Modelling copper-containing proteins

Bosch, Marieke van den

Chapter 3

A TEST CASE: N42C/H117G AZURIN MUTANT

Summary

Because the stability of the yellow species of H117G/N42C azurin is limited, structure determination with regular methods is prohibited. To obtain insight in the possible structure of this form of the protein, a model was built using molecular dynamics calculations. The calculations were not meant to simulate the formation of or to establish unequivocally the structure of the copper site. The purpose was to see whether a strain-free model of the copper site might be obtained with the most likely set of ligands and what rearrangement of the protein might be needed.

3.1 Introduction

Research on azurin has a long history and the properties of this protein have been investigated abundantly. It has been used many times for mutational studies and served as a model protein to test new techniques, experimentally as well as computationally, as, for instance, in the case of the double mutant Asn42Cys/His117Gly azurin. The His117 is a Cu-co-ordinating residue in the actoive site of the protein. In the His117Gly variant it was replaced by a Gly residue to create a gap in the protein surface by which the copper-containing active center becomes accessible for external ligands (den Blaauwen and Canters, 1993; den Blaauwen *et al.*, 1993; den Blaauwen *et al.*, 1991; Farver *et al.*, 2000; Jeuken *et al.*, 2000a; Jeuken *et al.*, 2000b). Type-I character is restored when monodentate ligands such as imidazole, pyridine (derivatives), halides or azide are added. The site obtains type-II character when bidentate ligands such as histidine and histamine are added. The additional N42C mutation was constructed in an attempt to investigate the electron self-exchange properties of azurin dimers (van Amsterdam *et al.*, 2003; van Amsterdam *et al.*, 2002a). The properties of this azurin variant in its monomeric form turned out to be quite unexpected (van Amsterdam *et al.*, 2002b) .

3.1.1 UV-visible Spectroscopy

When Cu(II) was added to a (colourless) solution containing H117G apoazurin (single mutation), the solution turns green, whereas the native protein is blue (den Blaauwen and Canters, 1993; den Blaauwen *et al.*, 1993; den Blaauwen *et al.*, 1991; Jeuken *et al.*, 2000b). The introduction of the (additional) N42C mutation in the H117G azurin mutant changes the absorption spectrum again, i.e., the solution becomes yellow upon adding Cu. In contrast to Cu-H117G azurin, addition of external ligands such as imidazole has no effect on the optical spectrum of Cu- N42C/H117G azurin. Furthermore, it turns out that blocking Cys 42 restores the H117G characteristics.

3.1.2 Resonance Raman Spectroscopy

The resonance Raman (RR) spectrum of the yellow N42C/H117G mutant is similar to that of other tetragonal type-II copper sites with copper cysteinates co-ordination, such as the His117→Gly azurin and superoxide dismutase mutants, in that it exhibits a

number of fundamental frequencies around 400 cm^{-1} (Andrew *et al.*, 1997; Andrew *et al.*, 1994; den Blaauwen *et al.*, 1993; van Pouderoyen *et al.*, 1996). Combination bands, however, are not observed for the single thiolate tetragonal sites of azurin and superoxide dismutase while they do occur in the RR spectrum of the Cu- loaded N42C/H117G variant. As in a superoxide dismutase variant, the combination bands are indicative of two cysteinate ligands (Liu *et al.*, 2000). Such combination bands are observed in copper-substituted liver alcohol dehydrogenase (Maret *et al.*, 1986) and other metal cysteinate proteins with multiple cysteinate ligands such as CuA (Alvarez *et al.*, 2001; Andrew *et al.*, 1996), rubredoxin (Czernuszewicz *et al.*, 1986), and ferredoxin (Han *et al.*, 1989) as well as in the copper-sulfide cluster Z_{ox} in N_2O reductase (Alvarez *et al.*, 2001) .

