PROTEINS: Structure, Function, and Genetics 41:299–315 (2000)

Molecular-Dynamics Simulation of the β Domain of Metallothionein With a Semi-Empirical Treatment of the Metal Core

Christian D. Berweger,[1] Walter Thiel,[2] and Wilfred F. van Gunsteren[1]*
[1]*Laboratorium für Physikalische Chemie, Eidgenössische Technische Hochschule Zürich, ETH Zentrum, Zürich, Switzerland*
[2]*Max-Planck-Institut für Kohlenforschung, Mülheim, Germany*

ABSTRACT **The three-metal-containing β domain of rat liver metallothionein-2 in aqueous solution was simulated with different metal contents. The Cd_3, the $CdZn_2$, and the Zn_3 variant were investigated using a conventional molecular dynamics simulation, as well as a simulation with a semi-empirical quantum-chemical description (MNDO and MNDO/d) of the metal core embedded in a classical environment. For the purely classical simulations, the standard GROMOS96 force-field parameters were used, and parameters were estimated for cadmium. The results of both kinds of simulations were compared to each other and to the corresponding experimental X-ray crystallographic and NMR solution data. The purely classical simulations were found to produce a too compact metal cluster with partially incorrect geometries, which affected the enfolding protein backbone structure. The inclusion of MNDO/d for the treatment of the metal cluster improved the results to give correct cluster geometries and an overall protein structure in agreement with the experiment. The metal cluster and the cysteine residues bound to it are structurally stable, while the irregular polypeptide backbone loops between the cysteines exhibit a considerable flexibility. MNDO without extension to d orbitals failed to maintain the structure of the metal core. Proteins 2000;41:299–315.** © 2000 Wiley-Liss, Inc.

Key words: zinc; cadmium; metalloprotein; MNDO/d; QM/MM simulation

INTRODUCTION

The role of metals in biological systems has gained increased interest in recent years.[1] One of the most important metal ions in biological systems is zinc. Its structural function is, for example, important in zinc fingers, and it enables special catalytic reactions in enzymes such as alcohol dehydrogenase, which is involved in the degradation of ethanol. However, cadmium has similar binding properties as zinc, but does not possess the same structural functions or catalytic capabilities. This is a reason why cadmium is toxic: it competes with zinc in binding to biomolecules, but does not provide the special properties of zinc.

Metallothioneins are a class of abundant small proteins with a high content of cysteines. They are capable of binding large amounts of heavy metals such as zinc, cadmium, and mercury. Rat liver metallothionein has 61 residues, of which 20 are cysteines. Binding to copper, zinc, and cadmium seems to be biologically most important, but it is also found to bind to mercury and platinum. It binds seven heavy metal ions in two domains, which are quite independent. In the β domain, consisting of residues 1–30, three metal ions are coordinated by nine cysteines. The α domain (residues 31–61) binds four metal ions to eleven cysteines. The cysteines are deprotonated and coordinate the metals in a tetrahedral fashion, similar to the structure of zincblende. The structure of rat liver metallothionein-2 has been solved by nuclear magnetic resonance spectroscopy (NMR)[2] and X-ray crystallography.[3] A previous X-ray structure[4] was proven to be incorrect.[3,5] The NMR structure contains seven cadmium ions, whereas the X-ray structure has four cadmium ions in the α domain, and a $CdZn_2$ composition in the β domain.

The function of metallothionein is unclear. The most likely ones are supplying metal ions, primarily copper and zinc, for the biosynthesis of other metalloproteins, and removal of toxic metals such as cadmium and mercury. This detoxification function requires a broad but strong affinity for various toxic heavy metal ions. An interplay with zinc fingers, and thus gene expression, is also possible.

Proteins involving heavy metal ions are not ideally suited for investigation by molecular dynamics (MD) computer simulations. The reason is the lack of reliable force fields for these ions. The GROMOS (Groningen molecular simulation computer program package) force field,[6] for example, contains parameters for zinc, but they have never been thoroughly tested. Parameters for cadmium are not generally available. The present work compares the performance of the standard GROMOS force field, extended by estimated parameters for cadmium, to simulations with a semi-empirical quantum-mechanical treatment of the metal core embedded in a classical-mechanical environment. A review of and references to hybrid quantum/classical simulation studies can be found in Gao and Thompson.[7] Metallorganic complexes as models of metallo-

*Correspondence to: Wilfred F. van Gunsteren, Laboratorium für Physikalische Chemie, Eidgenössische Technische Hochschule Zürich, ETH Zentrum, CH-8092 Zürich, Switzerland. E-mail: wfvgn@igc.phys.chem.ethz.ch.

Received 27 January 2000; Accepted 28 June 2000

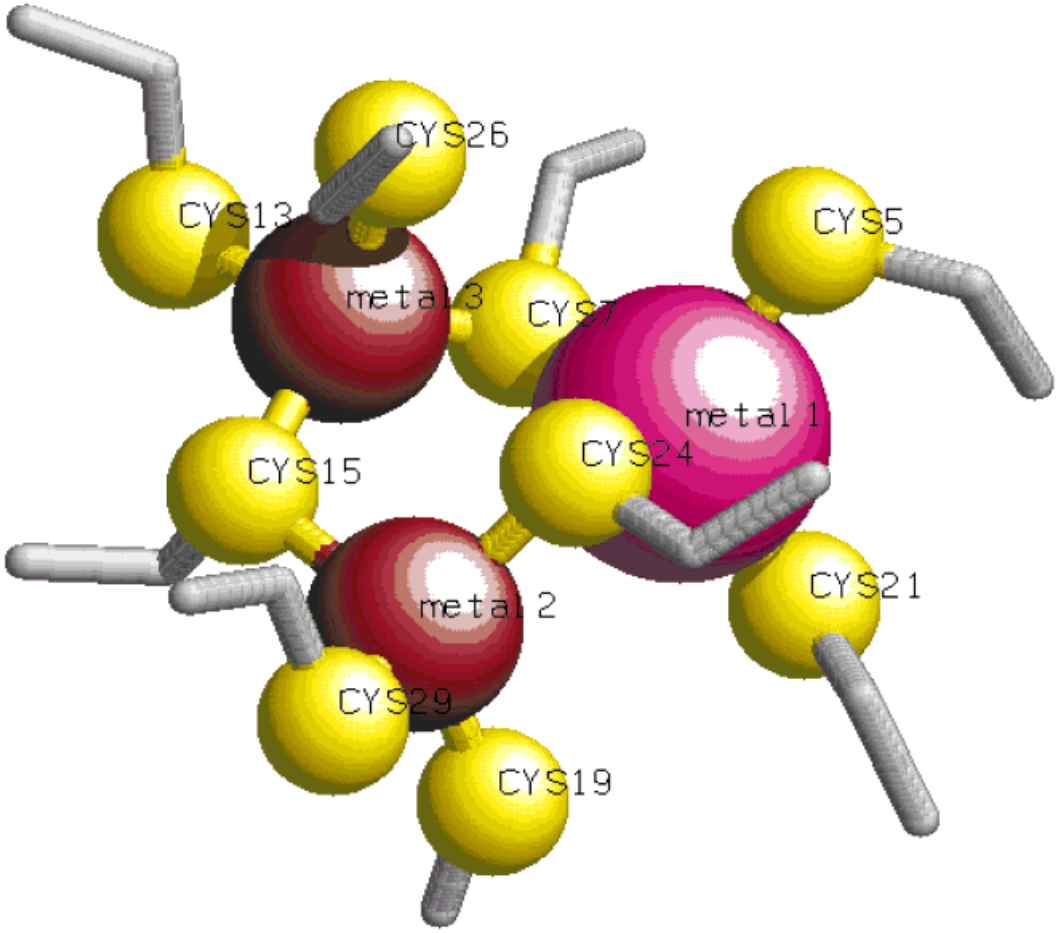

Fig. 1. Close-up of the metal core in the X-ray structure[3] of the β domain of rat liver metallothionein-2. The large red ball represents the cadmium ion, the two brown balls represent the zinc ions, and the small yellow balls represent the sulphur atoms. The cysteine side chains are displayed as thick gray sticks. The sulphur atoms of Cys7, Cys15, and Cys24 form bridges between the metals, while the other sulphur atoms are terminally coordinated.

protein cores are usually studied using density functional methods rather than semi-empirical methods.[8] In the case of metallothionein, primarily structural aspects are of interest. However, in many metallo-proteins, the metal core displays catalytic activity (e.g., zinc in alcohol dehydrogenase), enables electron transport (e.g., iron and copper in cytochromes), or captures light (e.g., magnesium in conjunction with porphyrin in the photosystems of plants). These properties and processes are certainly not suitable for a purely classical description.

Only the β domain of rat liver metallothionein-2 is considered in the present work. Experimental structures are available for the Cd_3 (NMR[2]) and the $CdZn$[2] (X-ray[3]) variants. Both structures are closely similar,[5] having the same metal-sulphur cluster geometries and a similar polypeptide fold. A more detailed comparison is given in Results. Figure 1 shows a close-up of the metal core of the β domain of metallothionein, containing two zinc ions and one cadmium ion. There are two types of sulphurs: a bridging type that is coordinated to two metal ions (from Cys7, Cys15, and Cys24), and a terminal type that is coordinated to a single metal ion. The metals and the bridging sulphurs form a twisted six-membered ring.

METHODS

In order to simplify notation, some abbreviations are introduced. We denote a purely classical molecular dynamics simulation as MDc. A molecular dynamics simulation with a combined quantum-chemical/force-field potential energy function is denoted as MDq. Three metal center variants were simulated: Cd_3, $CdZn_2$, and Zn_3. All three variants were simulated both fully classically (MDc) and combined with the semi-empirical method MNDO/d (modified neglect of diatomic overlap with extension to d orbitals)[9,10] (MDq). For the Zn_3 cluster, MNDO without d-orbital extension was employed in addition. The quantum-chemical part involves the three metal ions, the cysteinic sulphur atoms and β carbon atoms, and the attached hydrogen atoms. Thus, in the quantum-chemical core, the cysteines are reduced in size to methylthiolates. For the MDc simulations, force-field parameters for zinc were

taken from the GROMOS96 force field[6] and those for cadmium were estimated as described below.

The structure of the β-domain of rat liver metallothionein-2 (residues 1–30, containing the three-metal cluster) was obtained from the X-ray structure[3] (protein database entry 4MT2). The α domain was chopped off. The resulting structure of the $CdZn_2$ cluster was used as initial structure for the simulations of all three variants. A separate energy minimisation for each variant and force calculation scheme (MDc or MDq) was carried out prior to any dynamic simulation. The NMR structure[2] of the Cd_3 variant was used for structural comparisons (protein database entry 2MRT).

