Modelling copper-containing proteins

Bosch, Marieke van den

Chapter 2

MOLECULAR DYNAMICS SIMULATIONS OF TYPE-I CU-CONTAINING PROTEINS: DEVELOPMENT OF AN EMPIRICAL, NON-BONDED FORCE FIELD

Summary

To verify the stability of a metal structure using modelling techniques, common bonding force fields can not be used to describe the metal site. In this research a non-bonded force field was developed for a type-I Cu-site and applied to several Cu-containing proteins. The force field was developed using a potential energy surface. The Cu-position was varied in a grid and in every grid point the energy was calculated. The optimum position of the Cu-atom and the increase in energy when the Cu moves away from this point were determined while varying the force field parameters. After comparison to the crystal structure, a set of parameters was selected. The parameters were checked to validate the stability of the Cu-site of azurin in a Molecular Dynamics (MD) simulation. Next, MD-simulations were performed to check whether the force field was applicable to other type-I containing proteins, a type-II Cu-site and several azurin variants where the Cu-ligand Met 121 was replaced with several other residues.

2.1 Introduction

Proteins use the transition metal copper in many different ways (Holm *et al.*, 1996). The blue copper proteins are involved in electron transfer, others function as dioxygen binding proteins like the hemocyanins. Some catalyse reactions like tyrosinase that activate O_2 for hydroxylation of monophenols and nitrite reductase where a redox centre and a catalytic centre collaborate in the turnover of nitrite into NO. Finally, chaperones are used for the storage of Cu in soluble form.

Simulations of transition metals often are problematic because of the necessity to model the diversity of their co-ordination structures. Because of the high polarisability of transition metal ions, their charge and bonding properties vary easily (Jensen, 1999). The number of ligands may vary as well as the co-ordination structure at a given number of ligands, e.g. tetrahedral vs square-planar for 4-co-ordinated metals or square pyramidal vs trigonal bipyramidal for 5-co-ordinated metals. This has the consequence that the angle bend function has more than one ideal bond angle and a Taylor expansion around a single value is not suitable. Because of these difficulties, classical force field parameters describing the bonded and non-bonded interactions involving transition metal ions are not standardly incorporated in main force field libraries. Accurate force fields are developed for a specific co-ordination compound or metallo-protein, as mentioned in chapter one. Quantum chemical calculations like density functional theory (DFT) are often used to determine the charge distribution within a metal site. The bonding parameters are usually based on experimental data. Structural features like equilibrium distances and angles are often obtained from X-ray diffraction (XRD), the concomitant force constants that are related to the vibrational frequencies can be obtained from, for instance, Resonance Raman (RR) spectra (Banci *et al.*, 1992; Buning and Comba, 2000; Comba and Remenyi, 2002; de Kerpel *et al.*, 1999). There are several drawbacks of these force field parameter sets. First of all, DFT calculations are computationally expensive and are very dependent of the (input) structure. The consequence is that the obtained parameters are only valid for the specific structure used in the DFT calculation. Furthermore, when there are no experimental data like XRD measurements supplied, no force field parameter set can be obtained. Secondly, metal-ligand bonds are relatively easily built and broken; metals can be easily incorporated into a protein at a later stage.

The present models often make use of harmonic or quartic functions, which are not able to break, to represent the covalent bonds between the metal and its ligands. Therefore, the stability of Cu-sites cannot be modelled using these force fields. A Morse potential would be a better representation of these metal-ligand bonds but it is very hard to parameterise.

In this project a non-bonded copper force field was developed. By manipulation of the force field parameters for the copper site the parameters that describe the copper site of azurin best were extracted. In the type-I Cu-site, copper is co-ordinated by this $S\gamma_{Cys}$ and two $N\delta_{His}$ atoms at distances varying between 1.9 to 2.3 Å. At a significantly larger distance (2.5 to 3.1 Å) an axial ligand, usually methionine, closes the copper site. In some cases like stellacyanin a carboxyl group from a glutamine residue is assigned to this task. Azurin is the only blue copper protein known to have a weak, fifth co-ordination ligand: a peptide carbonyl from a glycine residue at 3.0Å, resulting in a trigonal bipyramidal co-ordination site. In a type-II Cu-site, like the catalytic centre of nitrite reductase, the Cu is co-ordinated by three $N\varepsilon_{His}$ atoms while the fourth ligand site is occupied by water or the nitrite.

A grid was designed within the type-I Cu-site. The total energy of the system was calculated as a function of the Cu-position on the grid. This was repeated for various force field parameters. Depending on the energy profile, a parameter set was obtained. Molecular Dynamics (MD) simulations were performed to check the structure and the stability of the metal site of azurin. Tests were also performed on other type-I Cu-containing proteins, the type-II Cu-site of nitrite reductase and on azurin mutants where the methionine ligand was replaced.

2.2 Methods

2.2.1 Potential Energy Surface

From the XRD-structure of oxidised azurin at pH 5.5 (Nar *et al.*, 1991), the residues co-ordinating the Cu (Gly 45, His 46, Cys 112, His 117 and Met 121) were distilled. The residues were cut-off at their Cα positions, except for residues Gly 45 and His 46 of which also the backbone atoms were taken into account. The result is 30 atoms

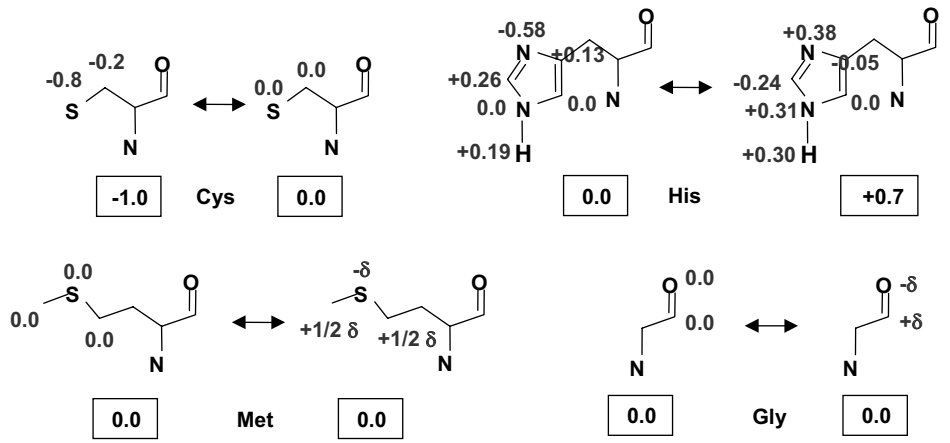


Figure 2.1: Partial charge distribution of the atoms in the copper-co-ordinating amin acids.