It was concluded that the UV-visible spectra are compatible with Cys 42 acting as one of the copper ligands in a tetragonal type-II copper site. Resonance Raman spectroscopy provides a strong indication that two cysteine residues are involved, which makes Cys 112 a second ligand. Cys 42 is in fairly close proximity to the gap created by the removal of the His117 side chain (Hammann *et al.*, 1997; Jeuken *et al.*, 2000a). Apparently, the loop containing residues 36–47 is flexible enough to enable Cys 42 to move toward the cavity where it can act as a copper ligand. Since the protein is not stable enough to permit NMR analysis and does not crystallise, the structure can not be experimentally obtained. MD-simulations were performed to investigate the Cu-co-ordination sphere.The loop was pulled inwards by gradually shortening the distance between the Cu and the $S\gamma_{Cys42}$ and the stability of the obtained Cu-site was checked using an empirical, non-bonded force field.

3.2 Methods

Modelling of the structure of the N42C/H117G azurin double mutant was done in three steps. First, the 3D-structure of wild type azurin was used as a starting point. His117 was replaced by a glycine residue and Asn42 by a cysteine residue using Swiss-Pdb Viewer version 3.51.2 (Guex and Peitsch, 1997). Using the GroMaCS simulation package (van der Spoel *et al.*, 1999b) and GROMOS96 43a2 force field (van Gunsteren *et al.*, 1996) the protein was placed in the centre of a cubic box together with 3240 SPC water

molecules (Berendsen *et al.*, 1981). The system was relaxed using a steepest descent algorithm.

The second step was to move the loop containing the Cys 42 inward using Molecular Dynamics (MD) simulations. In the dynamics simulations the protein and solvent were separately coupled to a temperature bath at 300 K using the weak coupling scheme (Berendsen *et al.*, 1984) with a relaxation time of 0.1 ps. An isotropic constant pressure was applied using a relaxation time of 1 ps. The bond lengths within the protein and the water geometry are constrained using the LINCS (Hess *et al.*, 1997) and SETTLE (Miyamoto and Kollman, 1992) algorithms respectively, allowing a timestep of 2 fs. A group-based twin range cut-off scheme was employed for the non-bonded interactions, with a short-range cutoff of 0.8 nm and a long-range cut-off of 1.4 nm, while the pairlist was updated every 5 steps. A Poisson-Boltzmann reaction field was used to approximate the electrostatic interactions beyond the long-range cut-off. The value for the dielectric permittivity of the continuum outside the long-range cut-off was set to 66 (Heinz *et al.*, 2001).

Several simulations under different conditions were performed to stepwise shorten the distance between the Cys 42 and the Cu. In the first case, a Cu-atom was constrained to $S\gamma_{Cys42}$ at 2.2 Å and the Cu was then pulled over towards the original binding site by restraining it to a dummy atom, which was placed at the average position between the two ligands: $S\gamma_{Cys112}$ and $N\delta_{His46}$. The distance restraint between copper and the dummy atom was shortened every 40 ps. During this procedure the residues Cys 112 and His 46 are not restrained and it was found that they exhibit a high mobility as long as the copper atom has not moved to its final position. Therefore, secondly, another methodology was applied by which a (harmonic) distance restraint was defined between the $S\gamma_{Cys42}$ and the Cu atom. The latter was constrained to $N\delta_{His46}$, $S\gamma_{Cys112,}$ and $S\delta_{Met121}$ at a distance of 2.0, 2.2 and 3.1 Å, respectively. Every 100 ps the distance restraint was shortened by 1 Å, starting at 9Å (simulation 1). When using the same force constant, 500 $kJ.mol^{-1}.nm^{-2}$, as used in the GROMOS96 force field for Fe-N bonds, this means that in every step 2.5 $kJ.mol^{-1}$ was added to the restraint potential. Once the distance restraint had reached a value of 2 Å, the distance restraint was kept constant and the force constant was increased by 2.5 kJ/mol^{-1} every 100 ps (simulation 2A). It appeared that even at relatively high force constants, Cys 42 was not easily pulled into the Cu-site. Therefore,

starting with the output of simulation 1, the constraint between the Cu and the methionine residue (simulation 2B) or both the Cu-Sδ_{Met121} and Cu-Nδ_{His46} constraints were lifted (simulation 3C). These simulations were stopped when the Cu- Sγ_{Cys42} distance was smaller than 0.25 nm.