Quantum-Classical Coupling Model

The Hamiltonian of a system that is modeled using quantum chemistry and a force field (FF) can be divided into (1) a purely classical part (the environment) for which the interaction is given by the force field, (2) the purely quantum-chemical part (the core) that is modeled using semi-empirical quantum-chemical electronic calculation, and (3) an interface (or coupling) part that involves both force-field and semi-empirical interactions. If there are polar groups in the environment close to the quantum-chemical core, their partial charges influence the electron density of the core. This effect can be included into the quantum-chemical calculation by means of so-called background charges that enter the core-Hamiltonian (which is part of the Fock matrix used in the Roothaan equation) in the same way as the nuclear charges of the core do. The background charges differ from the nuclear charges insofar as they are mostly fractional, and that they do not possess basis functions. So, the electrons will still gather around the real nuclei, but the wave function is polarised by the background charges. The forces due to the quantum-chemical interaction, acting on the nuclei as well as on the atoms carrying background charges, are added to the classical forces. There are no classical Coulomb interactions between the atoms of the classical environment and the nuclei of the quantum-chemical core, only van-der-Waals interactions (vdW). The Hamiltonian of this model reads

$$H = H^{\mathrm{FF}}_{\text{environment-environment}} + H^{\text{FF-vdW}}_{\text{core-environment}} + \mathcal{H}^{\mathrm{QC}}_{\text{intra-core}} + \mathcal{H}^{\text{QC-Coulomb}}_{\text{core-environment}} + H^{\text{Link}}_{\text{core-environment}} \quad (1)$$

Similar to the non-bonded forces in the force field, only partial charges within the cutoff distance to any quantum atoms are included in the quantum-chemical calculation. These atoms are called the neighbour atoms. In a molecular dynamics simulation, the neighbour atoms change with time. So there has to be a mechanism that dynamically builds up a list of neighbour atoms. This is relatively easily implemented by scanning through the non-bonded interaction pair-list.

When simulating a protein, it is necessary to keep the quantum-chemical core small due to the computational expense of the quantum-chemical method. Therefore, chemical bonds will cross the boundary between the classical part and the quantum-chemical core. This has two major implications. Firstly, the quantum part has some dangling bonds. Therefore, a commonly used approach is to attach additional atoms, mostly hydrogen, to saturate the dangling bonds. These auxiliary atoms are called link atoms. Secondly, atoms that are bonded across the boundary lie so close to each other that they have to be excluded from the normal cross-boundary interaction, similar to the exclusion of first and second bond neighbours from the non-bonded interactions in a force-field.[6] This special interaction is indicated by the link Hamiltonian $H^{\text{Link}}_{\text{core-environment}}$.

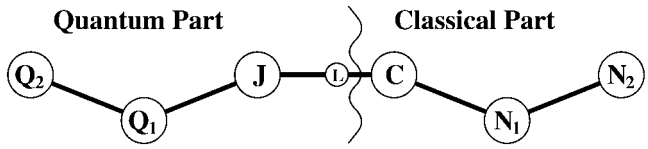

Fig. 2. Nomenclature used in the description of the link atom approach employed in the hybrid quantum/classical simulations.

Bond-Constrained Link Atom Model

In the widely used concept of link atoms,[11] auxiliary hydrogen atoms are added to the quantum-chemical part to saturate bonds across the quantum-classical boundary. As suggested previously,[11] the link atoms have no interactions with the classical part. A fundamental problem regarding the traditional link atom approach is that spurious atoms are introduced into the simulation, thus introducing unphysical degrees of freedom. For ease of description, let us introduce the following nomenclature (Fig. 2): The quantum-chemical atom that is bonded to a classical atom is called join atom **J**, its classical bonded partner is called connect atom **C**, and the link atom between the two is designated **L**. The neighbouring atoms on the quantum side are labeled with $\mathbf{Q}_1$, $\mathbf{Q}_2$, and so forth, the bonded atoms on the classical side are labeled $\mathbf{N}_1$, $\mathbf{N}_2$, and so on. The spurious extensions of degrees of freedom can be avoided if bond-constrained link atoms are employed. In this approach, the link atom is placed between the join atom and the connect atom in every time step. The exact location of the link atom $\vec{x}_{\mathbf{L}}$ is determined by a constant ratio s of atom distances

$$\vec{x}_{\mathbf{L}} = \vec{x}_{\mathbf{J}} + s(\vec{x}_{\mathbf{C}} - \vec{x}_{\mathbf{J}}) \quad (2)$$

where $\vec{x}_{\mathbf{J}}$ and $\vec{x}_{\mathbf{C}}$ are the positions of the join atom and the connect atom, respectively. The ratio s is chosen such as to reflect the ratio between the J-L and J-C standard bonds lengths. For example, if the quantum-classical boundary crosses a carbon-carbon bond (0.154 nm), which is replaced by a carbon-hydrogen bond (0.107 nm),[12] the ratio s is 0.695. A very similar approach is known as the scaled position link atom method (SPLAM),[13] and the same force redistribution scheme has been employed by Eichler et al.[14] So, the link atom does not move freely. In fact, it does not even exist as an atom that is propagated in time. Instead, the forces acting on it are distributed onto the join

atom and the connect atom, in such a way that the total force and the total torque is conserved,

$$\vec{F}'_{\mathbf{J}} = \vec{F}_{\mathbf{J}} + (1 - s)\vec{F}_{\mathbf{L}} \quad (3)$$

$$\vec{F}'_{\mathbf{C}} = \vec{F}_{\mathbf{C}} + s\vec{F}_{\mathbf{L}}. \quad (4)$$

This procedure gives the bond across the boundary quantum character, as the bond characteristics from the join atom to the link atom is transferred to the bond with the connect atom. Consequently, no classical bond-stretching force-field term is used to describe the bond across the boundary, which may not be very accurate. However, this does not matter much, since the focus of interest is usually at the center of the quantum-chemical core, not at its boundary. The following force-field terms across the quantum-classical boundary are still required (see Fig. 2): bond angle **J**-**C**-$\mathbf{N}_1$ and torsion angles $\mathbf{Q}_1$-**J**-**C**-$\mathbf{N}_1$ and **J**-**C**-$\mathbf{N}_1$-$\mathbf{N}_2$.

It should be mentioned that the bond-constrained link atom approach is compatible to a new development called adjusted connection atom (ACA).[15,16] This approach employs a special semi-empirical parameterisation for an atom that mimics a carbon atom, but is monovalent. Such an artificial atom is then used as a link atom, which is at the same time the connect atom. This corresponds to a bond length ratio *s* equal to unity.

It has been argued that in energy minimisation,[17,18] there are some advantages in using freely moving link atoms rather than fixed ones. However, in later work,[15] the differences were found to be small. When using link atoms in molecular dynamics simulations, it is more important to avoid the unwanted artificial increase of degrees of freedom since the presence of moving artificial link atoms affects the kinetic energy.

Covalently bonded atoms are normally so close to each other that any non-bonded interaction, either van-der-Waals or electrostatic, would be much too strong. So bonded atoms (first neighbours) are usually excluded from non-bonded interactions. The same applies for atoms connected by two bonds. An exception is the chemical bond with a predominant electrostatic character, as present in the binding of metal cations to negatively charged species. Such a situation is often modeled using the balance between the attractive electrostatic interaction, and the repulsive part of the van-der-Waals interaction. However, it is sometimes difficult to accurately model such a situation, and auxiliary bonded force-field terms can be used in addition to the non-bonded interactions. In the present work, we have not used such terms.

In analogy to classical force fields, there are no van-der-Waals interactions between first and second neighbours across the quantum-classical boundary. Hence, the following three atom pairs are excluded (see Fig. 2): **J**-**C**, **J**-$\mathbf{N}_1$, $\mathbf{Q}_1$-**C**.

Concerning the background charges that enter the quantum-chemical calculations, the partial charge on the connect atom is neglected. Partial charges on other atoms are included. A comparison of different options for treating shortest-range charge interaction has been presented by Antes.[15] The differences found are rather small when using the current embedding models and semi-empirical wave functions. The neglect of partial charges is not a major concern in the present case of metallothionein, because the connect atom (the C_α of cysteine) is uncharged in the GROMOS96 force field.

Computational Details

All four lysines in the β domain were protonated and the two aspartic acids deprotonated. The protein domain was simulated in a periodic box of water with a truncated-octahedral shape. A minimum protein-to-wall distance of 1.4 nm was used, giving a total of 2,745 water molecules and 8,446 atoms. The volume of the box was held constant at 86.6 nm^3.

For all classical atoms, the GROMOS96[6] force field 43A1 was employed. Water was modeled using the simple point charge model (SPC[19]). Classical bonds were constrained to a relative geometric accuracy of 10^{-4}.[20] For the non-bonded forces, a twin-range cutoff of 0.8/1.4 nm was used with a reaction field correction[21] ($\epsilon_{RF} = 54$, as determined for SPC water[22]). The short cutoff defined at the same time the interface region of background partial charges that enter the quantum-chemical calculations. More precisely, the classical partial charges of any charge group having at least one member closer than the short-range cutoff of 0.8 nm to any quantum atom, were included in the quantum-chemical calculation. This led to an interface region usually consisting of the whole protein domain plus a shell of water, totally comprising about 655 atoms.

In the quantum-chemical core, the metal ions had a formal oxidation state of +2 and were coordinated to deprotonated methylthiolate. Totally, the quantum core had the composition $M_3(CH_3S)_9$ and was charged −3 e, where M is a placeholder for any metal, Zn or Cd. One of the hydrogen atoms of each methylthiolate served as link atom. The quantum-chemical β-carbon atom was linked to the classical α-carbon atom by means of a bond-constrained link atom approach as described above. Hydrogen atoms were used as link atoms and the bond length ratio was 0.6948.

The protein and water were separately weakly coupled to a temperature bath of 300 K using a 0.1 ps coupling time.[23] The time step for the MDc simulations was 2 fs. In the MDq simulations, 0.5 fs was used to account for faster vibration of the unconstrained bonds in the quantum-chemical part. The non-bonded-interaction pair list was updated every 10 fs. The self-consistent field (SCF) procedure was iterated up to a convergence of the energy of 10^{-6} electron volt without an SCF convergence accelerator. Analytic gradients were employed. Simulation of a trajectory of 10 ps on a 450 MHz dual-processor pentium-II computer took roughly $1\frac{1}{4}$ h for an MDc simulation, 14 h for an MDq simulation based on MNDO, and 30 h for an MDq simulation with MNDO/d. The simulation lengths were 10 ns for the MDc simulations, and 460 ps for the MDq simulations.

TABLE I. Some Properties of Metal Sulfides and Non-Bonded-Interaction Force-Field Parameters for Zinc[6] and Cadmium

Mineral	Density g/cm^3	Metal-sulphur distance $l_{M\text{-}S}$ nm	van-der-Waals parameters $\sqrt{C_M^{six}}$ (nm^6 kJ/mol)$^{1/2}$	van-der-Waals parameters $\sqrt{C_M^{twelve}}$ 10^{-3} (nm^{12} kJ/mol)$^{1/2}$
Zincblende	4.102	0.234	0.02045	0.09716
Cadmiumblende	4.82	0.253	0.03267	0.24790

Estimation of Van-der-Waals Interaction Parameters for Cadmium

The van-der-Waals interaction parameters for cadmium were estimated using the GROMOS96[6] zinc parameters as a starting point. The basic structure of the metal clusters in metallothionein is equivalent to the mineral form of the metal sulfides, zincblende and cadmiumblende. Knowing the structure and the density of the latter, and the masses of the involved atoms, a metal-sulphur distance can be derived. The values shown in Table I are very close to those in the X-ray[3] and NMR[2] structures of metallothionein. However, the latter values may result from the bond restraints applied in the structure refinement process.