excluding aliphatic hydrogen atoms that are not explicitely defined. An internal co-ordinate system was defined where the **origin** is the mathematical average of the co-ordinates of the three equatorial ligands: $N\delta_{His46}$, $S\gamma_{Cys112}$ and $N\delta_{His117}$. The internal co-ordinates with respect to the midpoint are based on the position of the ligands. The plane spanned by the three equatorial ligands, the so-called N_2S-plane, is the xy-plane where the S_{Cys112} is situated on the x-axis at (-2.26, 0.00, 0.00)Å, $N\delta_{His46}$ at (+1.33, +1.50, 0.00)Å and $N\delta_{His117}$ at (+1.20, -1.77, 0.00)Å. The z-axis is perpendicular to the xy-plane where the positive direction points to the methionine residue. The $S\delta_{Met\ 121}$ is placed at (+0.95, +0.24, +3.02)Å and O_{Gly45} at (+0.58, +0.38, -2.76)Å. In the crystal structure, the average of the Cu-atom position is located at (-0.08±0.05, +0.05±0.04, +0.09±0.05) Å with respect to the origin.

The potential energy of the system is mapped while positioning the Cu-atom on a grid with steps of 0.05 Å in the x- and y-direction and 0.1 Å in the z-direction. This procedure has been repeated for various force field parameters, see further.The shape of the potential energy surface has been analysed in the three directions of the co-ordinate system by carrying out a harmonic fit ($y = ½\ k_{harm}x^2$). Unless mentioned otherwise, the fitting of the curves have a regression coefficient, R^2 value, larger than 0.99.

2.2.2 Force field

The van der Waals parameters for Cu were taken from the GROMOS96 ff (van Gunsteren *et al.*, 1996). However a small correction was made since it was noticed that the Cu-N distance was systematically modelled too large. In the GROMOS96 ff, the van der Waals interaction is defined as:

$$V_{vanderWaals} = \frac{C_{12}(i,j)}{r^{12}} - \frac{C_6(i,j)}{r^6} \qquad \textbf{(2.1)}$$

where $C_{12}(i,j)=\sqrt{C_{12}(i,i)}*\sqrt{C_{12}(j,j)}$ and similar $C_6(i,j)=\sqrt{C_6(i,i)}*\sqrt{C_6(j,j)}$. Depending on the atom type j, different values are used for $\sqrt{C_{12}(i,i)}$. We reduced the $\sqrt{C_{12}}$(Cu) and $\sqrt{C_{12}}$(N) parameters when interacting with each other to the lowest value defined in the force field library.

The charge distribution was systematically varied. The charge of the Cu was delocalised to the cysteine and the histidine ligands. The total charge of the cysteine is distributed over its atoms by means of an interpolation between a charge distribution according to the GROMOS96 ff as in the unprotonated state with a formal charge of –1 and in the protonated state with a formal charge of 0. For the histidines the charge distribution is interpolated between the charge distribution in the singly protonated state with a formal charge of 0 and the doubly protonated state with a formal charge of +1e of which 0.3e charge i

situated on the Hδ atom in the GROMOS ff. Since the Cu-atom replaces this atom the 0.3e remains situated on the Cu atom so that only 0.7e can be transferred to the histidine residues. The interaction with the axial ligands has been manipulated by altering the polarisation, δ, of the axial ligands, see Figure 2.1.

2.2.3 Molecular Dynamics simulations

The starting structures for the proteins were obtained from the Protein Data Bank (PDB) (Berman *et al.*, 2000) codes: 4AZU for azurin (Nar *et al.*, 1991), 1BAW for plastocyanin (Bond *et al.*, 1999), 1PZA for pseudoazurin (Vakoufari *et al.*, 1994), 1AAC for amicyanin (Durley *et al.*, 1993), 1RCY for rusticyanin (Walter *et al.*, 1996), 1QHQ for auracyanin (Bond *et al.*, 2001), 2CBP for Cucumber Basic Protein (CBP) (Guss *et al.*,

1996), 1JER for stellacyanin (Hart *et al.*, 1996), 2AFN for nitrite reductase (Murphy *et al.*, 1995), 1A4A for M121H azurin (Messerschmidt *et al.*, 1998), 2TSA for M121A azurin (Tsai *et al.*, 1996), 1ETJ for M121E azurin and 2AZA (Karlsson *et al.*, 1997) for M121Q azurin (Romero *et al.*, 1993).

The GROMACS package (v3.2) (Berendsen *et al.*, 1995; van der Spoel *et al.*, 1999a) was used for all simulations combined with the Gromos 43a2 force field (van Gunsteren *et al.*, 1996). The proteins were centred in a cubic box where the distance between the proteins and the box is at least 0.6 nm. The simulation boxes were filled with SPC (Berendsen *et al.*, 1981) water molecules. First the solvent was relaxed while the protein was frozen, next the total system was optimised. The optimisations were performed using a the steepest descent algorithm with a tolerance of 10 $kJ.mol^{-1}.nm^{-1}$. MD-simulations were performed at 300 K. Protein and solvent were separately coupled to a temperature bath at 300 K using the weak coupling scheme with a relaxation time of 0.1 ps. The initial atomic velocities of the optimised protein were generated randomly. An isotropic constant pressure was applied using a relaxation time of 1 ps. After removing fast fluctuations from the system by constraining the bond lengths within the protein (Hess *et al.*, 1997) and the water geometry (Miyamoto and Kollman, 1992) the use of a 2 fs time step is permitted. A group-based twin range cut-off-scheme was employed for the non-bonded interactions, with short range cu-toff of 0.8 nm and a long

Table 2.1: Position of energy minimum (Å) and harmonic fit (10 $kJ.mol^{-1}.Å^{-2}$) of the energy along the x, y and z axis in the grid upon charge (e) transfer from Cu to the cysteine residue.