Table 3.1: Overview of the applied Cu-Sγ_{Cys42} restraint during the simulations: force field parameters (r_0 and k) and actual distance (r(t)) and force (F(t)) at the start and end of every simulation.

Sim	Time (ns)		r_0 (nm)	k ($kJmol^{-1}nm^{-2}$)	r(t) (nm)		F(t)=k(r-r_0) ($kJmol^{-1}nm^{-1}$)	
	start	end			start	end	start	end
1	0.0	0.1	0.9	500	1.248	1.153	174	127
	0.1	0.2	0.8	500	1.153	1.005	177	103
	0.2	0.3	0.7	500	1.005	0.950	153	125
	0.3	0.4	0.6	500	0.950	0.981	175	191
	0.4	0.5	0.5	500	0.981	0.901	241	201
	0.5	0.6	0.4	500	0.901	0.865	251	233
	0.6	0.7	0.3	500	0.865	0.863	283	282
	0.7	0.8	0.2	500	0.863	0.569	332	185
2A	0.8	0.9	0.2	532	0.569	0.663	196	246
	0.9	1.0	0.2	555	0.663	0.516	257	175
	1.0	1.1	0.2	605	0.516	0.551	191	212
	1.1	1.2	0.2	645	0.551	0.548	226	224
	1.2	1.3	0.2	690	0.548	0.475	240	190
	1.3	1.4	0.2	755	0.475	0.466	208	201
	1.4	1.5	0.2	825	0.466	0.356	219	129
	1.5	1.6	0.2	1030	0.356	0.460	161	268
	1.6	1.7	0.2	1100	0.460	0.330	286	143
	1.7	1.8	0.2	1395	0.330	0.298	181	137
2B	0.8	0.9	0.2	500	0.569	0.462	185	131
	0.9	1.0	0.2	575	0.462	0.503	151	174
	1.0	1.1	0.2	625	0.503	0.471	189	169
	1.1	1.2	0.2	725	0.471	0.297	196	70
	1.2	1.3	0.2	1250	0.297	0.225	121	31
2C	0.8	0.9	0.2	500	0.569	0.303	185	52
	0.9	1.0	0.2	980	0.303	0.248	101	47
	1.0	1.1	0.2	3150	0.248	0.228	151	88

The final step was to check whether the copper sites obtained from simulations 2A, 2B and 2C are stable during the MD-simulations by applying a non-bonded force field (see chapter 2). It gives both the Cu as well as the ligands the opportunity to return to their initial positions. The Cu charge was herein +1.3e and both cysteine residues had a total charge of –0.3e, distributed over the Sγ (-0.24e) and Cβ (-0.06e). The His 46 residue was singly protonated at the Nε atom and was therefore neutral. The
charge distribution was taken from the GROMOS96 force field (Schuler and van Gunsteren, 2000; van Gunsteren *et al.*, 1996), where the ligand atom Nδ carries a charge of –0.58e. The side chain of the Met 121 is polarised: -0.25e on the sulphur atom and +0.125e on the adjacent carbon atoms. The polarisation of the carbonyl group of Gly 45 is ±0.20e. Two simulations of 0.4 ns were performed (simulations 3A, 3B and 3C), starting from the output structures of simulations 2A, 2B and 2C respectively.

