FOR THE RECORD

Quantum mechanical analysis of oxygenated and deoxygenated states of hemocyanin: Theoretical clues for a plausible allosteric model of oxygen binding

PIERO FARISELLI,¹ ANDREA BOTTONI,² FERDINANDO BERNARDI,² AND RITA CASADIO^{1,3}

¹Laboratory of Biocomputing, Centro Interdipartimentale per le Ricerche Biotecnologiche (CIRB), University of Bologna, 40126 Bologna, Italy

²Department of Chemistry, University of Bologna, 40126 Bologna, Italy

³Laboratory of Biophysics, Department of Biology, University of Bologna, 40126 Bologna, Italy

(RECEIVED December 30, 1998; ACCEPTED March 29, 1999)

Abstract: In this work with *ab initio* computations, we describe relevant interactions between protein active sites and ligands, using as a test case arthropod hemocyanins. A computational analysis of models corresponding to the oxygenated and deoxygenated forms of the hemocyanin active site is performed using the Density Functional Theory approach. We characterize the electron density distribution of the binding site with and without bound oxygen in relation to the geometry, which stems out of the crystals of three hemocyanin proteins, namely the oxygenated form from the horseshoe crab *Limulus polyphemus*, and the deoxygenated forms, respectively, from the same source and from another arthropod, the spiny lobster *Panulirus interruptus*. Comparison of the three available crystals indicate structural differences at the oxygen binding site, which cannot be explained only by the presence and absence of the oxygen ligand, since the geometry of the ligand site of the deoxygenated *Panulirus* hemocyanin is rather similar to that of the oxygenated *Limulus* protein. This finding was interpreted in the frame of a mechanism of allosteric regulation for oxygen binding. However, the cooperative mechanism, which is experimentally well documented, is only partially supported by crystallographic data, since no oxygenated crystal of *Panulirus* hemocyanin is presently available. We address the following question: is the local ligand geometry responsible for the difference of the dicopper distance observed in the two deoxygenated forms of hemocyanin or is it necessary to advocate the allosteric regulation of the active site conformations in order to reconcile the different crystal forms? We find that the difference of the dicopper distance between the two deoxygenated hemocyanins is not due to the small differences of ligand geometry found in the crystals and conclude that it must be therefore stabilized by the whole protein tertiary structure.

Keywords: *ab initio* computations; allosteric mechanism; arthropod hemocyanins; Density Functional Theory; ligand binding

Hemocyanins are large multisubunit proteins that transport molecular oxygen in a variety of invertebrates with molecular weights up to 3,500 kD. The three-dimensional structures of deoxygenated hexameric hemocyanin from *Panulirus interruptus* and of the deoxygenated and oxygenated forms of the hemocyanin subunit II of *Limulus polyphemus* were refined at atomic resolution 3.2 Å (Volbeda & Hol, 1989), 2.18 Å (Hazes et al., 1993), and 2.4 Å (Magnus et al., 1994a), respectively. A single subunit of hemocyanin comprises 647/628 amino acidic residues with local motifs of secondary structure, and three protein domains (D1, D2 comprising the oxygen binding site, and D3) in the crystals whose detailed structure is available in the Brookhaven Protein Data Bank (PDB). In *L. polyphemus* hemocyanin (oxygenated and deoxygenated), Domain 1 is mainly α -helical with one β -strand; Domain 2, which contains the oxygen-binding site, is also mostly α -helical; Domain 3 is characterized by a seven Greek key β -barrel topology and two long loops, containing a calcium-binding site and two disulfide bonds, respectively (Hazes et al., 1993). The disulfide bonds are remote from the oxygen binding site but are positioned so that they can affect its stability (Topham et al., 1998). Subunit A of *P. interruptus* hemocyanin has a similar three-domain tertiary structure, with three disulfide bonds, two of which stabilize the loop extending from Domain 3 and contacting Domains 1 and 2 (Volbeda & Hol, 1989), as in the *Limulus* protein. In both species, the oxygen binding center is a binuclear copper site, embedded in the protein matrix, with three histidine residues coordinating each copper ion. In the oxygenated form of *L. polyphemus* hemocyanin, the copper–copper distance is 3.60 Å, which is 1 Å less than that seen in the deoxygenated form of the same protein, and it is characterized by almost the same distance found in the deoxygenated

Reprint requests to: Rita Casadio, Department of Biology, Via Imerio 42, 40126 Bologna, Italy; e-mail: casadio@alma.unibo.it.