Charge		Position of energy minimum			k_{harm}		
Cu	Cys	x	y	z	x	y	z
+2.0	-1.0	-0.40	+0.05	+0.2	170	21	25
+1.8	-0.8	-0.30	+0.05	+0.2	105	33	22
+1.5	-0.5	-0.15	+0.05	+0.2	77	60	15
+1.3	-0.3	-0.05	0.00	+0.3	73	89	9
+1.2	-0.2	0.00	0.00	+0.3	74	107	6
+1.0	+0.0	+0.15	0.00	+0.7	44	98	16

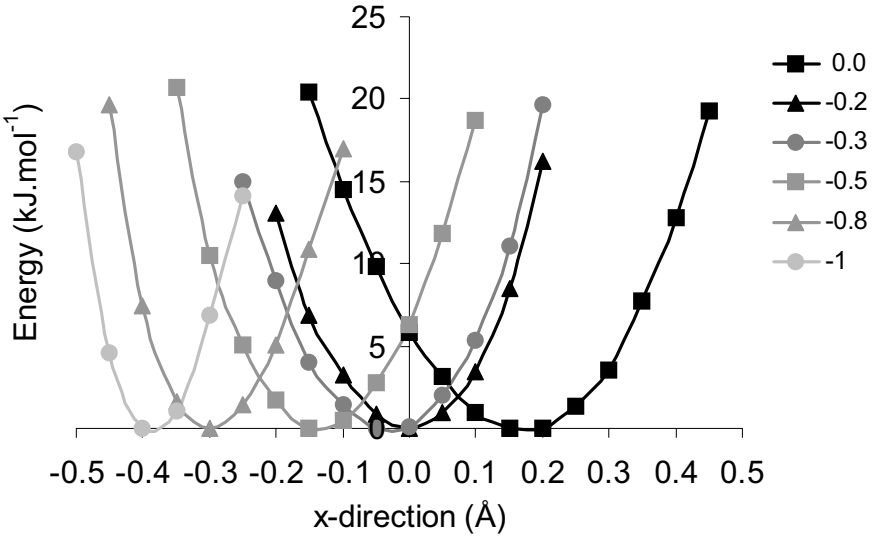


Figure 2.2: The effect of the energy surface in the x-direction upon different charges on the cysteine residue. The energy minima were set to 0 kJ.mol^{-1}.

range cut-off of 1.4 nm, while the pairlist was updated every 5 steps. To approximate the electrostatic interactions beyond the long-range cut-off, a Poisson-Boltzmann reaction-field force was used. The value for the dielectric permittivity of the continuum outside the long-range cut-off was set to 66.

2.3 Results and discussion

2.3.1 Force field development using a potential energy surface

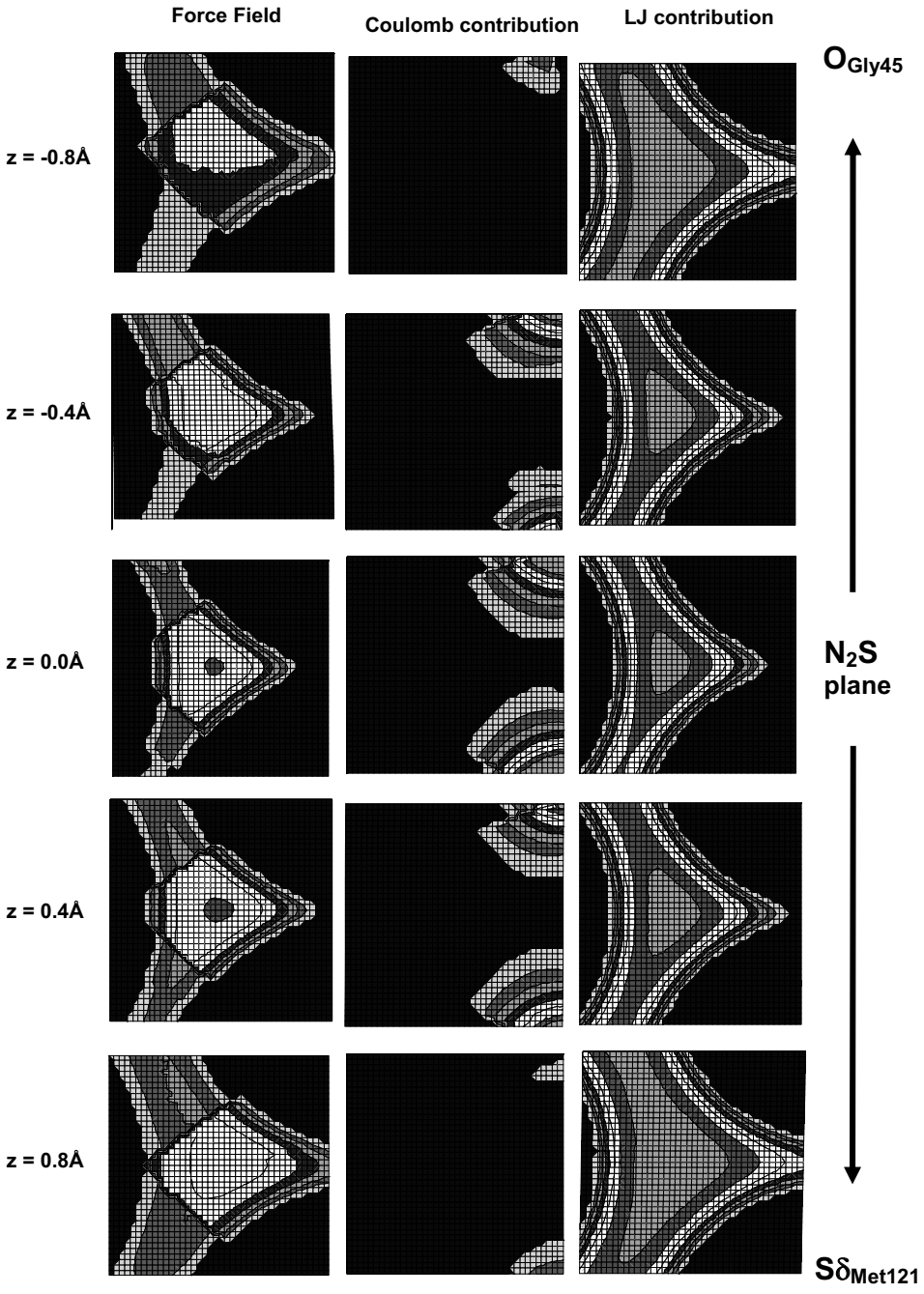
Table 2.1 shows the results of the variation of the energy minimum position of the Cu atom upon delocalising charge to the cysteine ligand. Herein, the charges of the other ligand residues were kept constant. The total charge on the histidine is zero and the polarisation of both axial ligands, δ=0.2 in Figure 2.1. When no charge is delocalised and the full +2e charge is located on the Cu atom, the attraction between the Cu and the cysteine residue is large, resulting in an energy minimum position located at 1.87 Å from

the $S\gamma_{Cys}$. Charge backdonation from the cysteine ligand consequently leads to a larger optimal Cu-$S\gamma_{Cys}$ distance. The charge delocalisation has a negligible effect on the energy minimum Cu position in the y- and z-direction. Only in the case of a full charge transfer from the Cu towards the cysteine ligand, a relatively large movement of the optimal Cu position in the z-direction is observed. A charge of +1.3e on the Cu ion and a remaining charge of –0.3e on the cysteine residue, leads to a Cu-$S\gamma_{Cys}$ equilibrium distance of 2.23Å. It is the optimal Cu-Cys charge transfer since at this point the Cu atom is only 0.03 Å removed from its crystallographic position in the x-direction.

