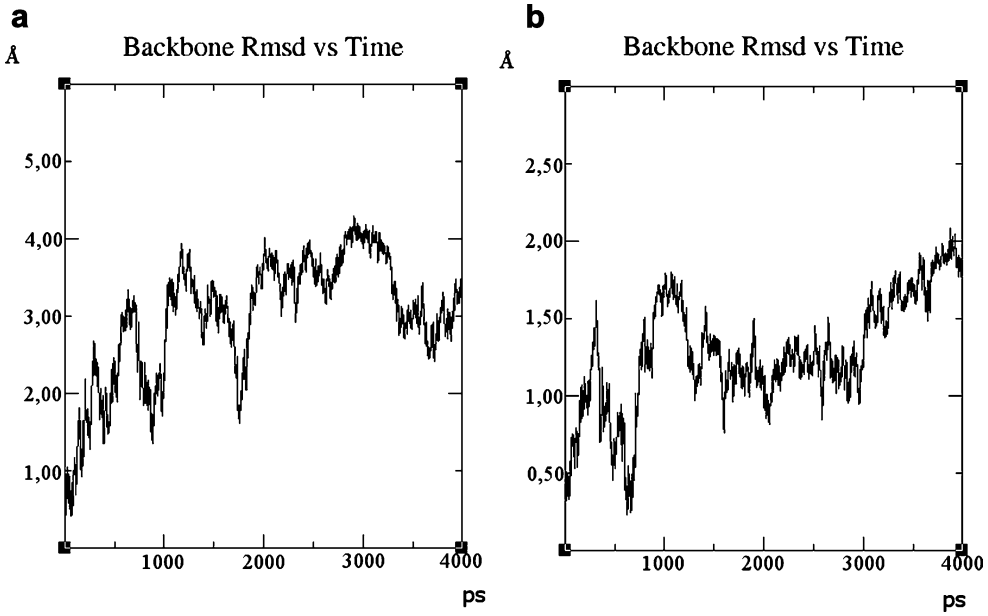


**Fig. 3** The figure shows two independent backbone RMSD profiles of molecular dynamics simulations of protein 010 with copper. The plot represents the RMSD variation as function of time

## Discussion

In this paper we report the investigation of designing extremely short peptides composed by only functional residues and lacking a traditional hydrophobic core. Results highlight the difficulties related to designing and arranging a stable three-dimensional structure in the absence of a hydrophobic core and the necessity to conceive novel structural solutions to circumvent this problem. In particular, in all cases where the exiguous residues number is not sufficient for the formation of a proper hydrophobic core, the utilization of metal ions can assist peptide folding and stability. The formation of a "metallic core" appears to ensure proper structural stability, effectively subsidizing the lack of a traditional hydrophobic core. It is also shown that a small protein requires big attention during the design phase since small spatial re-arrangements in the residues position can lead to huge changes that cannot be anticipated a priori. The formation of a metallic core requires that the cofactor is bound by a proper number of ligands. The square planar geometry indeed turns out to be the one that ensures greater stability. If a metal ion is meant to be used as a spindle for the organization of the entire mini protein scaffold, it will have to be surrounded by at least four ligands. This result must be interpreted within the framework of the limitations of classical molecular dynamics. In particular the force field used in this study does not

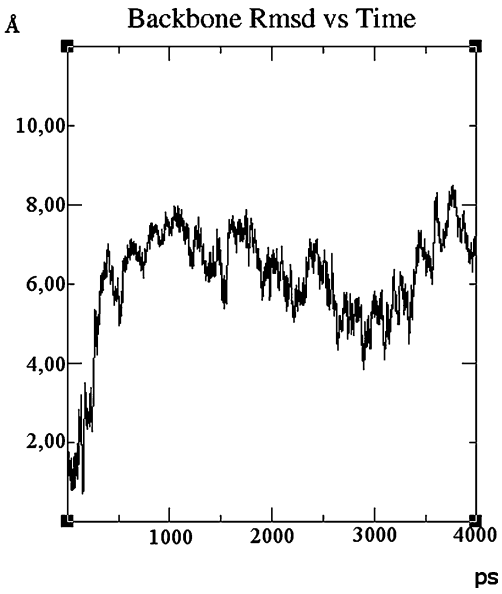


**Fig. 4** The figure shows the backbone RMSD profile of protein 010 without copper ion. The plot represents the RMSD variation as function of time

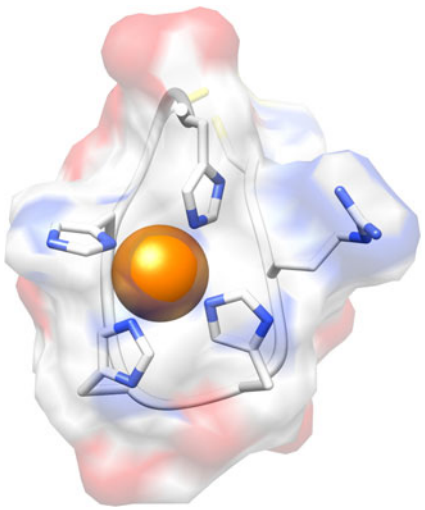

**Fig. 5** The picture shows the three-dimensional structure of protein 006. In orange the ion copper. Figure created with Chimera [25]