P. interruptus hemocyanin. Oxygen binding to the active site in *Limulus* promotes a local rearrangement of the site geometry (from trigonal to square planar) and particularly a contraction of the dicopper distance (from 4.6 to 3.6 Å, Magnus et al., 1994a). Away from the oxygen binding site the oxygenated and deoxygenated forms of *Limulus* hemocyanins are rather similar (with a root-mean-square deviation of 0.26 Å; Magnus et al., 1994a) showing that the structural changes detected upon oxygen binding are limited only to the active site. The hemocyanin covalent structures from *L. polyphemus* and *P. interruptus* are about 30% identical (Magnus et al., 1994b), and the overall folds are rather similar. However, a difference in the tertiary structure is evident when the deoxygenated *Limulus* and *Panulirus* hemocyanins are superimposed. The position of protein Domain 1 in the deoxygenated form from *P. interruptus* is rotated by 8° with respect to the active site containing domain as compared to the corresponding form from *L. polyphemus* (Magnus et al., 1994a). Due to this structural differences and to the observations that the crystals from *L. polyphemus* include a chloride atom bound far from the active site (less than 13 Å from CuA), a model of a two-state allosteric regulation was proposed. Analogously to hemoglobin (Monod et al., 1965), the basic oxygen binding behavior of hemocyanin was described with a two-state model, where the hemocyanin hexamer adopts conformational states of relatively high oxygen affinity (R-state) or upon chloride binding, low oxygen affinity (T-state) (Magnus et al., 1994a). Under physiological conditions, deoxygenated hemocyanin is typically in the low affinity state (represented by the deoxygenated form of *L. polyphemus*). In turn, the *Panulirus* deoxygenated hemocyanin would represent the deoxygenated state of the high affinity conformation (R-state). The question then poses as to whether the observed small structural differences at the active site are sufficient to justify the observed shortening of the dicopper distance upon oxygen binding or major rigid body movements as that noticed in the *Palinurus* structure are responsible for the maintenance of the geometry of the ligand binding site and the dicopper distance of 3.5 Å. Assessing mechanisms at a molecular level in the active sites of proteins have been

considered so far a problem not to be tackled with conventional ab initio methods. This is so, considering that the atomic complexity of even the simplest model of an active site (that can be selected from the databases of crystallized proteins) hampers the evaluation of the wave functions of the systems for its computational cost. For this reason tools of computational chemistry have been so far of little use in bioinorganic chemistry, dealing with many enzymes and proteins that have metal ions at their active sites and play a key role in catalysis.

Recently, ab initio computations, however, have demonstrated that it is possible to apply Density Functional Theory (DFT) approaches (Ziegler, 1991) to disclose the binding mechanism of dioxygen molecule to hemocyanin (Bernardi et al., 1996a, 1996b). According to this approach, the computations are performed using simplified models of the active site, whose geometry is taken from the crystal. In the specific case of hemocyanins, we focus on the two copper atoms (CuA and CuB, in the proteins), each trigonally coordinated by three histidine residues of the protein chain. In our models the three copper ligands are simulated by means of three imidazole rings, starting with the same geometry as in the protein active site (Figs. 1–3). All together the models contain 54 atoms of the copper ligands, two copper atoms and eventually the dioxygen molecule, in the oxygenated form, for a total sum of 56 and/or 58 atoms. In principle, the computational approach will highlight whether a given geometry for a given set of interacting atoms is consistent with that observed in the crystal after refinement. This will then allow to discriminate whether the small differences in the relative positions of the imidazole rings are responsible for the drawing of the copper atoms closer to each other upon oxygen binding or not. The geometry of the three different models was optimized with respect to the hydrogen, copper, and/or oxygen atoms, fixing carbon and nitrogen atoms of the imidazole rings. To simulate the protein environment in the models, only copper, hydrogen, and eventually oxygen atoms are free to move and to find the position corresponding to the electronic potential energy minimum. The results are shown in Figures 1–3 and in Table 1, where for sake of simplicity, only the relative distances between the cop-

Fig. 1. Stereo view of the model of the active site of the oxygenated *L. polyphemus* hemocyanin. Geometry is taken from the correspondent PDB file (1OXY) and shown after minimization (for the sake of clarity the nitrogens of the distal imidazole rings are not connected to the copper atoms). Drawing is performed with the RasMol program (Sayle & Milner-White, 1995). Imidazole rings corresponding to the histidine residues are labeled according to the residue position in the protein sequence. Nitrogens closest to the copper atoms are numbered according to the Table 1.