When plotting the potential energy as a function of the x-co-ordinate of the Cu for the different charge distributions, Figure 2.2 is obtained. It clearly shows the movement of the energy minimum of the Cu-atom to larger Cu-$S\gamma_{Cys}$ distances upon more delocalisation of the Cu charge. A harmonic fit was performed on the energy curves, see Table 2.1. Upon more delocalisation, $k_{harm,x}$ decreases while $k_{harm,y}$ increases. Furthermore, $k_{harm,z}$ is much smaller compared to the values in the x- and y-direction. These differences can be explained after taking a look at the individual non-bonding energy terms. The left part of Figure 2.3 shows the energy profile of Cu in the xy-plane for various z-co-ordinates in the case where the Cu is charged with +1.3e and the cysteine with –0.3e. The middle and right show the individual contributions of the non-bonding energy terms. Table 2.2 contains the results of the second order fit around the **origin** of the internal co-ordinate system of the three energy profiles. In the right colomn of Figure 2.3 we see a rapid increase of energy when Cu moves towards the ligands arising from the steric LJ-interaction ($\sim 1/r^{12}$) resulting in high values for k_{harm}: 990 and 1460 $kJ.mol^{-1}.Å^{-2}$ in the x- respectively y-direction. In the z-direction, the LJ-energy

Table 2.2: Second order fit (10 $kJ.mol^{-1}.Å^{-2}$) of the energy in the **origin** along the x, y and z axis of the total and individual non-bonding interactions.

	ff	LJ	Coulomb
k_x	82	99	-12
k_y	129	146	-16
k_z	5	-15	20


Force Field
Coulomb contribution
LJ contribution
O_{Gly45}
z = -0.8Å
z = -0.4Å
z = 0.0Å
z = 0.4Å
z = 0.8Å
N_2S plane
$S\delta_{Met121}$

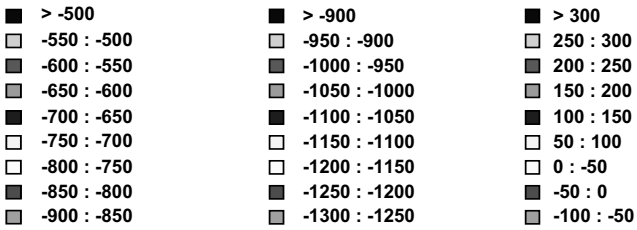


Figure 2.3: Energy profile of the copper atom in the type-I site of azurin using the force field as described in the text (left colomn). The energetic terms are subdivided in the Coulomb (middle colomn) and Lennard-Jones contribution (right colomn). The horizontal axis of the 15 pictures depict the x-axis, the vertical axis is the y-axis, see text for more information. On the right the value of the z-axis is depicted.

increases when approaching the N_2S plane (z=0.0 Å) due to the repulsion with the equatorial ligands. The second order fit results therefore in small negative values for k_z of -150 kJ.mol^{-1}.Å^{-2}. The Coulomb interaction, on the other hand, pulls the Cu atom towards the three equatorial ligands, see the middle column of Figure 2.3. This results in negative values for k of -120 and -160 kJ.mol^{-1}.Å^{-2} in the x- and y-direction, respectively and positive values in the z-direction of 200 kJ.mol^{-1}.Å^{-2}. Having analysed the individual non-bonding interactions, the differences of k_{harm} around the **energy minimum** in Table 2.1 can be explained. Upon more delocalisation and thus less Coulomb attraction between the Cu and its ligands, a less negative number for k_x would be expected. However, due to the movement of the energy minimum away from the $S\gamma_{Cys}$ the LJ-interaction decreases as well and a lower $k_{harm,x}$ is the result. This also affects the interaction with the histidine residues. Due to the charge delocalisation, the Cu atom has a lower attractive Coulomb interaction with the histidines and a higher LJ repulsion due to the movement of the energy position towards them, both responsible for an increase of $k_{harm,y}$. In the z-direction an almost linear relationship is observed between the Cu-cysteine charge delocalisation and the k_{harm}. Upon less charge being placed on the Cu atom, the overall lower Coulomb attraction with the equatorial ligands results in a lower k_{harm} value. Only in the case of a full charge transfer, the energy minimum shifts so much in the positive z-direction that the interaction with the axial methionine ligand becomes significant.

Table 2.3 provides information about the fitting parameters when transferring charge to the histidine residues. The force field parameters of the two histidine ligands was the same. The optimal charge delocalisation, as described above, of 0.7e from Cu to the cysteine residue was used. Overall, only small differences are observed. The only result of transferring more charge to the histidines is that the LJ interaction is able to push the Cu more out of the N_2S-plane decreasing all k_{harm}. Since the energy minimum is positioned at a distance of 2.06 and 2.19 Å for $N\delta_{His46}$ and $N\delta_{His117}$ respectively in the case that the histidine ligands are charged neutrally, further calculations were performed with no charge backdonation from the histidine ligands to the copper atom. The GROMOS96 force field sets the charge for a carbonyl group of –0.38e on the oxygen atom and +0.38e on the carbon atom, a methionine side-chain is not charged at all. The effect of a different polarisation for the axial ligands on the energy minimum is shown in