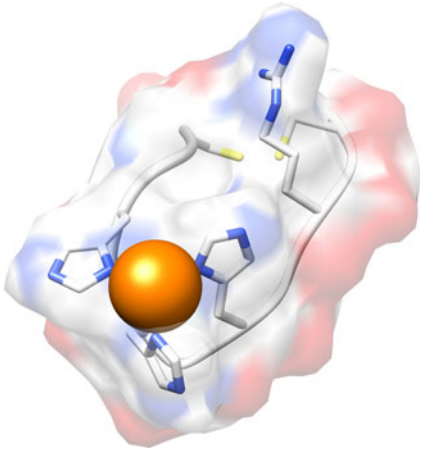

**Fig. 6** The picture shows the three-dimensional structure of protein 004. In orange the ion copper. Figure created with Chimera software [25]

consider some important contributions energy, including the polarization of the electron density. In particular, the copper ion is approximated by a sphere with a fix charge neglecting any possible redox potential transition. The presence of the Cu ion in the binding site is expected to alter the charge distribution of all atoms in the binding residues. Thus an investigation of charge value of every single atom of the ligand residues is advisable by means of either quantum-mechanical dynamics simulations or wet lab experiments. Better force fields may improve our predictions. However, previous studies carried out on very small protein suggest that results obtained in this research can be considered reliable and consistent. Indeed, in a similar research Sakharov et al. [23] investigated the structural properties of small metallo-proteins, i.e., Zinc fingers and zinc ions. Their results show that although classical MD is not able to reproduce the geometry of metal coordination the integration into the CHARMM27 [16], of experimental parameters for vdW and electrostatic interactions, yield reliable predictions. Consequently, we expect our results to be reliable despite the current limitation of MD simulation techniques. Doubtless the most significant result obtained in this research is the proof that a protein can be "cut and reduced" to obtain a protein formed by only the functional residues. These results allow us to imagine very complex scenarios where different desirable activities could be obtained starting from only a few residues.

## Conclusions

In this work we addressed the possibility of creating minimal proteins able to bind metals. The results obtained indicate that a metal ion can be bound by a very "basic" system, but also that the metal itself plays a predominant role in the organization and stabilization of the entire structure. This opens the possibility of designing synthetic enzymes focusing only on the functional residues. In particular, we envisage the ab initio design and construction of synthetic mini metallo-proteins embedded with redox activities such as superoxide dismutation for the treatment of diseases related to oxidative such as ischemia-reperfusion and central nervous system disorders, cardiovascular conditions, cancer, and diabetes [24]. Finally, the simplicity of the design used, allows to draw cambialistic enzymes with different desirable activities related to the different amino acids used to bind the metal and the different redox potential of the metal involved.

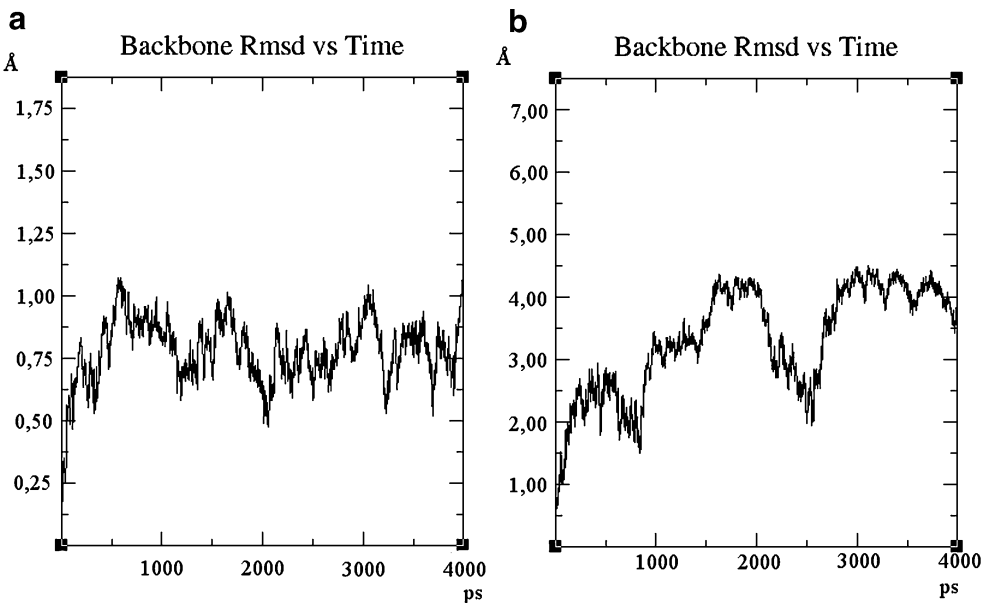


**Fig. 7** The figure shows the backbone RMSD profiles of molecular dynamics simulation of protein 006 with and without the copper ion. The plot represents the RMSD variation as function of time. Consistent fluctuation are visible when the protein is non-bounded to a copper ion