The zinc parameters were scaled to reflect the larger bond length with cadmium, while retaining the depth of the minimum of the van-der-Waals term in the force field. Using $r = l_{Cd\text{-}S}/l_{Zn\text{-}S}$, the ratio of metal-sulphur distances in the mineral, we obtain the scalings

$$\sqrt{C_{Cd}^{six}} = r^6\sqrt{C_{Zn}^{six}} \quad (5)$$

$$\sqrt{C_{Cd}^{twelve}} = r^{12}\sqrt{C_{Zn}^{twelve}} \quad (6)$$

where $\sqrt{C_{Zn}^{six}}$ and $\sqrt{C_{Zn}^{twelve}}$ are the GROMOS96 van-der-Waals parameters for zinc. The resulting values are listed in Table I.

RESULTS AND DISCUSSION

Both, the MDc and MDq simulations with the semi-empirical method MNDO/d were able to maintain the overall structure of the protein. However, MDq simulations using MNDO without extension to d orbitals failed in this respect. The metal cluster disintegrated after a short simulation period of 15 ps. This was not an accidentally observed unfolding event: Two more simulations with different starting velocities suffered the same fate. Therefore, only the MDq simulations with MNDO/d will be discussed.

The following subsections give more details and compare the results of the two methods against each other and against experimental data. The section is organised as follows. Tables and figures show results grouped in terms of properties such as bond lengths or NOE distances. To avoid confusion, the discussion in the text is grouped in terms of comparisons: comparisons between experimental and simulated structures, comparisons between MDc and MDq simulated results, or between Cd_3, $CdZn_2$, and Zn_3 variants.

Comparison of the $CdZn_2$ X-Ray Crystal Structure With the Cd_3 NMR Solution Structure

The structures of the $CdZn_2$ (X-ray[3]) and the Cd_3 variant (NMR[2]) derived from experimental data are very similar.[5] In particular, the metal cores have the same coordinative bonds and metal-sulphur cluster geometries. The polypeptide folds are closely similar. However, the polypeptide loops linking the metal-bonded cysteines are less well defined in the NMR structure. This finding is attributed to the absence of regular secondary structure and the high degree of dynamic structural disorder.

Figure 3 shows the experimental structures and the structures at the end of the simulations. The top row shows the NMR structure to the left and the X-ray structure to the right. Both structures look similar indeed: The metals and sulphurs are nearly identical and the overall fold is the same, however, with quite some variation in the loops between the cysteine residues. There is a difference in the direction of the side chain of Cys13: In the X-ray structure it faces the sulphur from the top-front, while in the NMR structure from behind.

Figure 4 shows the backbone C_α atom positional root-mean-square (RMS) deviations between the X-ray and the NMR structure. The average RMS deviations of the cysteine residues are 0.18 nm for the C_α atoms, 0.14 nm for the C_β atoms, and 0.03 nm for both the sulphurs and the metals. The metal-sulphur configurations are very similar. The RMS deviations increase with increasing distance from the metal core, both atom-wise within the cysteines and residue-wise in the entire domain. This is not surprising because the cysteine sulphur atoms were superimposed.

The difference in position of cysteine C_α atoms between the NMR and the X-ray structures is about 0.2 nm on average, which is reasonable considering the difference in metal atoms (Cd_3 vs. $CdZn_2$) and environment (solution vs. crystal). However, the average difference over all C_α atoms is about 0.3 nm, which reflects a high degree of flexibility, which is most probably due to the absence of regular secondary structure and its stabilising hydrogen bonds.

There is a large variation in the C_α positions of individual residues. Not surprisingly, the cysteines mostly exhibit a very low variation. An exception is Cys13, which was previously mentioned as having a different side chain orientation in the two structures. There is a region of large deviation from Thr9 to Ser14. This loop has different conformations in the two structures, see Figure 3a and b, loop at the top. The same applies to Gly17, which is at the

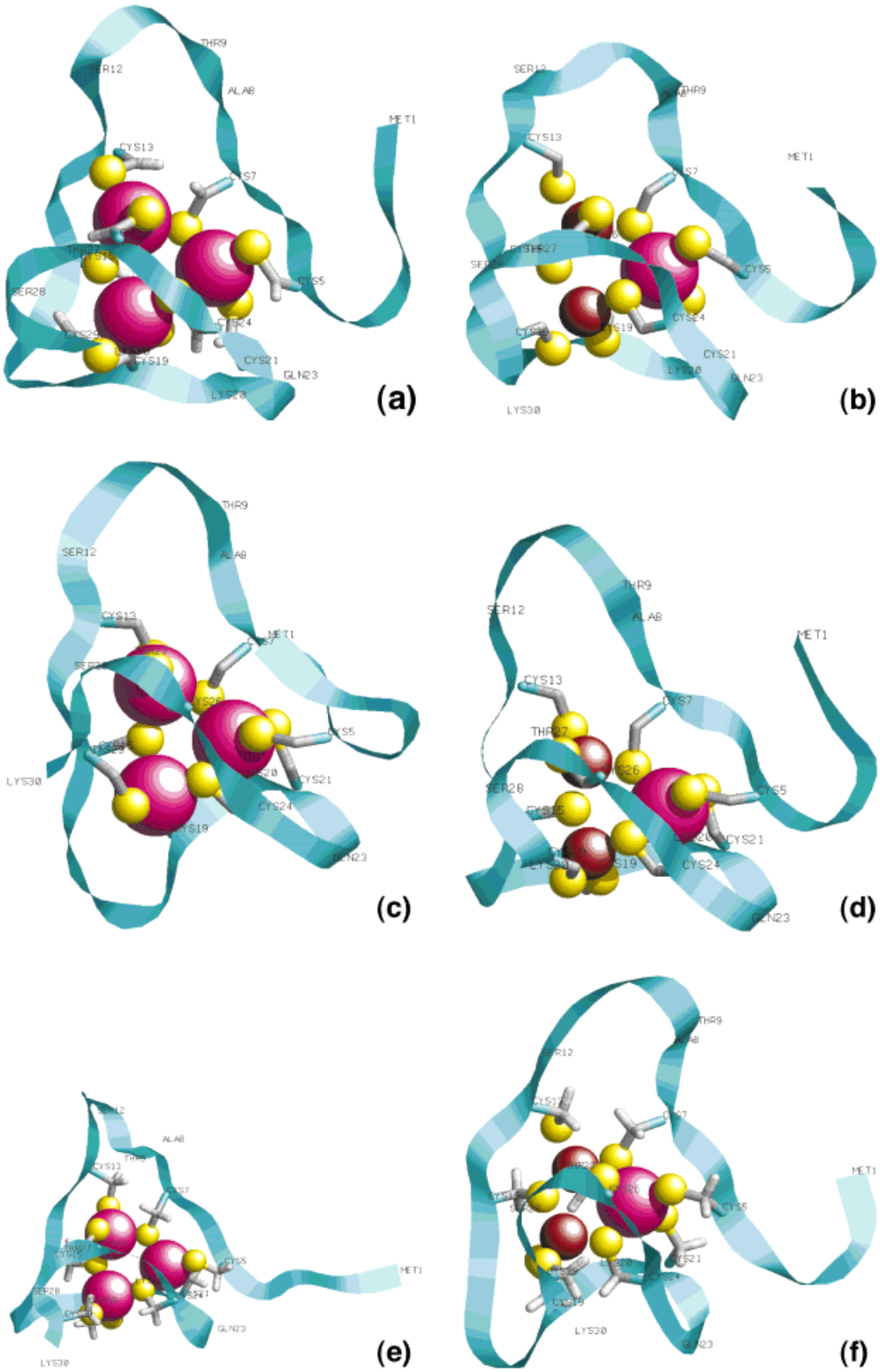

Fig. 3. Schematic structures of the β domain of rat liver metallothionein-2 containing three metal ions. All structures are rotated to display the same view on the metals. Cadmium ions are displayed as large red balls, zinc ions as medium-sized brown balls, sulfur atoms as small yellow balls, the rest of the cysteine side chains as thick gray sticks, and the protein backbone as blue ribbon. A selection of residues is labeled. **a:** Cd_3 NMR structure.[2] **b:** $CdZn_2$ X-ray structure.[3] **c:** Cd_3 final MDc simulation structure. **d:** $CdZn_2$ final MDc simulation structure. **e:** Cd_3 final MDq simulation structure. **f:** $CdZn_2$ final MDq simulation structure. The X-ray structure and the MDc simulations do not have hydrogens on the cysteine side chain, the NMR structure shows a single pseudo atom, and the MDq simulations have explicit hydrogens.

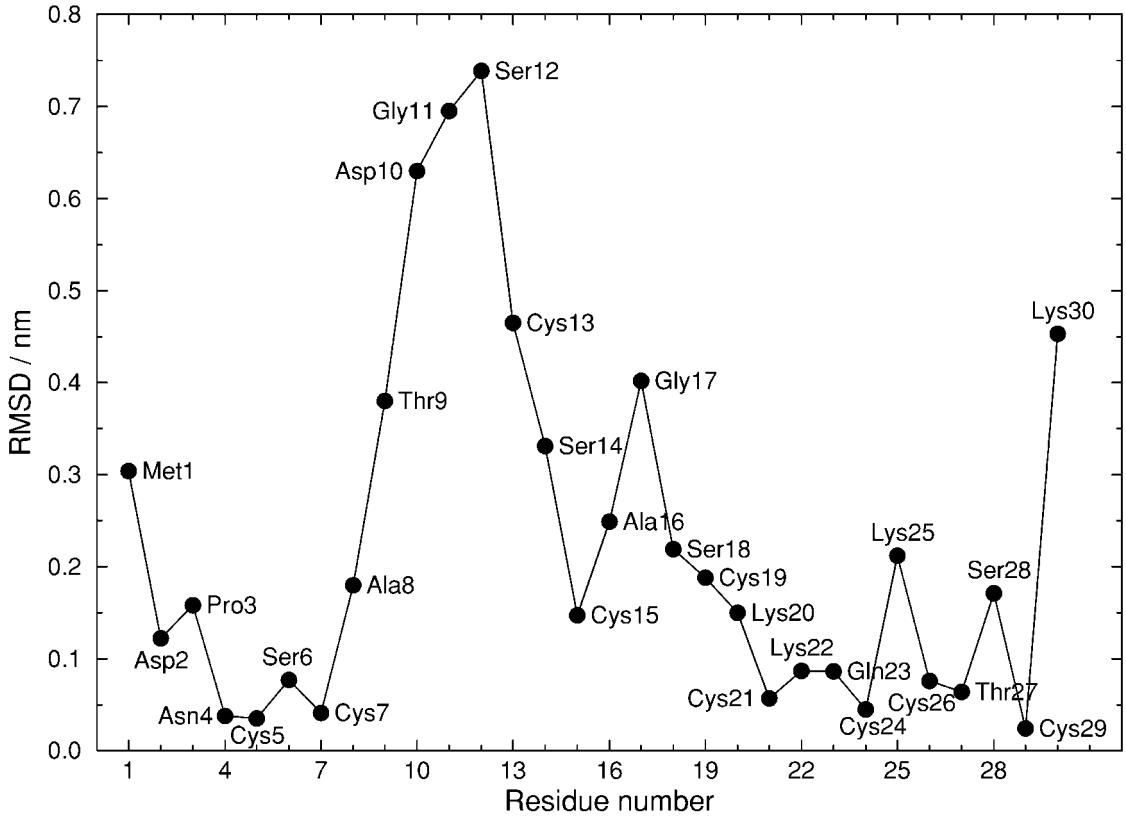

Fig. 4. Positional root-mean-square deviations for C_α-atoms between X-ray[3] and NMR structure.[2] The two structures were superimposed using a translational and rotational least-squares fit for the cysteine sulphur atoms.

edge of the loop to the lower back left in Figure 3a and b. The terminal residue Met1 and the linker to the second domain, Lys30, are expected to exhibit more structural variation. The X-ray structure only weakly defines residues Met1 and Asp2. These results suggest that the simulated RMS deviations from the experimental structures should be below 0.05 nm for the metal and cysteine sulphur atoms, below 0.15 nm for cysteine C_β atoms, below 0.2 nm for cysteine C_α atoms, and below 0.3 nm for all C_α atoms. These thresholds will be used as a tolerance level in analysing the simulations.