Table 2.4. The energy difference with respect to the minimum energy versus the z-co-ordinate is plotted in Figure 2.4. Herein, it is shown that there are two (local) energy minima, one on each side of the N_2S plane. The relative energy difference between the minima is small, <5 kJ.mol^{-1}, and slightly varies with the polarisation of the axial ligands and the x- and z-position of the energy minimum, see Table 2.4. At higher polarisations, the local minimum above the N_2S-plane shifts in both x- as well as z-direction, towards the methionine ligand, while the minimum below the N_2S-plane remains fixed at 0.6 Å. It is not clear why the minimum does not shift in the y-direction as both axial ligands are situated on the positive y-axis. Perhaps the Coulomb interaction of the Cu atom with the

Table 2.3: Location of the energy minimum with respect to the origin (Å) and harmonic fit of the energy along the x, y and z axis at the point of the energy minimum (10 kJ.mol^{-1}.Å^{-2}) for different charge transfers from Cu to the histidine residue

Charge		Position of energy minimum			k_{harm}		
Cu	His	x	y	z	x	y	z
+1.3	+0.0	-0.05	0.00	+0.3	73	89	9
+1.1	+0.1	-0.05	0.00	+0.3	66	84	8
+0.9	+0.2	-0.05	0.00	+0.3	67	85	8
+0.7	+0.3	-0.05	0.00	+0.4	62	76	7
+0.5	+0.4	-0.05	0.00	+0.4	66	83	6
+0.3	+0.5	-0.05	0.00	+0.5	55	68	4

Table 2.4: Location of the energy minimum with respect to the origin (Å) and harmonic fit of the energy in the x, y and z direction at the point of the energy minimum (10 kJ.mol^{-1}.Å^{-2}) for different amounts of polarisation of the axial ligands (δ is explained in Figure 2.1)

Polarisation		Position of energy minimum			k_{harm}		
Cu	δ	x	y	z	x	y	z
+1.3	0.2	-0.05	0.00	+0.3	73	89	9
+1.3	0.3	0.00	0.00	+0.3	73	108	8*
+1.3	0.4	+0.05	0.00	+0.5	57	105	8*
+1.3	0.5	+0.15	0.00	+0.7	41	101	12*
+1.3	0.6	+0.15	0.00	+0.7	44	103	15*

*R^2<0.99

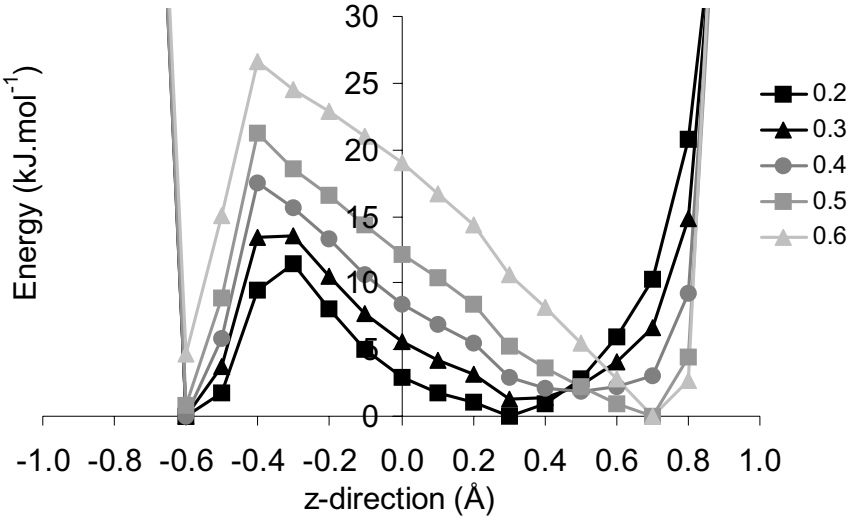


Figure 2.4: Effect of polarisation (δ in Figure 2.1) of the axial ligands on the energy map of the copper ion. The energy minima were set to zero kJ.mol^{-1}.

$S\delta_{Met}$ and O_{Gly} is too small compared to the LJ interaction with the histidine residues tobe of any influence. The movement of the minimum above the N_2S-plane in the x- and z-direction has the effect that the $k_{harm,x}$ decreases and $k_{harm,y}$ increases. In Figure 2.4 can be seen that the energy curve in the z-direction does not behave harmonically. The energy shows an increase of steepness around the minima at higher polarisations, most likely due to an increase of LJ interaction with the axial ligands. The energy barrier between the two minima increases from 10 kJ.mol^{-1} at a polarisation of 0.2e to 25 kJ.mol^{-1} for a polarisation of 0.6e. A polarisation of 0.2e provides the best results for the distance between the optimal Cu position and N_2S plane, though +0.3 Å is still too large compared to the crystallographic structure where the Cu is located at +0.09 ± 0.05 Å.

The aim is to achieve a force field for type-I Cu-containing proteins which will not fix the Cu atom into the site using bonds or constraints. Angular terms, however, will not completely strain the site and the effect of adding dihedral angles to the force field was tested by calculating a potential energy surface. The aromatic ring system of the histidine prefers a planar conformation and since it is known from quantum chemistry that the copper is part of the aromatic system, an improper dihedral angle was defined to keep

the Cu in one plane with both histidine rings. Table 2.5 shows the results on the position of the energy minimum in the grid and the harmonic fittings of the energy curves around the minimum. The addition of the dihedral angle in the force field does not affect the position of the energy minimum to a great deal. The k_{harm}'s in both the x- and y-direction are not significantly affected by the force constant, it changes only when the position of the energy minimum changes. $k_{harm,z}$ increases slightly with the force constant of the introduced dihedral angle interaction term, only in the case of an extreme value of 1000 $kJ.mol^{-1}.rad^{-2}$ the values differ significantly. Plotting the energy in the z-direction, however, clearly shows the effect of the interaction term on the local minimum below the N_2S-plane, see Figure 2.5. By introducing the two improper dihedral angles, the minimum above the plane is favoured. This is mainly due to the direction of the His46-ring which points to the positive z- axis, see Figure 1.1a. At a force constant of 100 $kJ.mol^{-1}.rad^{-2}$ the energy of the minimum below the plane has increased with more than 10 $kJ^{-1}.mol^{-1}$. Clearly it is seen that the dihedral angle has a small effect on the minimum above the N_2S-plane.