3.3 Results and discussion

3.3.1 Structure

Figure 3.1a shows the input structure of simulation 1: the optimised azurin double mutant structure after replacement of Asn 42 and His 117 by a cysteine and glycine residue respectively in *wt* azurin. The $S\gamma_{Cys42}$ distance with respect to the Cu atom at this point is 1.248 nm. Using a harmonic force, the loop containing the Cys 42 residue was pulled inward by stepwise decreasing the equilibrium distance between Cu and $S\gamma_{Cys42}$ (simulation 1) and increasing the force constant (simulation 2A), see Table 3.1. The output structures of simulations 1 and 2B are presented in Figure 3.1b and 3.1c, respectively. Figure 3.2 (black line) shows the distance between the Cu and the $S\gamma_{Cys42}$ atoms during the simulations. Pulling the loop towards the original copper-site takes a long time. Table 3.1 shows that the Cys 42 residue remains at a relatively large distance from the copper atom, even at very high force constants. It also shows that the force acting on the two atoms involved remains fairly high, indicating strain. In the output structure, Figure 3.1c, it can be seen that the methionine residue is located at the other side of the Cu-atom with respect to the Cys 42. Since the experimental data do clearly

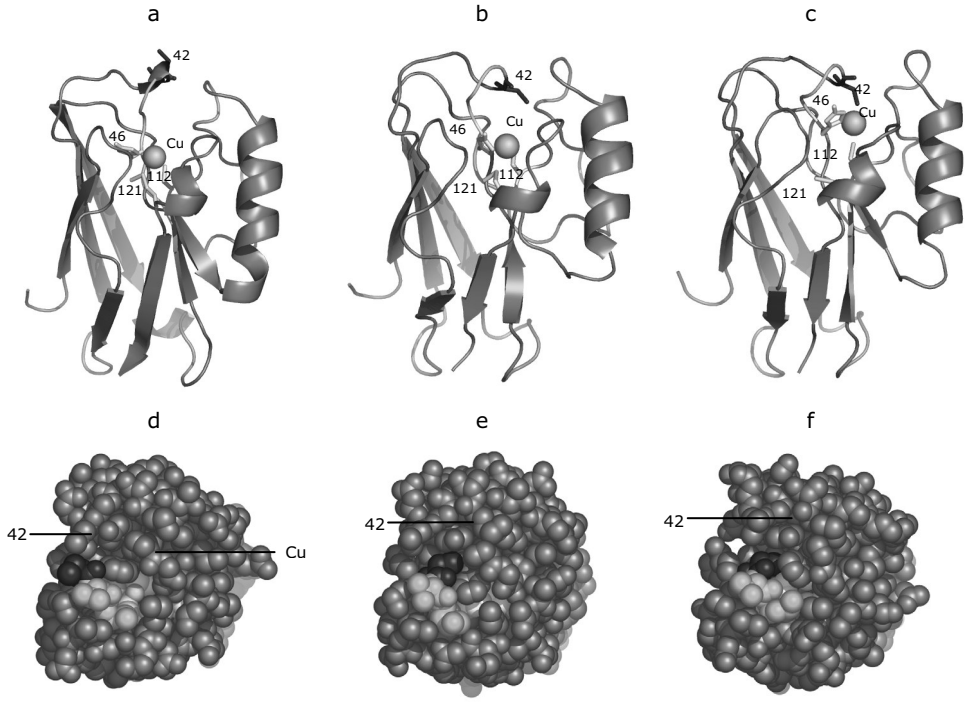

Figure 3.1: Ribbon representation of the side-view and of the input structure of simulation 1 (a), the output structure of simulation 1 (b) and the output structure of simulation 2B (c) and a sphere representation of the top view of the same three structures (d), (e) and (f) respectively.

show that the methionine co-ordinates to the Cu atom, a second simulation was performed starting from the output of simulation 1 where the constraint between Cu and the methionine was removed (simulation 2B). In the output structure, Figure 3.1d, it can be seen that now the copper atom is able to move towards the $S\gamma_{Cys42}$. Figure 3.2 confirms this by showing that the Cu- $S\gamma_{Cys42}$ distance becomes smaller at a shorter time scale (see Table 3.1).