Similar regions of high structural variation are also found amongst the set of the ten best NMR structures (fig. 2 in Braun et al.[5]). There, regions around Asp10, Ala16, and Thr27 exhibit large variations, apart from the chain ends. It is likely that the flexibility of the intercysteinal loops causes these increased deviations.

Table II lists hydrogen bonds present in the experimental structures. Both experimental structures were relaxed by an energy minimisation with atom-positional restraints (force constants equal 2.5 10^4 kJmol^{-1} nm^{-2}) to the original structure in order to remove strain and to adapt the structure to the GROMOS force field and its criterion of hydrogen bonds (see footnote to Table II). The asterisks mark hydrogen bonds already present in the original experimental structures.

There are relatively few hydrogen bonds, and the hydrogen bonding pattern differs quite much between the X-ray and the NMR structures. Two hydrogen bonds are present in both experimental structures, both involving Ser12 and Asp10. Interestingly, these residues are in the region of maximum RMS difference between the two structures (Fig. 4). It seems that a relatively stable turn is formed, whose relative position with respect to the metal core is different. The other hydrogen bonds are present in either of the structures.

$CdZn_2$ MDc Simulation Compared to the X-Ray Structure

A crucial property of the metal cluster are the metal-sulphur bond lengths. For this reason, bond lengths were averaged in time, after an equilibration phase of 50 ps, and summarised in Table III. The experimental bond lengths are quite close to the ones found in the minerals (Table I). Compared to the experimental X-ray structure, the MDc simulation yields bonds about 0.025 nm too short for the cadmium ion. For zinc, the difference is worse, 0.031 nm. It should be noted that the experimental values were refined against 0.24 nm for zinc and 0.25 nm for cadmium.

The structures of the entire domain were analysed by means of atom-positional root-mean-square deviations (RMSD; see Fig. 5). They were calculated by comparing

TABLE II. List of Hydrogen Bonds That Occur in Any Experimental Structure or in Any Simulation for More than 20%†

Hydrogen bond specification		Experimental		Percentage of occurrence					
				MDc			MDq		
		NMR Cd_3	X-ray $CdZn_2$	Cd_3	$CdZn_2$	Zn_3	Cd_3	$CdZn_2$	Zn_3
Asp2-Lys25	N-H-O	—	—	22.4	—	—	—	—	—
Asp2-Cys5	N-H-O	—	—	—	7.8	56.1	5.2	—	4.0
Asn4-Asp2	NDx-HDx-ODx	—	—	47.9	48.6	13.5	21.3	32.3	65.5
Asn4-Asp2	N-H-ODx	—	100.0*	90.0	86.5	80.8	20.8	17.5	60.0
Cys5-Asp2	N-H-ODx	—	100.0	33.4	63.6	62.6	—	6.7	2.6
Gly11-Ala8	N-H-O	—	—	3.3	12.5	21.1	14.5	—	17.9
Ser12-Asp10	N-H-ODx	100.0*	100.0*	—	16.9	2.4	23.4	6.9	36.0
Ser12-Asp10	OG-HG-ODx	100.0	200.0	11.1	23.1	4.7	63.9	70.6	54.0
Ser12-Thr27	OG-HG-OGx	—	—	21.5	3.0	—	—	—	—
Ser12-Ala8	N-H-O	—	—	—	—	—	20.2	—	7.8
Cys13-Thr27	N-H-OGx	—	—	—	9.5	—	—	21.6	—
Cys13-Ala8	N-H-O	—	—	8.8	13.9	—	33.5	—	31.8
Ser14-Ser28	OG-HG-OG	—	—	24.5	—	—	5.9	—	—
Cys15-Ser28	N-H-OG	—	—	—	93.3	—	—	75.9	—
Cys15-Ser28	N-H-O	—	—	21.2	—	—	—	—	—
Ala16-Ser28	N-H-OG	—	100.0*	—	3.8	—	—	9.4	—
Gly17-Ser28	N-H-O	—	—	—	—	—	—	—	67.6
Ser18-Cys15	N-H-O	—	100.0*	—	—	—	16.8	—	—
Cys19-Gly17	N-H-O	—	—	26.0	7.4	—	18.2	—	2.5
Lys20-Ser6	NZ-HZx-O	—	—	8.2	—	—	42.9	—	—
Lys22-Asn4	N-H-ODx	—	—	—	—	—	25.5	—	—
Lys22-Asn4	N-H-O	—	100.0	6.4	9.1	27.8	49.0	2.6	8.7
Gln23-Lys22	NEx-HEx-O	—	—	—	—	2.7	—	23.3	15.4
Gln23-Gln23	NEx-HEx-O	100.0	—	—	—	—	—	—	—
Gln23-Asn4	N-H-ODx	—	—	—	3.0	22.7	45.1	—	4.1
Gln23-Asn4	N-H-O	—	100.0*	—	85.5	81.2	66.1	85.5	48.6
Lys25-Gln23	NZ-HZx-O	—	100.0*	—	—	—	—	21.9	2.9
Lys25-Asp2	NZ-HZx-ODx	—	100.0*	—	13.0	2.1	—	2.6	—
Cys26-Cys29	N-H-O	100.0	—	—	8.8	—	—	—	—
Thr27-Lys25	N-H-O	100.0*	—	—	—	—	—	—	—
Ser28-Cys13	N-H-O	—	—	—	—	—	—	23.9	—
Ser28-Cys13	OG-HG-O	—	100.0	—	5.4	—	—	64.1	—
Ser28-Ser14	N-H-OG	—	—	—	—	71.7	3.4	—	—
Cys29-Cys26	N-H-O	—	100.0*	16.8	17.3	27.0	41.5	39.9	62.9
Lys30-Ser28	N-H-O	100.0*	—	—	8.6	2.6	9.7	13.8	24.6

†Values larger than 100% result from three-center hydrogen bonds, when both single components are present at the same time. The character x denotes two equivalent atoms. In such cases, the percentages for equivalent atoms were added. The experimental structures were relaxed by an energy minimisation to remove conformational strain. Hydrogen bonds present in the original experimental structure are marked by an asterisk. A hydrogen bond is considered to exist if the hydrogen-acceptor distance is smaller than 0.25 nm and the donor-hydrogen-acceptor angle is larger than 135°. A dash denotes that the hydrogen bond has not been observed in the corresponding simulation.

structures from the trajectory with a reference structure after a translational-rotational fit over the cysteine sulphur atoms. Separate RMSD curves were calculated for the metals, for the cysteine sulphur atoms, for the cysteine β-carbon atoms, for the cysteine α-carbon atoms, and all α-carbon atoms of the domain.

Figure 5 shows RMS deviations from the initial structures. Figure 5c for the $CdZn_2$ MDc simulation shows that the deviation of the metal and sulphur atoms is above the tolerance level of 0.05 nm. This can be explained as a consequence of the too short bond lengths in the metal cluster. The deviations of the cysteine C_α and C_β atoms slowly drift up to 0.2 nm. The total of the C_α atoms is below 0.3 nm. Summarising, the MDc simulation shows too large a deviation for atoms of the metal core (metal, sulphur, and cysteine C_β atoms), whereas the backbone deviations are comparable to the difference between the X-ray and NMR structures in Figure 4. The same applies to the RMS deviations from the X-ray structure shown in Figure 6c. This is no surprise, as the initial structure for the simulation was obtained using energy minimisation from the X-ray structure.

The X-ray structure exhibits bond angles of the bridging sulphur in a narrow range around 104° (Table IV). In contrast, the MDc simulation yields average angles larger than 130°. As a consequence, the angles at the metals between bridging sulphurs are smaller in the simulation (95°) than in experiment (between 104° and 123°). These results are further discussed below when comparing the quantum-chemical and classical treatment.