The effect on the potential energy map of adding L-Cu-L' and C-L-Cu angle bend terms to the force field was also investigated. The GROMOS96 ff uses a value of 405 and 575 $kJ.mol^{-1}$ to describe the C-S-H respectively C-N-H bending interaction. Various force constant were used for all eight angles wherein the Cu was involved. A force constant of

Table 2.5: Location of the energy minimum with respect to the origin (Å) and harmonic fit of the energy in the x, y and z direction at the point of the energy minimum (10 $kJ.mol^{-1}.Å^{-2}$) for different force constant of the $N\delta_{His}$-$C\gamma_{His}$-$C\varepsilon_{His}$-Cu improper dihedral, force constant in $kJ.mol^{-1}.rad^{-2}$. The force constant of the $N\delta_{His}$-$C\gamma_{His}$-$C\varepsilon_{His}$-H improper dihedral in the GROMOS96 FF is 167 $kJ.mol^{-1}.rad^{-2}$.

force constant	Position of energy minimum			k_{harm}		
	x	y	z	x	y	z
1000	+0.05	0.00	+0.3	97	148	34
300	0.00	0.00	+0.3	79	109	17
200	0.00	0.00	+0.3	77	109	14
100	0.00	0.00	+0.3	75	108	12
0	-0.05	0.00	+0.3	73	89	9

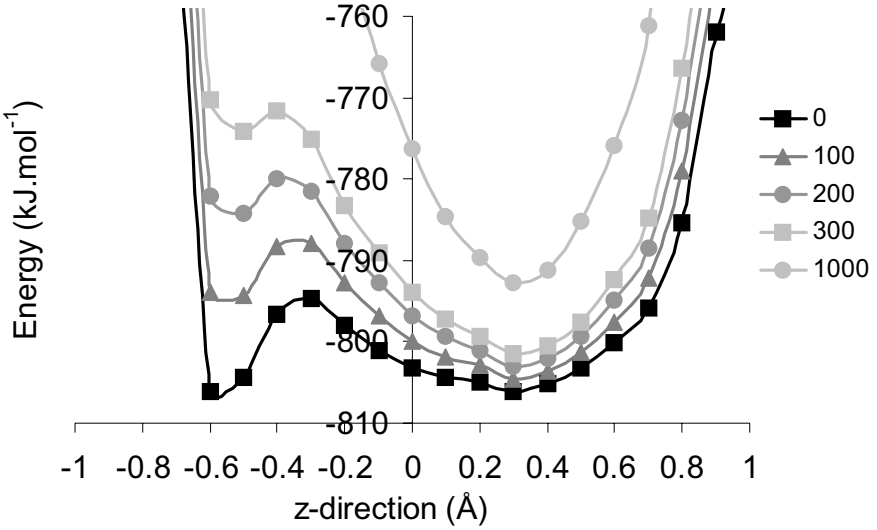


Figure 2.5: Effect of variation of the force constant for the $N\delta_{His}$-$C\gamma_{His}$-$C\varepsilon_{His}$-Cu improper dihedral on the energy map of the Cu-ion.

500 kJ.mol^{-1} has a minimal effect, the optimal position for the Cu did not change significantly and the k_{harm} increased for a maximum of 20% in the x- and y-direction and 50% in the z-direction but still resulting in very low values around 13 kJ.mol^{-1}.Å^{-2}.

2.3.2 Force field testing

A non-bonded force field proposal was obtained looking at the effects on the energy surface while varying the force field parameters. This resulted in a force field where a charge of 0.7e was delocalised to the unprotonated cysteine residue. The histidine residues were charged neutrally, leaving the Cu with a charge of +1.3e. The best results were obtained when the glycine and methionine residues were 0.2e polarised. Since the effect of the angular terms is minimal, none of these interaction terms were added to the force field. At a particular temperature every atom has a thermal energy of 1/2 kT per degree of freedom. MD-simulations were performed to check whether the type-I Cu-site is stable using these force field parameters.

Azurin

The Cu-site was stable during the 100 ps MD-simulation, however, the average Cu-position was situated too far from the methionine residue and very near the glycine residue. Therefore, the polarisation of the axial ligands was finetuned by increasing the polarisation of the methione residue to δ=0.25e in Figure 2.1. The average distance using this force field is shown in Table 2.6. The Cu maintains a high flexibility especially in the direction of the axial ligands, as shown in the standard deviation. The interaction of Cu with the equatorial ligands is much more rigid than with the axial ligands which agrees with the obtained k_{harm} as presented above. The Cu occasionally crosses the energy barrier at the N_2S-plane and occupies the local minimum on the other side of the plane.

Table 2.6: Cu-ligand distances (Å) for azurin in the XRD versus average distance during a 100 ps MD run including standard deviation.

	XRD	MD	Std. Dev.
r45	2.97	2.85	0.3
r46	2.08	2.01	0.05
r112	2.24	2.26	0.09
r117	2.01	2.01	0.05
r121	3.15	2.59	0.5
N_2S	0.08	0.31	0.2

Type-I Cu-containing proteins

The type-I Cu-site in azurin is stable, however flexible. The next step was to test whether the force field is also suitable to describe other cupredoxins. Azurin is the only known protein that contains a type-I Cu-site to have the fifth, glycine residue co-ordinating the copper. In the other cupredoxins the polarisation of this residue (δ=0.2 in Figure 2.1) was copied to the carbonyl of the residue preceeding the N-terminal histidine, to prevent Cu-co-ordination by this carbonyl resulting in the same co-ordination sphere as for azurin. A 100 ps MD run was performed and the average distances calculated, see Table 2.7. Most important, the Cu-sites were stable in all proteins. The Cu-N_{His} distances are well reproduced for every cupredoxin. The XRD distances between the Cu-atom and

ordinates the Cu atom, see Table 2.7. The Cu-Sδ_{Met} distance is harder to model since the Sγ_{Cys} are in most selected cupredoxins smaller than azurin. However the force field, which is based on azurin, models the same, large Cu-Sγ_{Cys} for all cupredoxins. In case of plastocyanin the carbonyl of the residue preceeding the N-terminal histidine ligand co-force in this direction is not very strong ($k_{harm,z} << k_{harm,x}$, $k_{harm,y}$). In general, the modelled Cu-methionine distances are too small, only in the case of the outliers, azurin and stellacyanin, the distance is modelled fairly well. This is also shown in the Cu-N_2S distance. Except for plastocyanin, the Cu is situated far above the plane reducing the distance with the methionine ligand.