3.3.2 Stability

The number of hydrogen bonds was monitored to check the stability of the protein. Figure 3.3 shows the number of hydrogen bonds during the three simulations between backbone atoms and within the whole protein. The number of backbone hydrogen bonds

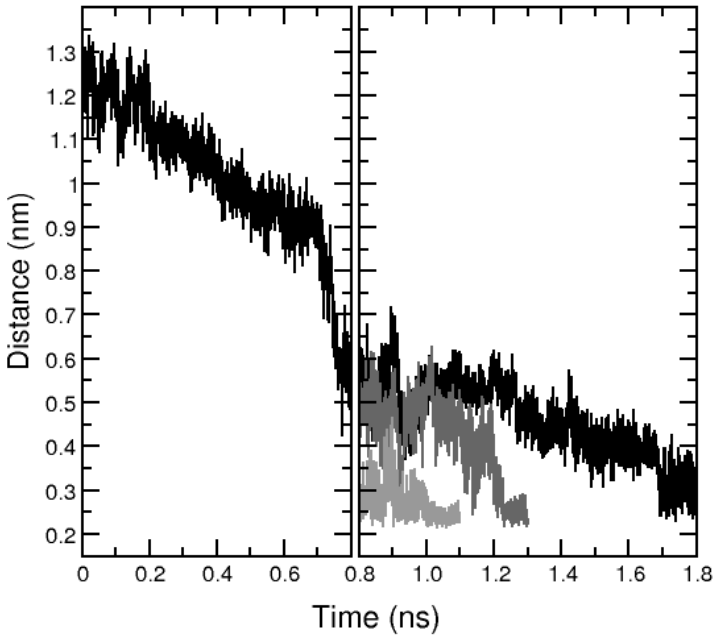

Figure 3.2: Cu-$S_{\gamma Cys42}$ distance during simulations 1 (left panel), 2A (right panel; black), 2B (right panel; dark grey) and 2C (right panel; grey)

during simulation 1A remains stable indicating that there are no forces that disrupt the overall structure of the protein. During simulation 2A, a small decrease of the number of backbone hydrogen bonds is observed. This implies that the structure of the protein has to be disrupted in order for the loop to be pulled toward the Cu-site. When the Cu is not constrained to the methionine residue, the metal is able to move towards the Cys 42 making it easier to shorten their distance. There is no need to disrupt the protein judging by the stable amount of hydrogen bonds during the simulation 2B. The Root Mean Square (RMS) deviation of the backbone of the protein with respect to the input structure shows that the backbone of the protein is stable. After 1.8 ns the deviation is below 0.17nm. Figure 3.4 shows the RMS deviation of the backbone of the 36-47 loop. It emphasizes the difference between the simulations 2A and 2B. In the case of simulation 2A where the methionine remains constrained to the Cu-atom, the loop needs to be pulled much further into the protein resulting in relaticely large RMSD values while