## References


1. Hardy LW, Malikayil A (2003) The impact of structure-guided drug design on clinical agents. Curr Drug Discov 3:15–20
2. Dwyer Mary A, Looger Loren L, Hellinga Homme W (2004) Computational design of a biologically active enzyme. Science 304:1967–1971
3. Röthlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, Albeck S, Houk KN, Tawfik DS, Baker D (2008) Kemp elimination catalysts by computational enzyme design. Nature 453:190–195
4. Alexandrova AN, Roethlisberger D, Baker D, Jorgesen WL (2008) Catalytic mechanism and performance of computationally designed enzymes for kemp elimination. J Am Chem Soc 130:15907–15915
5. Faiella M, Andreozzi C, Martin T, de Rosales R, Pavone V, Maglio O, Nastri F, DeGrado WF, Lombardi A (2009) An artificial di-iron oxo-protein with phenol oxidase activity. Nat Chem Biol 5:882–884
6. Cozier GE, Leese MP, Lloyd MD, Baker MD, Thiyagarajan N, Acharya KR, Potter BV (2010) Structures of human carbonic anhydrase II/inhibitor complexes reveal a second binding site for steroidal and nonsteroidal inhibitors. Biochemistry 49:3464–3476
7. Tollin G, Hanson LK, Caffrey M, Meyer TE, Cusanovich MA (1986) Redox pathways in electron-transfer proteins: correlations between reactivities, solvent exposure, and unpaired-spin-density distributions. PNAS 83:3693–3697
8. Lee S, Doddapaneni K, Hogue A, McGhee L, Meyers S, Wu Z (2010) Solution structure of Gfi-1 zinc domain bound to consensus DNA. J Mol Biol 397:1055–1066
9. Tokuriki N, Stricher F, Serrano L, Tawfik DS (2008) How protein stability and new functions trade off. PLoS Comput Biol 4:e1000002
10. Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS (2006) Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. Nature 444:929–932
11. Nagarajan H, Butler JE, Klimes A, Qiu Y, Zengler K, Ward J, Young ND, Methe BA, Palsson BO, Lovley DR, Barrett CL (2010) De Novo assembly of the complete genome of an enhanced electricity-producing variant of Geobacter sulfurreducens using only short reads. PLoS ONE 5:e10922
12. Remenant B, Coupat-Goutaland B, Guidot A, Cellier G, Wicker E, Allen C, Fegan M, Pruvost O, Elbaz M, Calteau A, Salvignol G, Mornico D, Mangenot S, Barbe V, Medigue C, Prior P (2010) Genomes of three tomato pathogens within the Ralstonia solanacearum species complex reveal significant evolutionary divergence. BMC Genomics 11:379
13. Rohl CA, Strauss CE, Misura KMS, Baker D (2004) Protein structure prediction using Rosetta. Meth Enzymol 383:66–93
14. Tainer JA, Getzoff ED, Beem KM, Richardson JS, Richardson DC (1982) Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase. J Mol Biol 160:181–217
15. James P, Braun C, Wang R, Gumbart W, Tajkhorshid J, Villa E, Chipot E, Skeel C, Kalé DR, Schulten L, Klaus S (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802
16. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616
17. Jorgensen WL, Chandrasekhar J, Madura J, Impley RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
18. Ungar L, Scherer N, Voth G (1997) Classical molecular dynamics simulation of the photoinduced electron transfer dynamics of plastocyanin. Biophys J 72:5–17
19. Ullmann GM, Knapp EW, Kostić NM (1997) Computational simulation and analysis of dynamic association between Plastocyanin and Cytochrome f. Consequences for the Electron-Transfer Reaction. J Am Chem Soc 119:42–52
20. Wiesemann F, Teipel S, Krebs B, Howeler U (1994) Force field calculations on the structures of transition metal complexes. Application to Copper(II) Complexes in Square-Planar Coordination. Inorg Chem 33:1891–1898
21. https://www.scientificlinux.org/
22. Nordlund A, Leinartaite L, Sarabojia K, Aisenbreyc C, Gro G, Zetterstro P, Danielssona J, Loganb DT, Oliveberga M (2009) Functional features cause misfolding of the ALS-provoking enzyme SOD1. PNAS 106:9667–9672
23. Sakharov DV, Carmay L (2005) Zn protein simulations including charge transfer and local polarization effects. J Am Chem Soc 127:4921–4929
24. Batinić-Haberle I, Rebouças JS, Spasojević I (2010) Superoxide dismutase mimics: chemistry, pharmacology, and therapeutic potential. Antioxid Redox Signal 13:877–918
25. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612