Using this universal type-I Cu force field, differences arising from the different protein matrices become visible. Since the modelled structure will only differ due to forces from

Table 2.7: XRD and modelled distances for several blue copper proteins.

	plastocyanin		Amicyanin		auracyanin		stellacyanin	
	XRD	MD	XRD	MD	XRD	MD	XRD	MD
His-1	3.51	2.63	3.92	3.93	3.52	3.97	3.98	4.66
His	1.96	2.03	1.95	2.00	2.01	2.02	1.96	2.04
Cys	2.18	2.29	2.11	2.47	2.19	2.27	2.18	2.26
His	2.12	2.02	2.04	2.03	2.03	2.06	2.04	2.03
Met	2.73	2.55	2.90	2.34	2.84	2.35	2.21 *	2.07 *
N_2S	0.48	0.19	0.30	0.70	0.24	0.61	0.32	0.77

	pseudoazurin		CBProtein		NiR, typeI		rusticyanin	
	XRD	MD	XRD	MD	XRD	MD	XRD	MD
His-1	3.75	3.21	3.85	3.22	4.28	4.31	5.85	6.09
His	2.10	2.01	1.93	2.01	2.11	2.02	2.04	2.05
Cys	2.15	2.24	2.16	2.28	2.11	2.28	2.26	2.24
His	2.15	2.01	1.95	2.01	1.97	2.03	1.89	2.03
Met	2.76	2.37	2.61	2.37	2.63	2.42	2.89	2.38
N_2S	0.48	0.44	0.39	0.42	0.53	0.67	0.33	0.74

* Cu-Gln distance

the protein matrix and not from quantum effects, the effect of the protein can be enters the site quickly. In the other mutants, the axial ligands are modelled fairly well considering the high fluctuations it gave for the earlier described proteins.monitored. Only in the case of azurin and plastocyanin the residue preceeding the N-terminal histidine is able to co-ordinate the Cu using the same parameters for the Cu site. In these cases the protein might force this residue towards the Cu-site. Furthermore, the modelled distance between Cu and $S\gamma_{Cys}$ is in all the cases equal to azurin. The observed, smaller Cu-$S\gamma_{Cys}$ distances in the other cupredoxins are most likely due to quantum effects.

Non type-I Cu-containing proteins

Cu-sites show a large plasticity. The introduction of a mutation can therefore have a large effect on the Cu-site. The developed force field was tested on non-type-I Cu-sites to determine whether the developed ff is suitable to model a larger variation in the Cu-sites. A major advantage of a non-bonded ff is that there is no need to define the ligands nor bonding properties like the equilibrium distance.

The first test-case is the homo-trimer nitrite reductase which contains two Cu-sites in each monomer: a type-I Cu-site, which was modelled as described above, and a type-II Cu-site (see Figure 1.1b). For the type-II Cu-site the same atomic charges were used as in the type-I Cu-site: a charge of +1.3e was placed on the Cu-atom, the three histidine ligands (singly protonated at the Nδ–atom) have total charge of 0e. The additional Cu-co-ordinating water molecule was modelled by SPC (Berendsen *et al.*, 1981). The result

Table 2.8: Cu-ligand distances (Å) for for the type-II site of nitrite reductase in the XRD versus average distance during a 100 ps MD run including standard deviation.

	NiR, typeII		
	XRD	MD	Std. Dev.
His	2.02	2.04	0.06
His	2.13	2.06	0.05
His	2.16	2.15	0.11
H_2O	2.14	2.52	0.10
N_3	0.99	0.86	0.19

is a total non-integer charge of +1.3 on the Cu-site. Table 2.8 shows the average Cu-ligand distances with standard deviation of the type-II Cu-site of NiR during a 100 ps MD-simulation. The Cu-Nε_{His} distances are modelled very well. Even the small differences in Cu-Nε_{His} distances are reproduced where as in all previous simulations the Cu-Nδ_{His} distances are equal within 0.01Å. The distance between Cu and the plane spanned by the three Nε-ligands (N_3) is modelled larger than for the type-I Cu-sites, in accord with the XRD data. Furthermore, four azurin mutants where the Met 121 was replaced by other residues were tested as well. The introduced residues (Glu, His, Ala and Gln) all distort the type-I Cu-site. The charge of the replaced residues was taken from the developed non-bonding force field. The introduced histidine 121 residue (singly protonated at the Nε atom) was given the same charge distribution as the other two co-ordinating histidine residues. The introduced alanine or glutamine were both polarised with 0.25e on the Cβ and Oε1 respectively and -0.25e on the Cα and Cδ. The glutamate residue has a total charge of -1e and its distribution was copied from the standard GROMOS96 force field. The Cu-sites of the mutants are stable in all simulations. The Cu-Sγ_{Cys} distances in the azurin mutants are modelled too large just as in the case of all other protein s except *wt* azurin, see Table 2.9. The Cu-Nδ_{His} distances are reproduced correctly, only in the case of the M121A mutant the XRD measurements show an enlargement of the Cu-Nδ_{His} distance which is not modelled. Further, the Cu-O_{Gly} distance becomes too large to co-ordinate the Cu-atom which is mainly due to the intrusion of a solvent molecule entering the Cu-site. The water molecule replaces the Ala

Table 2.9: XRD and modelled distances for azurin mutants.

	M121Q		M121H		M121A		M121E	
	XRD	MD	XRD	MD	XRD	MD	XRD	MD
Gly_{45}	3.37	3.73	3.89	3.65	2.74	3.71	3.17	3.33
His_{46}	1.93	2.05	2.01	2.06	2.20	2.04	2.10	2.02
Cys_{112}	2.12	2.30	2.16	2.26	2.12	2.29	2.08	2.34
His_{117}	2.05	2.04	2.14	2.05	2.37	2.04	2.01	2.04
X_{121}	2.26	2.57	2.28	2.07	4.69	4.76	1.97	2.44
N_2S	0.26	0.54	0.57	0.78	0.08	0.71	0.27	0.58

121 position and pulls the Cu away from the glycine far away from the N_2S-plane. Upon manually removing this water molecule from the Cu-site an other solvent molecule

2.4 Comparison of force fields

2.4.1 Azurin

Swart et al. have developed a force field for the type-I site of azurin based on quantum chemical calculations (Swart *et al.*, in preparation). The charge distribution was calculated using Density Functional Theory. The Cu-site was described using five covalent bonds, see Table 2.9. Force constants were developed directly from the Hessian matrix (a matrix of second derivatives of the energy with respect to atomic coordinates) as obtained in quantum chemical calculations. No (dihedral) angle terms were used to describe the Cu-site. The non-bonded interactions between Cu and its ligands were excluded. Using this force field the potential energy surface of the Cu-ion was calculated as described above. Table 2.10 shows the comparison of this force field with the non-

Table 2.10: Cu-ligand force constants ($kJ.mol^{-1}.nm^{-4}$) as used in literature for azurin and plastocyanin (PC) converted to a quartic function by $k_q=k_h/2r_0^2$ where r_o is the equilibrium distance.