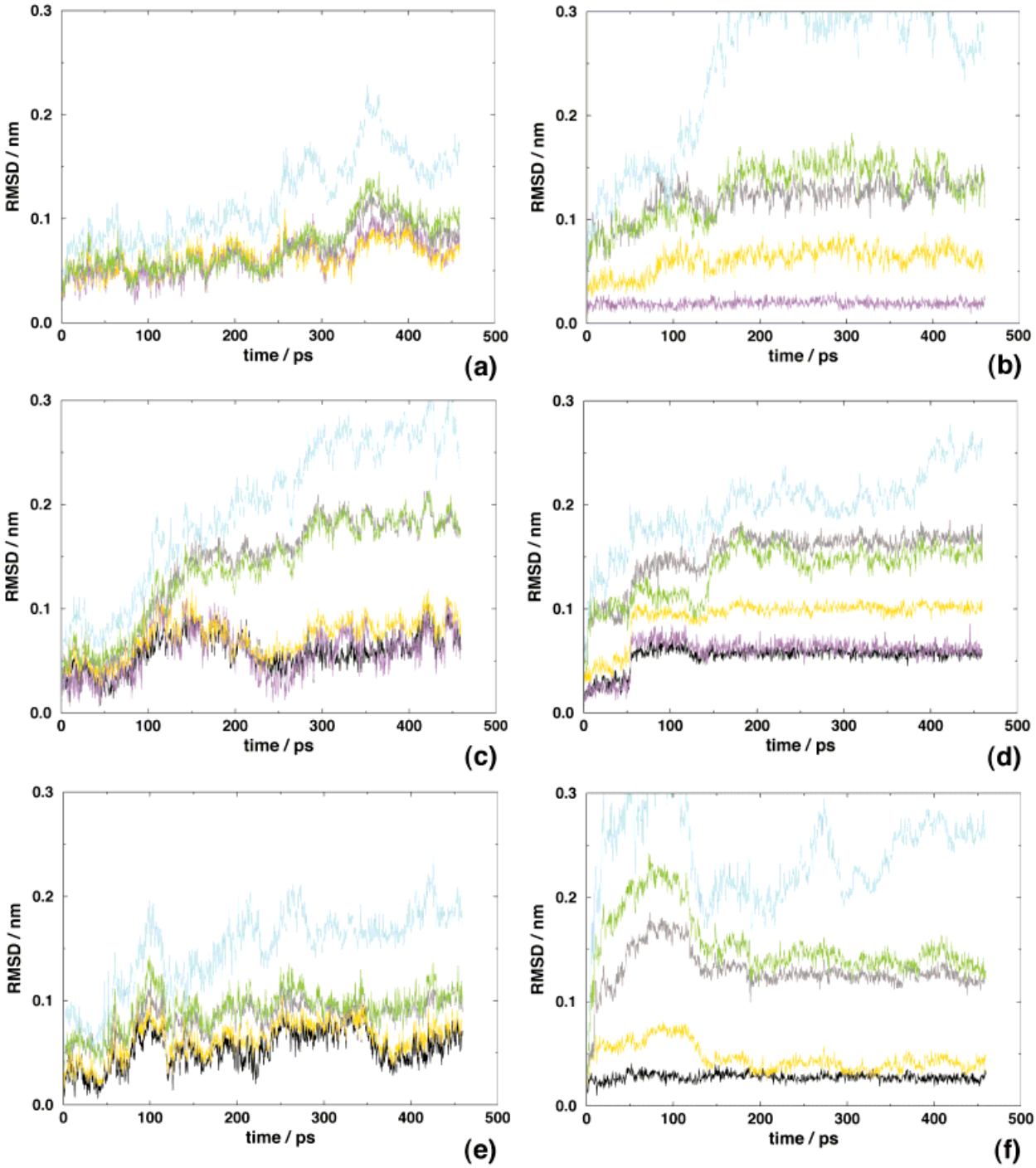

Fig. 5. Atom-positional RMS deviations from the initial structures over a period equal to the length of the MDq simulations. The RMS deviation was calculated after a translational-rotational fit over the cysteine sulphur atoms. **a:** MDc simulation of Cd_3 variant; **b:** MDq simulation of Cd_3 variant; **c:** MDc simulation of $CdZn_2$ variant; **d:** MDq simulation of $CdZn_2$ variant; **e:** MDc simulation of Zn_3 variant; **f:** MDq simulation of Zn_3 variant. Black: zinc ions; purple: cadmium ions; yellow: sulphur atoms; gray: cysteine C_β atoms; green: cysteine C_α atoms; blue: all C_α atoms.

In the $CdZn_2$ MDc simulation, most hydrogen bonds (Table II) of the X-ray structure are observed, Lys25-Gln23 NZ-HZx-O and Ser18-Cys15 N-H-O are lost. Most of the frequent hydrogen bonds in the simulation are present in the crystal; however, two prominent new ones appear in the simulation: Asn4-Asp2 NDx-HDx-ODx and Cys15-Ser28 N-H-OG. Residues Lys25, Ser18, Ser28, and Cys15 have deviations larger than 0.14 nm between the X-ray

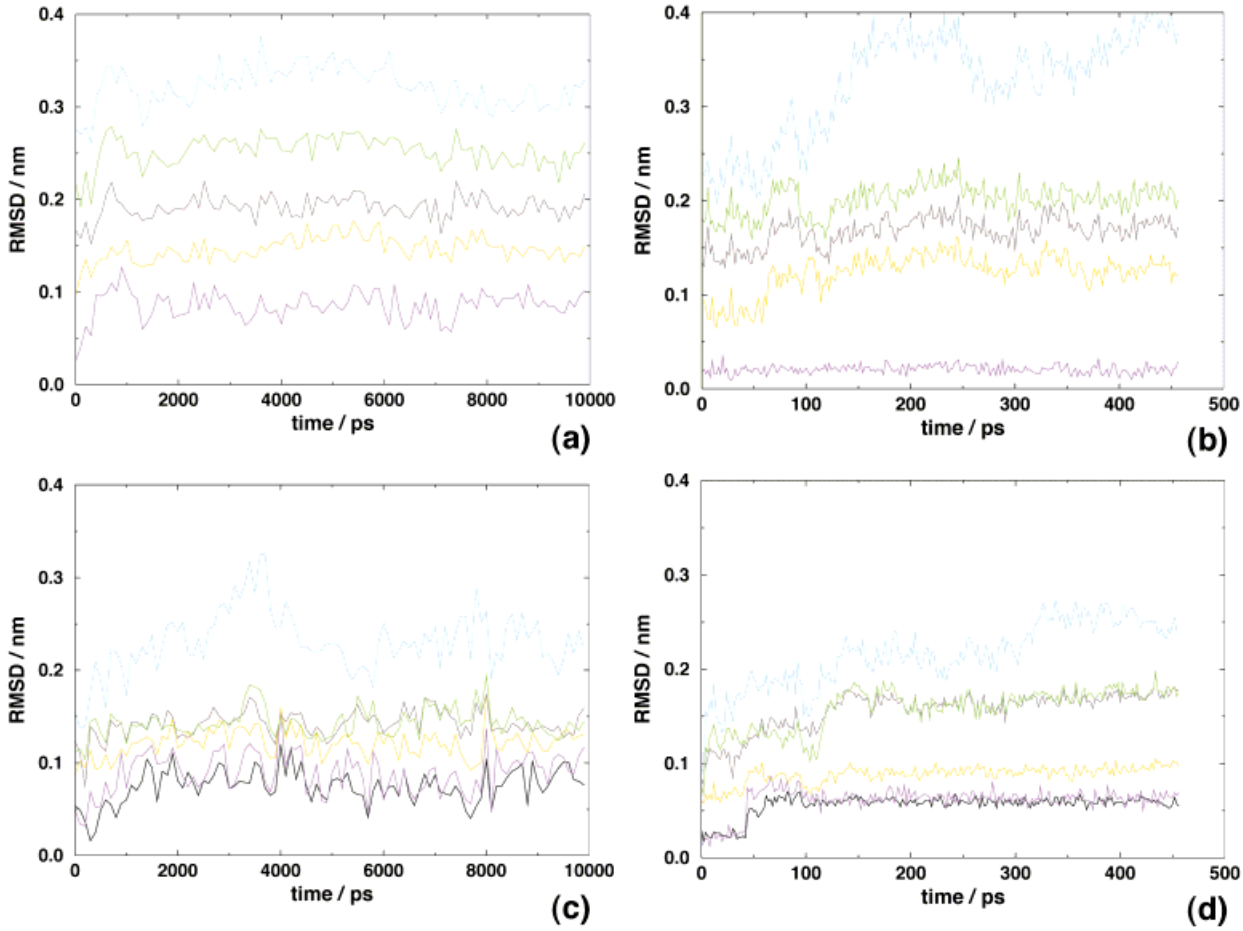

Fig. 6. RMS deviations from the experimental structures. For the Cd_3 cluster (a and b), this is the NMR structure[2] and for the $CdZn_2$ cluster (c and d), it is the X-ray structure.[3] **a:** MDc simulation of Cd_3 variant; **b:** MDq simulation of Cd_3 variant; **c:** MDc simulation of $CdZn_2$ variant; **d:** MDq simulation of $CdZn_2$ variant. Black: zinc ions; purple: cadmium ions; yellow: sulphur atoms; gray: cysteine C_β atoms; green: cysteine C_α atoms: blue: all C_α atoms.

and NMR structures (Fig. 4), which indicates different behaviour in the crystal and in solution. Asp2 is weakly defined in the X-ray structure. This might explain the absence of the corresponding hydrogen bond in the X-ray structure.

$CdZn_2$ MDq Simulation Compared to the X-Ray Structure

The length of the cadmium-sulphur bond is well reproduced (Table III). More difficulties arise with zinc, whose bond to sulphur appears about 0.016 nm too short.

The RMS deviations from the initial conformation (Fig. 5d) for all atom types are initially below the thresholds mentioned earlier. However, there is a stepwise increase around 50 ps where the levels for the metals and the sulphurs rise above the thresholds. Then the deviations stay around these levels with the same small fluctuations as before. The metal core seems to undergo a transition to another stable structure. The deviations for the two C_α atom types also increase a bit, but remain well below the tolerance threshold. Virtually the same behaviour is observed in comparison to the X-ray structure in Figure 6d. As the initial structure was derived from the X-ray structure, this is not surprising. However, there is a significant difference in the deviation of the sulphur atoms: it is initially already above 0.05 nm, but the increase around 50 ps is smaller, the final level lower than compared to the initial structure. Overall, the MDq simulation gives small deviations, mostly below 0.2 nm, which is considered to be a normal structural variation for the α carbon atoms. Compared to the MDc simulations that exhibit a steadily growing RMSD, they stabilise in the MDq simulations, and their fluctuations are smaller.

The bond angles of the MDq simulations yield an irregular picture (Table IV). The average of the sulphur bridge angles (M-S_b-M) is quite close to the experimental values, but the differences between the individual angles are large, in contrast to experiment. The metal ring bridge angles (S_b-M-S_b) are too large, but give the correct trend in the variation of the three individual angles. Two of the terminal angles (S_t-M-S_t) are in good agreement with the experiment; the other one deviates by 15°.

TABLE III. Time-Averaged Bond Lengths (Aver) and Fluctuations in Time (Fluct) for Bonds between Metal (M) and Sulphur (S) Atoms (in nm)[†]

	Metal 1				Metal 2				Metal 3			
	M-S_t		M-S_b		M-S_t		M-S_b		M-S_t		M-S_b	
MDc Cd_3												
Aver	0.223	0.225	0.226	0.225	0.229	0.227	0.230	0.231	0.225	0.224	0.225	0.228
Fluct	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002	0.002	0.002	0.001	0.002
Mean		0.225				0.229				0.226		
MDc $CdZn_2$												
Aver	0.223	0.224	0.227	0.227	0.205	0.204	0.211	0.210	0.204	0.205	0.208	0.207
Fluct	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001
Mean		0.226				0.207				0.206		
MDc Zn_3												
Aver	0.203	0.203	0.207	0.207	0.205	0.202	0.209	0.207	0.207	0.211	0.213	0.211
Fluct	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002	0.002	0.002
Mean		0.205				0.206				0.210		
MDq Cd_3												
Aver	0.244	0.245	0.245	0.240	0.243	0.236	0.241	0.243	0.245	0.234	0.242	0.246
Fluct	0.002	0.002	0.002	0.001	0.003	0.002	0.002	0.003	0.002	0.002	0.003	0.002
Mean		0.243				0.241				0.242		
MDq $CdZn_2$												
Aver	0.247	0.251	0.252	0.242	0.220	0.230	0.214	0.218	0.231	0.223	0.216	0.223
Fluct	0.002	0.002	0.004	0.002	0.002	0.003	0.003	0.003	0.002	0.002	0.003	0.004
Mean		0.248				0.220				0.223		
MDq Zn_3												
Aver	0.228	0.229	0.216	0.212	0.221	0.233	0.205	0.208	0.230	0.229	0.207	0.207
Fluct	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.002	0.002	0.002	0.003
Mean		0.221				0.216				0.218		
NMR Cd_3												
Aver	0.251	0.251	0.254	0.264	0.254	0.247	0.249	0.268	0.251	0.253	0.252	0.265
Mean		0.255				0.255				0.255		
X-ray $CdZn_2$												
Aver	0.249	0.249	0.254	0.254	0.230	0.241	0.241	0.248	0.237	0.237	0.237	0.233
Mean		0.251				0.240				0.236		

[†]The two left columns of every metal group represent the bonds to the terminal sulphur (S_t), the two right columns represent the bonds to the bridging sulphur (S_b). The means of the four bonds per metal are also given. Metal 1 is the cadmium atom in the $CdZn_2$ variants.