Fig. 2. Stereo view of the model of the active site of the deoxygenated *L. polyphemus* hemocyanin. Geometry is taken from the correspondent PDB file (1LLA) and shown after minimization. Others remarks are as in the caption of Figure 1.

per atoms, as well as the distance from the nitrogens in the imidazole rings bound to the coppers, are listed and compared to those in the crystals. It is evident that both in the oxygenated and deoxygenated models for the *Limulus* hemocyanins the values obtained with our computational approach are rather similar to those found in the crystals (compare 3.55 Å (DFT) with the value of 3.6 ± 0.2 Å reported by Magnus et al. (1994a) and 4.8 Å with the value of 4.6 ± 0.2 Å found by Hazes et al. (1993), for *Limulus* oxygenated and deoxygenated hemocyanins, respectively). This indicates that the optimization procedure promotes only small local rearrange-

ments and confirms that the crystal geometry of the active site is only slightly different from that obtained after geometry optimization of the models with and without bound oxygen (Table 1; Figs. 1, 2). The models are therefore good representatives of the protein active site of *Limulus*, both in the presence and absence of bound oxygen, suggesting that in this case the protein environment can be as well neglected for simulating the active site geometry. Moreover, when oxygen is bound to the model, the charges of the two copper atoms increase from +0.40 to +0.67 electronic unit. This is in agreement with the previously computed reaction mech-

Fig. 3. Stereo view of the model of the active site of the deoxygenated *P. interruptus* hemocyanin before and after minimization. The position of the copper atoms after minimization (black spheres) is shown superimposed to the original position (gray spheres) in the model. Geometry is taken from the correspondent PDB file (1HC6). Others remarks are as in the caption of Figure 1.

Table 1. Geometry of the oxygen binding site of hemocyanins in crystals and after minimization

Hemocyanin, distances ^a	Oxygenated (<i>Limulus</i> , 1OXY ^b) ^c		Deoxygenated (<i>Limulus</i> , 1LLA ^b) ^c		Deoxygenated (<i>Panulirus</i> , 1HC6 ^b) ^c	
	Crystal	DFT ^d	Crystal	DFT ^d	Crystal	DFT ^d
CuA-CuB	3.60	3.55	4.60	4.80	3.55	4.58
N1-CuA	2.40	2.43	2.10	1.97	2.62	2.24
N2-CuA	2.23	2.10	2.01	2.10	1.95	2.03
N3-CuA	2.10	2.00	1.94	1.98	1.97	2.02
N1-CuB	2.44	2.50	2.20	2.23	2.72	2.89
N2-CuB	2.10	1.93	2.10	1.92	2.11	1.90
N3-CuB	1.92	1.93	1.92	1.93	1.95	1.87
O-O	1.42	1.44	—	—	—	—

^aDistances (in Å) are between the two atoms listed in the first column.

^bThe mean coordinate errors are approximately 0.2 Å for 1OXY and 1LLA (Magnus et al., 1994a) and 0.35 Å for 1HC6 (Volbeda & Hol, 1989). Ni are the nitrogens of the ligand histidine rings closest to the copper atom.

^cAcronym of the correspondent PDB file.

^dDistances are obtained after geometry optimization as described in Materials and methods.

anism of oxygen binding to the active site that involves an electron transfer from the copper atoms to the dioxygen molecule (Bernardi et al., 1996b). A quite different result is obtained for the deoxygenated *Panulirus* hemocyanin: the minimized geometry becomes trigonal (with charges of the copper atoms equal to +0.40 electronic unit) and shows a copper–copper distance 1 Å larger than that found in the correspondent crystal (Table 1; Fig. 3). This minimized geometry is very similar to the geometry observed in the model of the deoxygenated *Limulus* hemocyanin. The results clearly indicate that the simulation of the protein backbone that is sufficient to reproduce the *Limulus* crystal geometry fails in the case of the *Panulirus* active site. We, therefore, conclude that in the absence of the protein backbone the more likely copper–copper distance for the deoxygenated form is 4.6 Å, and therefore 1 Å larger than in the presence of the oxygen molecule. The 3.6 Å dicopper distance in the absence of oxygen cannot be explained unless a strained geometry, supported by the protein backbone is assumed. This could be obtained by rigid body movements and domain rotation upon chloride release, as proposed (Magnus et al., 1994a). In this model the deoxygenated and oxygenated *Limulus* hemocyanins correspond to the two low affinity states (T-states) with the ligand unbound and bound, respectively. In turn, the deoxygenated *Panulirus* hemocyanin represents the deoxygenated high affinity state (R-state) and the oxygenated R-state (not crystallized) is postulated to complete the allosteric mechanism (Magnus et al., 1994a). At this stage, due to computational limits, the DFT approach cannot be applied to active site models so large to include also the chloride binding site to simulate possible constraints. However, our present results clearly indicate that the optimized geometry of the model of the deoxygenated form of *Panulirus* shows a minimum dicopper distance that is similar to that of the correspondent form of *Limulus*. From this finding it can be assumed that the whole protein is then necessary to strain the geometry of the active site observed in the *Panulirus* deoxygenated crystal. Further experimental investigations, however, will be necessary to assess whether the strained geometry in the protein is supported by chloride and/or calcium release (the crystal from *Panulirus* does not contain Ca and Cl ions as opposite to those from *Limulus*) and/or other long-range interactions.