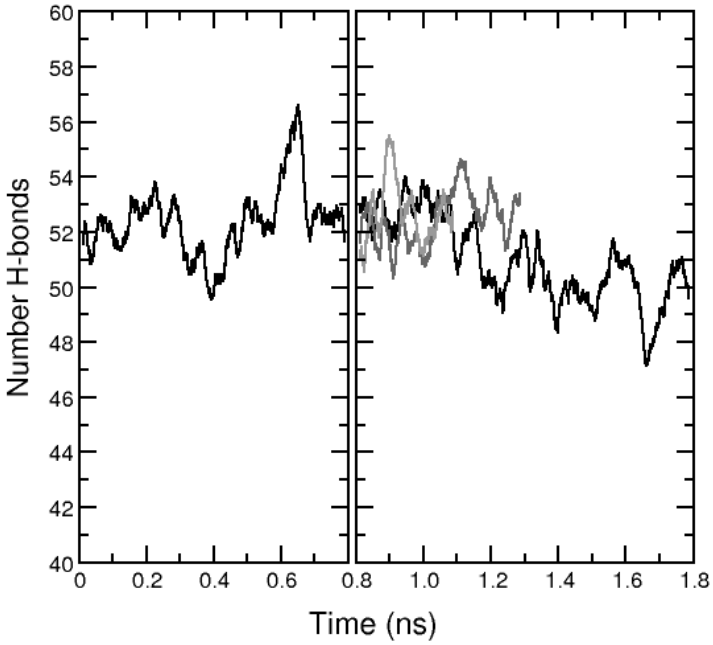

Figure 3.3: Number of NH-CO hydrogen bonds during simulations 1 (left panel), 2A (right panel; black), 2B (right panel; dark grey) and 2C (right panel; grey)

when discarding the Cu-methionine constraint, minimising of the Cu-Sγ_{Cys42} distance does not result in a large displacement of the 36-47 loop containing Cys 42.

3.3.3 Structure and stability of the Cu-site

Finally, three simulations were performed of 0.4 ns starting from the output structures of simulations 2A, 2B and 2C where the stability of the Cu-site was tested (simulations 3A, 3B and 3C, respectively). The Cu-ligand constraints and Cu- Sγ_{Cys42} restraint were removed and an empirical charge distribution was used to describe the Cu-site on a non-bonded level. Figure 3.5a, b and c show the Cu-ligand distances for the three runs. Figure 3.5a shows that the Cu-Nδ_{His46} and Cu-Sγ_{Cys112} distances are stable during the simulation 2A. The Met 121, however, immediately moves away from the Cu-site. In Figure 3.5b it is seen that when discarding the methionine residue as a possible Cu-ligand, the three remaining, proposed ligands are stable including the Cys 42 residue.

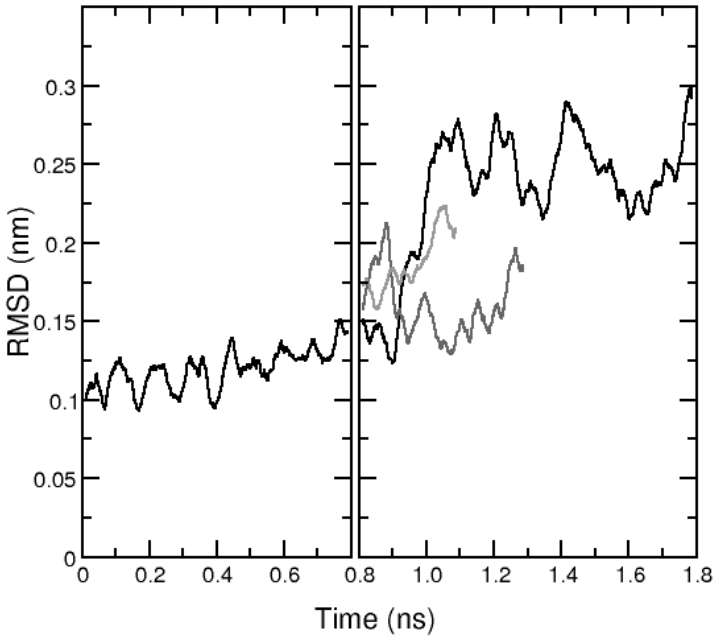

Figure 3.4: RMSD of the backbone of the loop during simulations 1 (left panel), 2A (right panel; black), 2B (right panel; dark grey) and 2C (right panel; grey). A running average of 50 points was applied on the data.

Furthermore, it can be seen that in simulation 3A the Cu- $S\gamma_{Cys42}$ distance varies much more than in simulation 3B while they are described by the same force field parameters. Especially when looking at the last 0.1 ns, it is clear that in this structure the Cys 42 is not stable as a Cu-ligand and competes with the Met 121. This is not the case in simulation 3B, where both Cu-S_{Cys} distances have a similar behaviour. When discarding the His 46 as a ligand as well, the Cu-site is not stable as seen in Figure 3.5c.