Hydrogen bonds present in the X-ray structure are mostly observed in the MDq simulation; however, several with very low percentage. The only hydrogen bond lost in the simulation is Ser18-Cys15 N-H-O. Prominent new hydrogen bonds established in the simulation are Asn4-Asp2 NDx-HDx-ODx and Cys15-Ser28 N-H-OG, as in the MDc simulation.

Comparison of the Cd_3 MDc Simulation With NMR Data

The bond lengths (Table III) are systematically too short, with differences between MDc simulations and experiment of 0.028 nm. This has two reasons. Firstly, the parameters for cadmium were estimated using the zinc parameters as a basis. So the inaccuracy of the results for zinc is propagated to cadmium. Secondly, the bond lengths in the NMR structure are larger than in the X-ray structure, due to different bond-length parameters (0.26 nm) used in the structure derivation.[24]

Figure 5a shows that the backbone and the cysteine C_β atoms exhibit small deviations well below the tolerance thresholds. However, the deviations for the metal and sulphur atoms drift above the tolerance. Again, this seems to be a consequence of the too short bond lengths induced by the estimated van-der-Waals parameters for cadmium.

As the simulations were started using the X-ray structure as a basis, and the differences between the X-ray and NMR structures are partially large, it is expected that the deviations of the Cd_3 cluster against the NMR structure

TABLE IV. Time-Averaged Bond Angles (aver) and Fluctuations (fluct) In the Metal Core, In Degrees

	Sulphur ring bridges M-S_b-M			Metal ring bridges S_b-M-S_b			Metal-terminal angles S_t-M-S_t		
	S_{b1}	S_{b2}	S_{b3}	M_1	M_2	M_3	M_1	M_2	M_3
MDc Cd_3									
Aver	132.0	133.0	147.7	101.5	87.8	92.2	96.9	156.7	125.8
Fluct	2.8	3.1	2.8	3.5	2.0	1.6	5.2	14.4	4.0
Mean		137.6			93.8			126.4	
MDc $CdZn_2$									
Aver	131.1	134.9	141.2	93.2	93.8	97.7	96.2	115.1	112.0
Fluct	2.4	2.0	2.5	2.4	1.9	1.4	1.9	6.4	3.0
Mean		135.7			94.9			107.8	
MDc Zn_3									
Aver	131.7	137.2	135.7	95.8	96.4	98.8	115.2	113.9	95.8
Fluct	2.5	2.1	2.1	1.5	1.5	2.2	3.4	3.6	4.6
Mean		134.9			97.0			108.3	
MDq Cd_3									
Aver	132.8	122.4	122.9	105.0	104.0	110.6	110.6	97.7	119.1
Fluct	3.0	3.2	2.5	3.7	2.0	3.4	4.8	4.6	4.3
Mean		126.1			106.5			109.1	
MDq $CdZn_2$									
Aver	115.9	82.8	100.7	123.3	115.0	154.2	97.0	97.3	106.8
Fluct	3.7	4.8	7.3	2.7	7.1	6.0	1.7	4.5	2.4
Mean		99.8			130.8			100.4	
MDq Zn_3									
Aver	112.7	111.2	109.7	100.2	108.1	128.7	94.8	92.2	92.5
Fluct	6.0	5.3	8.1	2.7	4.2	8.3	2.9	3.7	1.8
Mean		111.2			112.3			93.2	
NMR Cd_3									
Aver	108.5	110.0	109.4	102.9	101.6	106.5	115.9	112.3	110.1
Mean		109.3			103.7			112.8	
X-ray $CdZn_2$									
Aver	102.3	103.3	106.2	109.0	103.7	122.5	113.1	99.1	105.6
Mean		103.9			111.7			105.9	

†Left group: angle at bridging sulphurs between two metal ions; middle group: angle at a metal between two bridging sulphurs; right group: angle at a metal between two bridging sulphurs; right group: angle at a metal between two non-bridging sulphurs. The means of the groups are also given. The bridging sulphur atoms S_{b1}, S_{b2}, and S_{b3} belong to residues Cys24, Cys7, and Cys15, respectively.

are quite large. As there is no significant conformational change observed, there is also no progress towards the NMR structure (Fig. 6a).

The bond angle problems (Table IV) are mostly analogous to the MDc $CdZn_2$ case: The sulphur-bridge angles (M-S_b-M) are too large. The metal site 2 is heavily distorted: the ring-bridge angle is smaller than 90°, while the angle to the terminal sulphurs is larger than 150°. This raises the question whether the classical electrostatics is capable of maintaining a tetrahedral configuration around a relatively large ion.

The NMR structure exhibits a small number of hydrogen bonds (Table II), of which two are not reproduced by any simulation: Gln23-Gln23 NEx-HE2x-O and Thr27-Lys25 N-H-O. The first is an intraresidual hydrogen bond that depends on the side-chain conformation. However, the latter is a backbone hydrogen bond, as are two more, Cys26-Cys29 N-H-O and Lys30-Ser28 N-H-O, all three of which are not recovered in the Cd_3 MDc simulation. Out of the two Ser12-Asp10 experimental hydrogen bonds, one is reproduced in the Cd_3 MDc simulation at a low percentage. These are the only hydrogen bonds in the NMR structure that are also present in the X-ray structure. Hydrogen bonds established during the Cd_3 MDc simulation are generally similar to those in the $CdZn_2$ MDc simulation. A new hydrogen bond is formed at the beginning of the chain at Asp2. However, it seems that the character of the initial

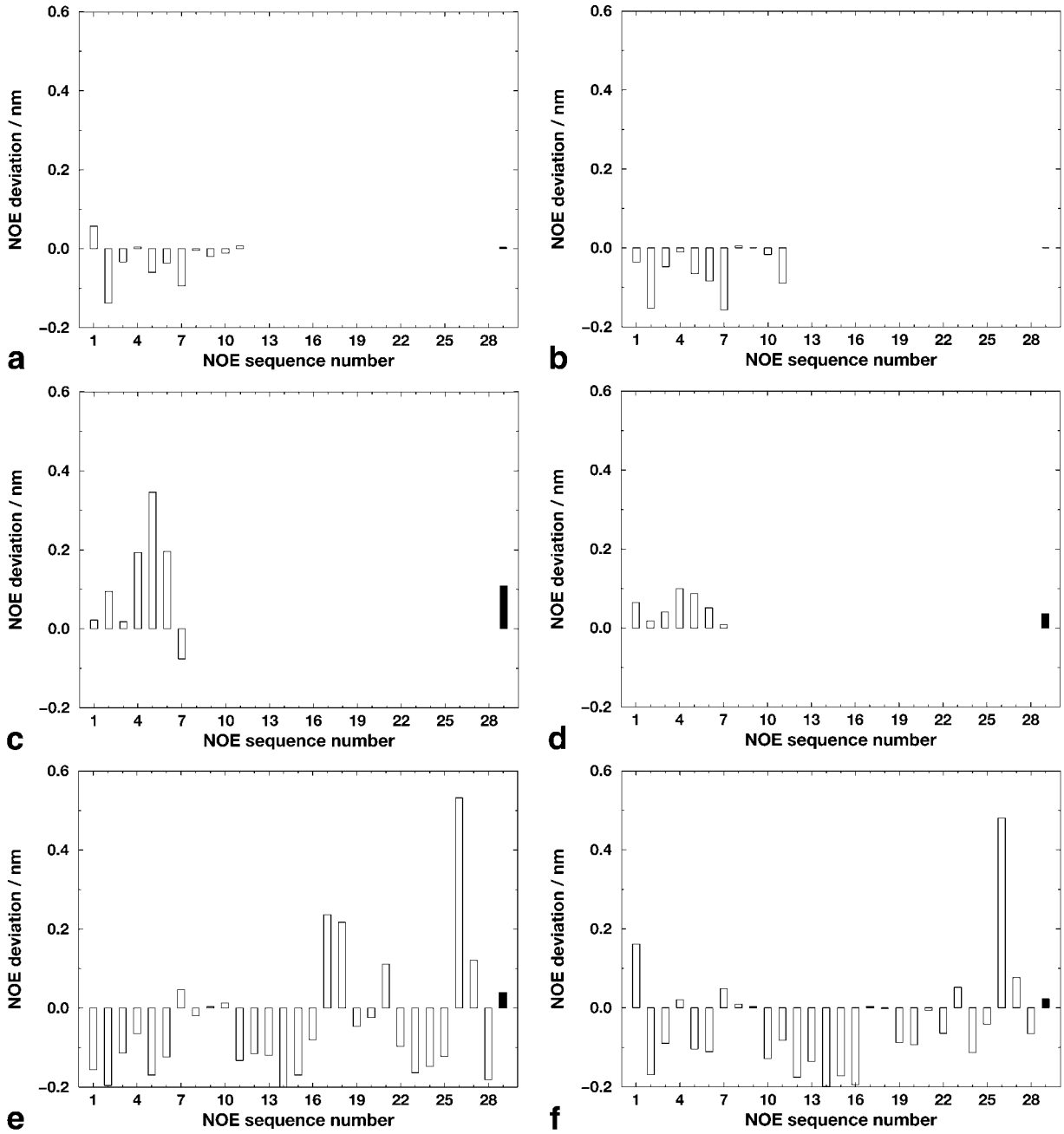

Fig. 7. Comparison of simulated NOE atom-atom distances with experimental[2] NOE bounds. The simulated NOE average distances are r^{-3} averages. Values greater than zero violate experimental data. **a:** sequential backbone bounds in MDc simulation; **b:** sequential backbone bounds in MDq simulation; **c:** medium-range backbone and long-range backbone bounds in MDc simulation; **d:** medium-range backbone and long-range backbone bounds in MDq simulation; **e:** interresidual bounds with side-chain protons in MDc simulation; **f:** interresidual bounds with side-chain protons in MDq simulations. NOE values appear only once; the second entry from the NOE list (table 1 in Schultze et al.[2]) was omitted. The sequence of the NOEs is the same as in table 1 of Schultze et al.[2] The solid black bar to the far right in each graph gives the average over all NOE violations.

(X-ray-based) structure is largely retained (see Fig. 3), except for Gln23-Asn4 N-H-O, which disappears, and Cys15-Ser28, which is rearranged.

The compatibility of the simulated trajectories with NOE distance limits derived from the experiment by Schultze et al.[2] Schultze et al.'s[2] table 1 groups the NOE values in three categories: (1) Sequential backbone bounds, (2) medium-range backbone and long-range backbone bounds, (3) interresidual bounds with side-chain protons. The NOE atom-atom distances were calculated from the trajectories using r^{-3} averaging.