In conclusion, our computations indirectly support the previously proposed allosteric mechanism, in which two different conformations switch their dioxygen affinity from a low affinity R-state to a high affinity T-state (Magnus et al., 1994a).

Materials and methods: All DFT calculations were performed with DGauss 2.3.1 within a UniChem environment and run on a CRAY C90. Each minimization took 20,000 s CPU time. We have used the local spin density (LSD) approximation for the exchange correlation potential in the form given by Vosko, Wilk, and Nusair (VWN computations). Furthermore we have considered non-local spin density (NLSD) corrections based on the Becke–Perdew (VWNBP) functional. An LSD-optimized basis set of double-zeta quality in the valence shell plus polarization functions (DZVP) was used. The degree of the SCF convergence is set at a medium level, corresponding to a density fitting coefficient of 5.0E-5 and a total energy of 5.0E-7 (UniChem manual, <http://www.oxmol.com>).

Geometry optimizations are carried out in Cartesian coordinates. For all molecules (Figs. 1–3) the heavy atoms in the imidazole rings are kept fixed to simulate protein rigidity, while the copper, hydrogen, and oxygen atoms are left free to move.

Acknowledgments: This work was supported partially by a grant for a target project in Biotechnology from the Italian Centro Nazionale per le Ricerche (C.N.R.) and by a grant from Ministero della Università e della Ricerca Scientifica e Tecnologica (MURST) given to the project “Biocatalisi e Biconversioni,” both delivered to R.C. Calculations were performed at CINECA, Bologna, Italy using a Cray C90 machine thanks to the grant bolbob41 given to R.C.

References

- Bernardi F, Bottoni A, Casadio R, Fariselli P, Rigo A. 1996a. An ab-initio study of dioxygen binding site of hemocyanin: A comparison between a CASSCF and DFT approach. *Int J Quant Chem* 56:109–116.
- Bernardi F, Bottoni A, Casadio R, Fariselli P, Rigo A. 1996b. An ab-initio study of the mechanism of the binding process of triplet O₂ to hemocyanin. *Inorg Chem* 35:5207–5212.
- Hazes B, Magnus KA, Bonaventura C, Bonaventura J, Dauter Z, Kalk KH, Hol WGI. 1993. Crystal structure of deoxygenated *Limulus polyphemus* subunit

- II hemocyanin at 2.18 Å resolution: Clues for a mechanism for allosteric regulation. *Protein Sci* 2:597–619.
- Magnus KA, Hazes B, Ton-That H, Bonaventura C, Bonaventura J, Hol WGJ. 1994a. Crystallographic analysis of oxygenated and deoxygenated states of arthropod hemocyanin shows unusual differences. *Proteins* 19:302–309.
- Magnus KA, Ton-That H, Carpenter JE. 1994b. Recent structural work on the oxygen transport protein hemocyanin. *Chem Rev* 94:727–735.
- Monod J, Wyman J, Changeaux JP. 1965. On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12:88–118.
- Sayle RA, Milner-White EJ. 1995. RASMOL: Biomolecular graphics for all. *Trends Biochem Sci* 20:374.
- Topham R, Tesh S, Cole G, Mercatante D, Westcott A, Bonaventura C. 1998. Active-site disruption in native limulus hemocyanin and its subunits by disulfide-bond reductants: A chemical probe for the study of structure-function relationships in the hemocyanins. *Arch Biochem Biophys* 352:103–113.
- Volbeda A, Hol WGJ. 1989. Crystal structure of hexameric hemocyanin from *Panulirus interruptus* refined at 3.2 Å resolution. *J Mol Biol* 209:249–279.
- Ziegler T. 1991. Approximate density functional theory as a practical tool in molecular energetics and dynamics. *Chem Rev* 91:651–670.