	Swart (azurin)	De Kerpel (PC)	Comba (PC)
Gly	73	-	-
His	569	794	1286
Cys	688	1145	420
His	950	676	1287
Met	77	48	86

Table 2.11: Comparison of different Cu force fields applied to azurin.

ff	Position of energy minimum			k_{harm}		
	x	y	z	x	y	z
XRD	-0.08	+0.05	+0.04	-	-	-
Empirical	-0.05	0.00	+0.3	73	89	9
Swart	-0.10	+0.05	+0.1	125	89	36

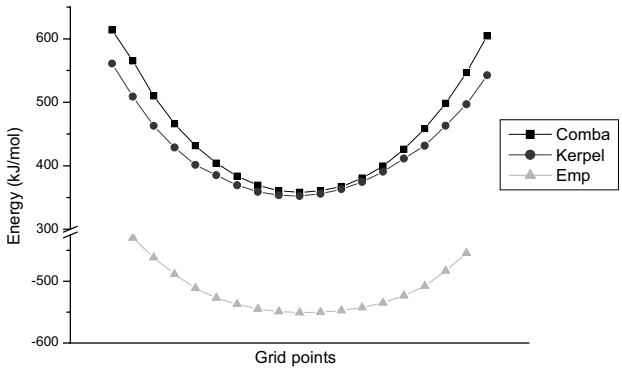
bonded force field as described above. It shows that both force fields correctly position the energy minimum in the x- and y-direction. In the z-direction the energy minimum for the force field developed by Swart et al. is more located towards the N_2S-plane. The $k_{harm,Swart}$ in the x- and z-direction are larger than the ones obtained using the non-bonded force field most likely due to the introduction of covalent bonds. It is therefore remarkable that no difference is observed for the $k_{harm,y}$.

2.4.2 Plastocyanin

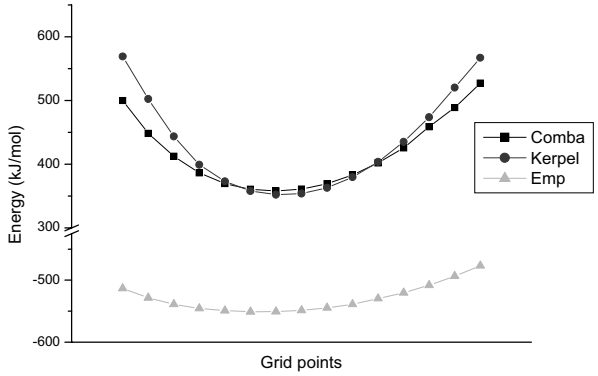
More force fields have been developed for other cupredoxins. de Kerpel et al. and Comba et al. both have used quantum chemical calculations to develop a force field for plastocyanin (Comba and Remenyi, 2002; de Kerpel and Ryde, 1999). The charge distribution was calculated using DFT. Both have used four Cu-ligands, where the force constant were developed using the second derivative of the Hessian (de Kerpel) and a quantum chemically derived energy surface (Comba). Additionally both have added 13 angular terms involving the Cu-atom with force constant varying from 20 to 530 $kJ.mol^{-1}$. No exclusions were defined for the Cu-ligand non-bonding interactions. The potential energy surface of the type-I Cu-site of plastocyanin is calculated using the same method as in azurin. The system consisted of the sidechains of the four Cu-co-ordinating residues: His39, Cys89, His92 and Met97. The origin is defined by the mathematical average of the $S\gamma_{Cys}$ and both $N\delta_{His}$ atomic co-ordinates. The position of the Cu-atom is varied and the potential energy of the system is calculated using the non-bonded force field and the force fields developed by de Kerpel et al. and Comba et al. Figure 2.6 displays the change in energy upon a variation of the Cu atom in the x- (a) and y direction (b) in plastocyanin for the three force fields. It is clearly shown that the energy

Table 2.12: Comparison of different Cu force fields applied to plastocyanin.

force field	Position of energy minimum			k_{harm}		
	x	y	z	x	y	z
XRD	-0.13	-0.06	+0.48	-	-	-
Empirical	-0.10	0.00	+0.6	87	96	15
Comba	-0.10	0.00	+0.6	250	229	45
de Kerpel	-0.15	+0.05	+0.5	344	173	49



a



b

Figure 2.6: Comparison of different force fields on the mobility in (a) the x- and (b) the y-direction of the Cu atom in plastocyanin. The resolution is 0.05Å per grid point.

profile for the QM force field is steeper than for the empircal force field as also shown in Table 2.11. The position of the minimum energy varies only slightly for the diverse force fields, whereas the k_{harm} of the QM force fields show an increase for all directions. Also compared to the QM force field in azurin by Swart the k_{harm} in the x- as well as y-direction are very large. Most likely this is due to the higher Cu-ligand force constant used in the force field by de Kerpel and Comba as well as the maintained non-bonding Cu-ligand interactions. Also the addition of 13 angular terms might have a small effect.

2.5 Concluding remarks

A potential energy surface is used as a relatively simple method to extract the ideal ff parameters to describe the Cu-site. Properties as observed using the surface comply with the MD-simulations. The optimal copper position is in accord with the average distances and the steepness of the energy curve when the Cu-atom was positioned around the optimal energy value relates to the mobility in the simulations.

The MD-simulations reproduce this observation by showing that the distance between Cu and the equatorial ligands has a low standard deviation and is fairly constant over all the modelled proteins. The distances between the Cu and the equatorial ligands were modelled throughout all proteins with a very small differentiation whereas the distances to axial ligands show a much higher variation. Also the two positions of the Cu atom situated above and below the N_2S-plane found in the energy surface were observed in the MD-simulations. Furthermore, the potential energy surface is able to quickly compare different force fields analysing for example the influence of bonding terms without having to perform dynamics simulations.

The developed non-bonding Cu force field using a potential energy surface is stable in the dynamics simulations and is also applicable to other cupredoxins and mutated Cu-sites. This opens the possibility to use the force field for structure and stability prediction of Cu-sites of which the structure is not available.