Despite the large deviations in the MDc trajectories to the NMR structure (Fig. 6a), the sequential NOE bounds (panel a in Fig. 7) are quite well satisfied. However, there

TABLE V. List of NOE Bounds Violated by More Than 0.1 nm in Any Simulation of the Cd_3 Variant of the β Domain of Rat Liver Metallothionein-2[†]

NOE sequence number	Residue	Proton		Residue	Proton
Backbone medium-range and long-range NOE bounds					
4	Cys5	C_α	—	Cys21	C_α
5	Cys5	C_α	—	Gln23	Amide N
6	Cys5	C_α	—	Cys24	C_α
Side-chain interresidual NOE bounds					
1	Asp2	C_β (methylene)	—	Cys5	Amide N
17	Ala8	C_β (methyl)	—	Cys13	Amide N
18	Thr9	C_γ (methyl)	—	Asp10	Amide N
21	Lys20	C_β (methylene)	—	Cys21	Amide N
26	Thr27	C_α	—	Lys30	C_β (methylene)
27	Ser28	C_β (methylene)	—	Lys30	Amide N

[†]See also Figure 7.

are sizeable NOE bound violations (larger than 0.1 nm) in the other two categories (c,e in Figure 7). The details of the violated NOE bounds are listed in Table V.

The large violations of NOE bounds numbered 26 and 27 of the side-chain interresidual bounds are striking. However, these NOE bounds involve protons on Lys30 (the β-carbon protons and the amide proton). This residue is the linker to the second domain of metallothionein, which was omitted in the simulations. It is likely that these violations are artifacts of the omission of the second domain: It is the last residue retained in the β domain simulations and its structure may therefore be different from the one in the complete protein.

The interresidual NOE number 18 involves Thr9, which is located in the flexible loop from Thr9 to Ser14, in which already large differences between the X-ray and the NMR structure were observed. All other violated NOEs involve cysteine residues. The limited accuracy of the description of the metal-sulphur cluster configuration in the MDc simulation could be the origin of the NOE bound violations involving cysteine residues.

The average values for backbone ϕ angles are mostly in accordance with the ranges derived from experimental J-value couplings (table 2 in Schultze et al.[2]). However, the average ϕ angle value in Gln23 is 68° instead of in the range[2] of −170° to −70°. Average ϕ angle values for Ala8 and Cys29 are off by less than 10°.

Comparison of the Cd_3 MDq Simulation With NMR Data

Also in the MDq Cd_3 simulation, the cadmium-sulphur bond lengths are systematically too short, on average 0.013 nm shorter than in the NMR structure (Table III).

The RMS deviation from the initial structure is shown in Figure 5b. The deviation of the cadmium atoms is very low and the fluctuations very small. The deviation of the sulphurs drifts slowly upwards and exceeds the tolerance level of 0.05 nm. Cysteine C_α and C_β atoms deviate little from the initial structure, as do the backbone C_α atoms before 100 ps. Afterwards, there is an increase up to an RMSD of 0.3 nm, seemingly stabilising there.

Compared to the NMR structure (Fig. 6b), the results are similar, but the deviations are larger except for cadmium. Specifically, the increase of the backbone's C_α atoms in Figure 5b is not a progress towards the NMR structure, a similar increase is evident in Figure 6b, being even larger. Figure 3e shows that the amino terminus (Met1) turned around during the simulation, causing the mentioned increase. Figure 8 confirms that this is the primary cause by showing residue-wise RMS deviations for the C_α atoms from the NMR structure.

The bond angles (Table IV) of the sulphur bridges are considerably larger than in the NMR structure, by about 15° on average. In turn, one would expect that the metal-bridge angles would be too small. This is not the case: these angles match well, the simulation angles being on average only 3° larger than the NMR results. It should be noted that a tetrahedral symmetry on the sulphur and cadmium atoms was assumed in the structure derivation.[24]

The hydrogen-bonding situation (Table II) is similar to the MDc simulation. Some prominent new hydrogen bonds emerging during the simulation are Cys13-Ala8 N-H-O, Lys20-Ser6 NZ-HZx-O, and Gln23-Asn4 N-H-ODx. In contrast to the MDc simulation, one of the experimental hydrogen bonds, Lys30-Ser28 N-H-O, is reproduced by the MDq simulation at a low percentage. Several high-percentage hydrogen bonds occur that are also present in the X-ray structure, but not in the NMR structure.

The sequential NOE bounds (panel b in Fig. 7) are quite well satisfied. The differences between the MDc and the MDq simulations are minor. The backbone medium-range and long-range NOE bounds (c,d) are better represented in the MDq simulation. The same applies to the side-chain interresidual bounds (e,f). Only the NOE bounds involving Lys30 are still heavily violated, and one involving Asp2 located in the turning amino terminus (see Fig. 8). In total, the NOE bounds are very well satisfied using the MDq simulation method, although the structures in the trajectories are quite different from the NMR structure.

The MDq simulation satisfies most of the backbone ϕ angle ranges (table 2 in Schultze et al.[2]). However, the average values of the ϕ angle in Cys26 is 143° instead of in the range of −180° to −40°. The missing α domain induces

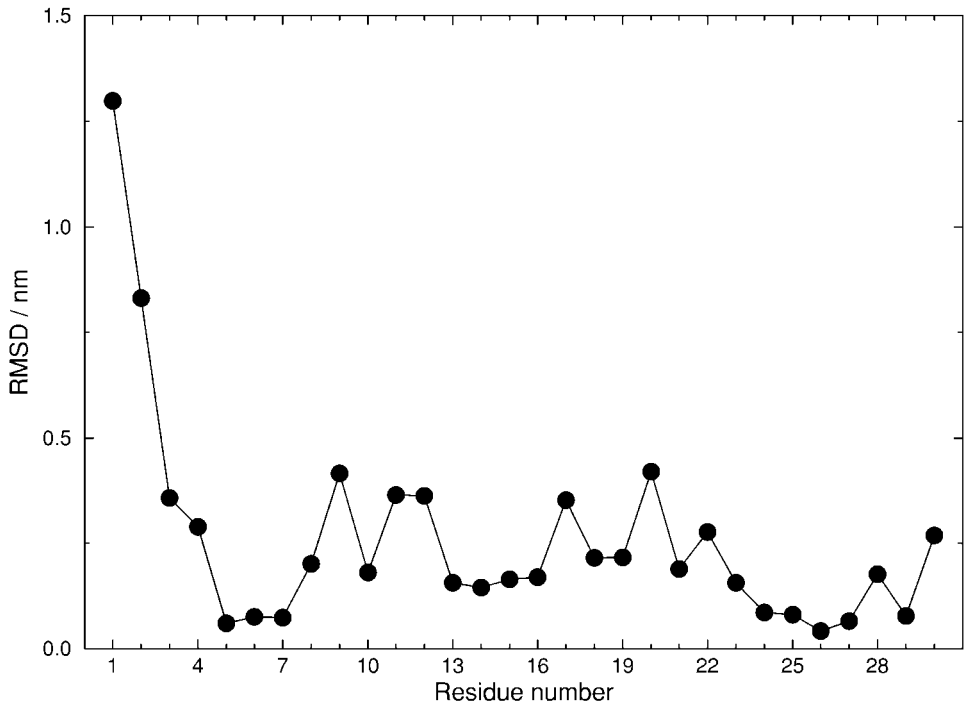

Fig. 8. Residue-wise RMS deviation of the final structure of the MDq Cd_3 simulation from the NMR structure.

again some perturbation at the chain end at Lys30, which has an average ϕ angle value of 78° instead of in the range from −180° to −60°.

Comparison of the Classical MDc and Quantum-Chemical MDq Simulations

In terms of the metal-sulphur bonds, the bond lengths are about 10% too short in the MDc simulations, whereas the MDq simulations yield smaller errors, the bond lengths still being systematically too short. However, the bond lengths in the X-ray and NMR structures are largely determined by refinement parameters. It is also not clear whether the bond lengths should be the same as in the corresponding metal sulfide minerals.

A pronounced difference between the MDc and the MDq simulations is exhibited in the metal-sulphur-metal bridges. The MDc simulations yield average M-S_b-M bridge angles of 136°, with all single angles larger than 130°. In contrast, the MDq simulations give an average of 114°, with all single angles except one being below 130°. The MDq values are much closer to the ideal tetrahedral angle of 109°. This clearly points at the deficiency of classical methods in describing metal-sulphur binding. With electrostatics only, thus neglecting the lone pair, the two metals would favour a linear configuration. This leads to the elongated M-S_b-M bridge angle. Without the aid of bonded interactions, such a situation cannot properly be described by the non-bonded interaction terms of a force field (van-der-Waals and electrostatic). One could think of adding bonded terms or virtual atoms to maintain the proper metal-sulphur geometry.[25] However, this would restrict the possibility of a substantial change of the geometry due to strong non-bonded forces.

Through the six-membered ring of metals and bridging sulphurs, the elongated angles at the sulphur bridges in the MDc simulations naturally lead to smaller ring bond angles at the metals (second column in every metal group in Table IV), which are in the range of 90° to 100°. In the MDq simulations, these bond angles are larger than 100° throughout. The angles at the metal centers are irregular and do not allow clear conclusions.

Comparing the three metal sites to each other, differences larger than the fluctuations are evident even when the sites contain the same metal. The short-range situation is equal for all sites: a metal ion tetrahedrally coordinated by four deprotonated cysteines. So the aforementioned differences must be caused by long-range interactions such as anisotropic charge distributions or steric hindrance by the protein backbone.

When comparing structural differences of the polypeptide backbone or NOE atom-atom distances, one should bear in mind the large difference in simulation lengths between the MDc and MDq simulations: 10 ns vs. 460 ps, respectively. Thus, the former are much better relaxed and equilibrated than the latter.