3.3.4 Model structure

The best model is obtained from simulations 2B and 3B. Although the Cu-site and the overall protein structure are stable during these simulations, there is some slight residual strain in the loop around residue 42 as residues Lys 41 and His 46 occur in a disfavored region of the Ramachandran diagram. When the His 46 is not constrained to the Cu-atom

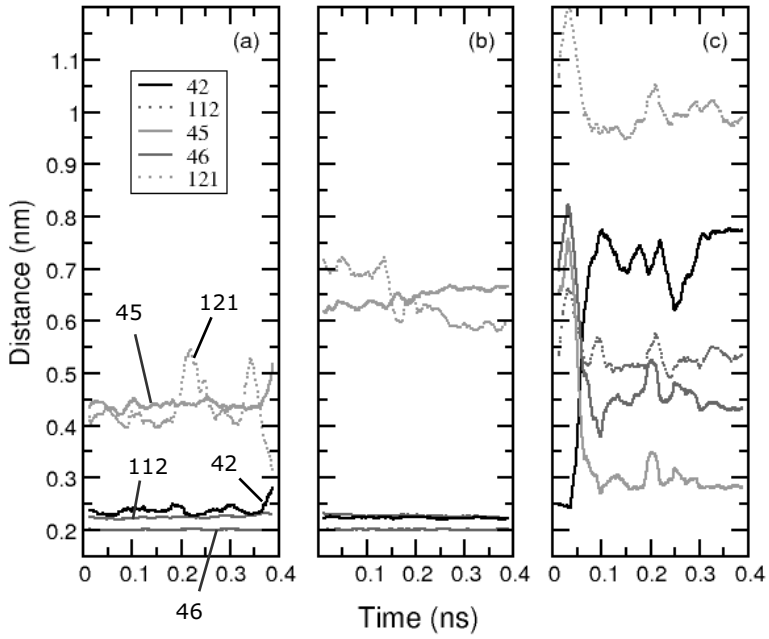

Figure 3.5: Cu-ligand distances during simulation 3A (a), 3B (b) and 3C (c). A running average of 50 points was applied on the data.

like in simulations 2C, 3C and in the procedure where the Cu was constrained to the Cys 42 and pulled over to a dummy atom between $N\delta_{His46}$ and $S\gamma_{Cys112}$ (data not shown), the histidine turns away from the original

Cu-site. When the histidine ligand is constrained to the Cu-atom while pulling the loop inward, simulation 2B, the residue is pushed into the disfavoured region of the Ramachandran diagram. The site is not stable however without this histidine ligand. It can be a reason why this mutant is less stable than the *wt* protein (van Amsterdam *et al.*, 2002b).Since it can be concluded that the Met 121 does not act as a Cu-ligand in the double mutant, a possible fourth co-ordination ligand for the copper in this site might be a water molecule. When starting the simulation a water molecule is present at 0.25 nm distance from the Cu-site, as expected from single mutant, H117G azurin studies (Jeuken *et al.*, 2000b). During simulation 1 the water molecule leaves the site and no bulk water molecule returns to the site so that in the continuation runs there is no water present near the Cu-atom (> 0.5 nm). The time-scale might be too small for water to re-enter the Cu-

site. More sampling is needed to be certain of the water behaviour with respect to the site.

3.4 Conclusions

The simulation demonstrates that it is possible, indeed, to obtain a more or less strain free form of the protein that allows for binding of at least three ligands to a single copper. The structures in Figure 3.1 clearly show that the loop (residues 36–47) has rearranged, enabling Cys 42 to participate in copper binding. The participation of Met 121 in copper binding is not very likely since the distance between $S\delta_{Met121}$ and copper is relatively large and a lot of work needs to done to adjust the protein structure to move the loop inward so far. Though positioned in a disfavoured region of the Ramachandran diagram the His 46 is very stable.