In terms of the RMS deviations (Figs. 5 and 6), the fluctuations for the metal and the sulphur atoms are significantly smaller in the MDq simulations. The quantum-chemical description seems to yield a more realistic potential-energy surface than the classical electrostatics combined with van-der-Waals terms. The crystal structure of the $CdZn_2$ variant is better represented in the MDq

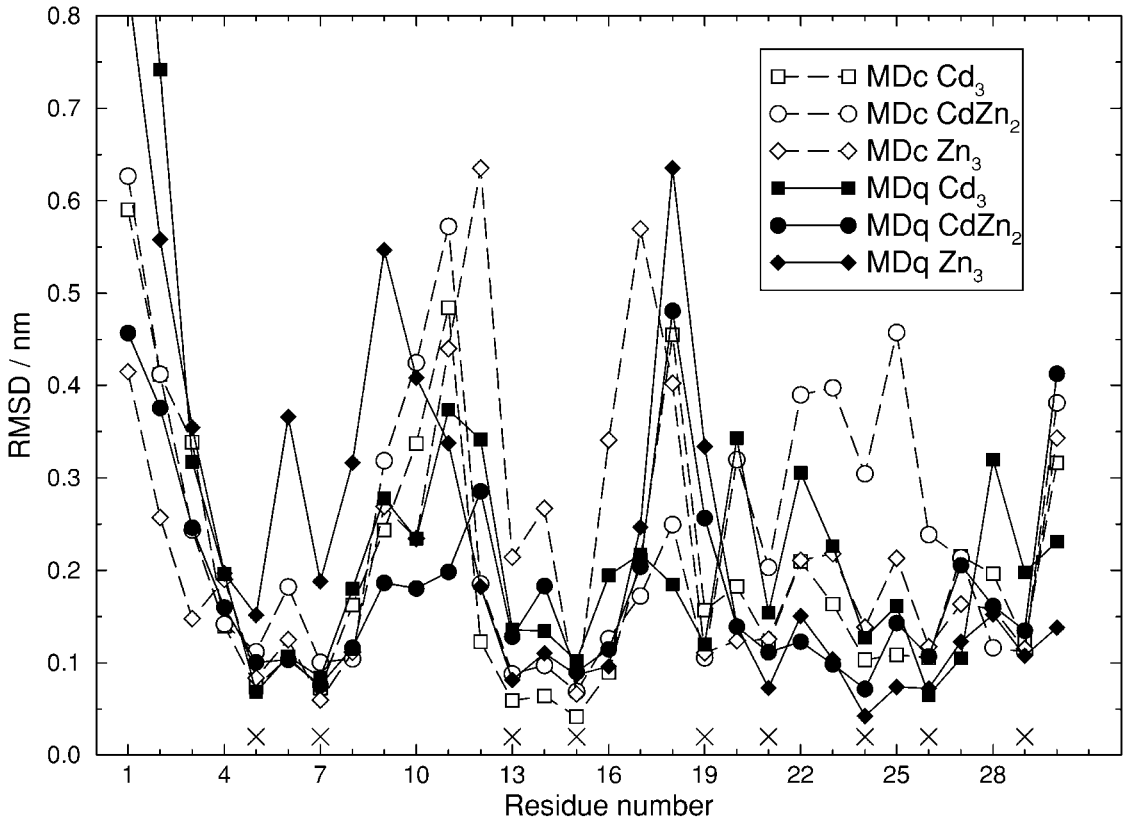

Fig. 9. C_α atom-positional RMS deviations per residue between the simulated trajectories and their initial structures, averaged over the entire simulations. The × symbols denote the cysteine residues.

simulation, giving mostly lower deviations than the MDc simulation (Fig. 6c,d).

The MDq simulation yields virtually no significant NOE bound violations, while the MDc simulation shows a handful violations. However, this difference could be due to the difference in averaging times.

Comparison of the Simulations of the Cd_3, $CdZn_2$, and Zn_3 Variants

The results for the Zn_3 variant are mostly in line with what was said previously for zinc in the $CdZn_2$ variant. Interestingly, the atom-positional RMS deviations of the C_α atoms in the MDc simulation (Fig. 5e) are smaller than for the $CdZn_2$ variant, although the initial structure is from the $CdZn_2$ variant, and the smaller zinc ions should induce a further contraction of the metal core. The deviations for the sulphur and zinc atoms are slightly higher, as expected.

The MDq simulation exhibits low deviations for the metals, and also the other atom types return below the tolerance thresholds after an initial disorder of about 100 ps, which is caused by motions of the N-terminus and the Thr10 loop. Changes in the metal types do not induce significant changes in the overall structure.

Figure 9 shows atom-positional RMS deviations per residue from the initial structures averaged over the whole simulations. The difference between residues is remarkable. The cysteines (see the × markers in Fig. 9) generally have a low deviation. The chain ends deviate clearly more, as is often observed in proteins. However, the loops between the cysteines exhibit an extraordinary flexibility, especially the loops around Gly11 and around Gly17, without disrupting the geometry of the metal core. The increased flexibility is present in all simulations and all variants. This supports the hypothesis (see Braun et al.[5] and citations therein) that the flexibility of the loops accounts for metallothionein's broad diversity in binding different metal ions.

It has recently been shown[26–31] that the GROMOS force field is capable of folding polypeptides into the correct secondary structure. The present work demonstrates that it is not a feature of the GROMOS force field to inevitably fold anything into a regular secondary structure: Metallothionein's variable loops remain irregular and flexible.

It is also appropriate to note that MNDO/d performs well on the metal cluster, despite the fact that it was not parametrised against zinc-sulphur or cadmium-sulfur bonds, neither against clusters involving several metal ions.

CONCLUSIONS

The classical MDc simulation of Cd_3 metallothionein-2 of rat liver in aqueous solution satisfies most of the NMR NOE data on this molecule; 8 atom-atom distance bounds

out of a total of 46 NOE bounds are violated by more than 0.1 nm. The structure of the metal-sulphur cluster is approximately maintained. Yet, the use of the semi-empirical method MNDO/d for the description of the metal core improved the simulated results. Specifically, bond lengths between the metal ions and the coordinated sulphur atoms are closer to experimental values. The MDc simulation yields a too compact metal core, thus affecting the whole protein. This is illustrated by the NOE bound violations. In the MDq simulation, in contrast to the MDc simulation, only one violation larger than 0.1 nm is observed.

The cost of the improved simulation is the increased computational expense. This is illustrated by Figures 5 and 6. Both the MDc and the MDq simulations ran simultaneously on the same type of computer. While the MDc simulations reached 10 ns, the MDq simulations only reached 460 ps. It should be kept in mind that this is partly due to the use of a shorter time step in the MDq simulations (0.5 vs. 2 fs, see Computational Details).

The MDc simulations indicate that the force-field parameters for zinc and cadmium would need improvement. Accidentally, the parameters derived for cadmium give bond lengths that would be appropriate for zinc. However, the current cadmium parameters are nevertheless unsuitable as improved parameters for zinc, as other zinc properties such as bond angles are not correctly reproduced by them. Moreover, a clear deficiency of a purely classical electrostatic and van-der-Waals description of the metal core shows up at the sulphur bridges, which tend to have too wide angles. Therefore, a proper description would require additional force-field terms for bond angles and maybe also bond lengths. However, it is not clear how these terms should look exactly, as the involved angles and bonds are hard to determine experimentally. The danger remains to choose "reasonable" parameters and, consequently, obtain "reasonable" results. In contrast, the treatment of the difficult core by quantum chemistry embedded in a classical environment provides an unbiased description, which is generalisable to other systems.

REFERENCES

1. Lippard SJ, Berg JJ. Principles of bioinorganic chemistry. Mill Valley CA: University Science Books, 1994.
2. Schultze P, Wörgötter E, Braun W, Wagner G, Vašák M, Kägi JHR, Wüthrich K. Conformation of [Cd_7]-metallothionein-2 from rat liver in aqueous solution determined by nuclear magnetic resonance spectroscopy. J Mol Biol 1988;203:251–268.
3. Robbins AH, McRee DE, Williamson M, Collett SA, Xuong NH, Furey WF, Wang BC, Stout CD. Refined crystal structure of Cd, Zn metallothionein at 2.0 Å resolution. J Mol Biol 1991;211:1269–1293.
4. Furey WF, Robbins AH, Clancy LL, Winge DR, Wang BC, Stout CD. Crystal structure of Cd, Zn metallothionein. Science 1986;231:704–710.
5. Braun W, Vašák M, Robbins AH, Stout CD, Wagner G, Kägi JHR, Wüthrich K. Comparison of the NMR solution structure and the X-ray crystal structure of rat metallothionein-2. Proc Natl Acad Sci 1992;89:10124–10128.
6. van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG. Biomolecular simulation: The GROMOS96 manual and user guide. Zürich: vdf Hochschulverlag AG an der ETH Zürich, 1996.
7. Gao J, Thompson MA. Combined quantum mechanical and molecular mechanical methods. ACS Symp Ser 1998;712.
8. Woo TK, Margl PM, Deng L, Cavallo L, Ziegler T. Towards more realistic computational modeling of homogenous catalysis by density functional theory: combined QM/MM and ab initio molecular dynamics. Catalysis Today 1999;50:479.
9. Thiel W, Voityuk AA. Extension of the MNDO formalism to *d* orbitals: integral approximations and preliminary numerical results. Theor Chim Acta 1992;81:391–404.
10. Thiel W, Voityuk AA. Extension of MNDO to d orbitals: parameters and results for the second-row elements and for the zinc group. J Phys Chem 1996;100:616–626.
11. Field MJ, Bash PA, Karplus M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. J Comput Chem 1990;11:700–733.
12. Christen HR. Grundlagen der allgemeinen und anorganischen Chemie. 9th ed. Frankfurt am Main: Salle + Sauerländer; 1988.
13. Eichinger M, Tavan P, Hutter J, Parrinello M. A hybrid method for solutes in complex solvents: density functional theory combined with empirical force fields. J Chem Phys 1999;110:10452–10467.
14. Eichler U, Kölmel CM, Sauer J. Combining ab initio techniques with analytical potential functions for structure prediction of large systems: method and application to crystalline silica polymorphs. J Comput Chem 1996;18:463–477.
15. Antes IS. Combined quantum mechanical and molecular mechanical methods: from link atoms to adjusted connection atoms. Ph.D. thesis, Universität Zürich, 1998.
16. Antes I, Thiel W. Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. J Chem Phys A 1999;103:9290–9295.
17. Bakowies D. Hybridmodelle zur Kopplung quantenmechanischer und molekülmechanischer Verfahren. Ph.D. thesis, Universität Zürich, 1994.
18. Bakowies D, Thiel W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. J Phys Chem 1996;100:10580–10594.
19. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. Intermolecular Forces, Dordrecht: Reidel, 1981; 331–342.
20. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 1977;23:327–341.
21. Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krüger P, van Gunsteren WF. The GROMOS biomolecular simulation program package. J Phys Chem A 1999;103:3596–3607.
22. Smith PE, van Gunsteren WF. Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K. J Chem Phys 1994;100:3169–3174.
23. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
24. Arseniev A, Schultze P, Wörgötter E, Braun W, Wagner G, Vašák M, Kägi JHR, Wüthrich K. Three-dimensional structure of rabbit liver [Cd_7]metallothionein-2a in aqueous solution determined by nuclear magnetic resonance. J Mol Biol 1988;201:637–657.
25. Pang YP. Novel zinc protein molecular dynamics simulations: steps toward antiangiogenesis for cancer treatment. J Mol Model 1999;5:196–202.
26. Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Reversible peptide folding in solution by molecular dynamics simulations. J Mol Biol 1998;280:925–932.
27. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide folding: when simulation meets experiment. Angew Chem Int Ed 1999;38:236–240.
28. Daura X, van Gunsteren WF, Mark AE. Folding-unfolding thermodynamics of a β-heptapeptide from equilibrium simulations. Prot Struct Funct Genet 1999;34:269–280.
29. Bonvin AMJJ, van Gunsteren WF. β-hairpin stability and folding: molecular dynamics studies of the first β-hairpin of tendamistat. J Mol Biol 2000;296:255–268.
30. Daura X, Mark AE, van Gunsteren WF. Peptide folding simulations: no solvent required? Comput Phys Commun 1999;123:97–102.
31. Bürgi R, Daura X, Mark A, Bellanda M, Mammi S, Peggion E, van Gunsteren WF. Folding study of an Aib-rich peptide in DMSO by molecular dynamics simulations. J Pept Res 2000; in press.