Protonation status of metal-bound ligands can be determined by quantum refinement

Nilsson, Kristina; Ryde, Ulf

Published in:
Journal of Inorganic Biochemistry

DOI:
10.1016/j.jinorgbio.2004.06.006

2004

Document Version:
Peer reviewed version (aka post-print)

Protonation status of metal-bound ligands can be determined by quantum refinement

Kristina Nilsson and Ulf Ryde *

Department of Theoretical Chemistry

Chemical Centre

Lund University

P. O. Box 124

S-221 00 Lund

Sweden

Abstract

The protonation status of key residues and bound ligands are often important for the function of a protein. Unfortunately, protons are not discerned in normal protein crystal structures, so their positions have to be determined by more indirect methods. We show that the recently developed quantum refinement method can be used to determine the position of protons in crystal structures. By replacing the molecular-mechanics potential, normally used in crystallographic refinement, by more accurate quantum chemical calculations, we get information about the ideal structure of a certain protonation state. By comparing the refined structures of different protonation states, the one that fits the crystallographic raw data best can be decided using four criteria: the R factors, electron density maps, strain energy, and divergence from the unrestrained quantum chemical structure. We test this method on alcohol dehydrogenase, for which the pK_a of the zinc-bound sol-

vent molecule is experimentally known. We show that we can predict the correct protonation state for both a deprotonated alcohol and a neutral water molecule.

Key words: crystallographic refinement, density-functional calculations, alcohol dehydrogenase, metal-bound solvent molecules, acid constants.

---

* Phone: +46-46 222 45 02, Fax +46-46 222 45 43
Email address: Ulf.Ryde@teokem.lu.se (Ulf Ryde).
URL: www.teokem.lu.se/∼ulf (Ulf Ryde).

1 Introduction

X-ray crystallography is the major source of structural information for large biomolecules, such as proteins. Unfortunately, the resolution typically obtained for proteins is fairly low, so some information is missing in the resulting structures. In particular, hydrogen atoms can normally not be discerned, except in the most accurate structures. This is unfortunate, because protons are involved in most reaction mechanism of enzymes. Therefore, a detailed knowledge of the positions of the protons in the structure would give a better understanding of the function of the protein. Today, such information has to be obtained by more indirect methods, e.g. by studying how the reaction rate depends on pH.

Another effect of the restricted resolution of protein crystal structures is that the positions of the atoms in the structure are not accurately known. Therefore, the data are normally supplemented by some sort of empirical information, typically in the form of a molecular-mechanics force field. This force field is used to ensure that bond lengths and angles are chemically reasonable and that aromatic systems are planar. Thus, for low- and medium resolution crystal structures, the general fold of the protein and the dihedral angles are determined by the experimental data, whereas the bond lengths are mainly determined by the molecular-mechanics force field.

Consequently, the quality of the resulting crystal structures will depend on the force-field used in the crystallographic refinement [1,2]. For standard amino acids and nucleic acids, accurate force fields exist, which are based on statistical analysis of small-molecule data [3]. However, for more unusual molecules, such as substrates, inhibitors, coenzymes, and metal centres, i.e. *hetero-compounds*,

experimental data are often partly lacking or are less accurate. In particular, force constants are not available and the force field has to be constructed by the crystallographer, a complicated and error-prone procedure.

Even with an accurate empirical force field, the atomic positions in protein structures are quite uncertain, with an average error in bond lengths of ∼10 pm [4,5] and appreciably larger errors are occasionally found [2]. This uncertainty contributes to the problem of determining the protonation status of various molecules in the crystal structure: Different protonation states of a molecule give rise to more or less pronounced differences in the bond lengths and angles of the surrounding atoms. This is especially evident for metal-bound water molecules, for which the metal–O bond length decreases by ∼30 pm if the water molecule is deprotonated. Thus, if the structure was accurate enough, it would be possible to deduce the positions of the protons by studying the geometry of the surrounding atoms. However, this would require detailed information of the ideal structure of the two protonation states in the environment encountered in the protein (the metal–O distance varies with the nature of the other ligands of the metal). Such information is normally not available.

A conceivable way to solve these problems is to replace the molecular-mechanics force field for the site of interest by more accurate quantum chemical calculations: Density functional calculations with a medium-sized basis set typically reproduce experimental bond lengths within 2 pm for organic molecules and within 0–7 pm for bonds to metal ions [6–9], making them more accurate than standard low- and medium-resolution crystal structures. We have recently developed such a method, *quantum refinement* [10], in which we replace the empirical force field for a small part of the protein in a standard crystallo-

graphic refinement by quantum chemical calculations. We have shown that it works properly and that it can be used to locally improve crystal structures of hetero-compounds, e.g. inhibitors and metal sites [9,10]. In this paper, we show that we can also use this method to determine the protonation state of metal-bound solvent molecules. Thus, we show that we can reproduce the correct protonation status of zinc-bound solvent molecules in two crystal structures of alcohol dehydrogenase, for which the protonation is known by kinetic experiments [11].

2 Methods

2.1 Quantum refinement

Quantum refinement [10,12] is essentially standard crystallographic refinement supplemented by quantum chemical calculations for a small part of the protein. Crystallographic refinement programs change the protein model (coordinates, occupancies, B factors, etc.) to improve the fit of the observed and calculated structure-factor amplitudes (usually estimated as the residual disagreement, the R factor). Owing to the limited resolution normally obtained for biomolecules, the experimental data are supplemented by chemical information, usually in the form of a molecular-mechanics (MM) force field [1]. Then, the refinement takes the form of a minimisation or simulated annealing calculation by molecular dynamics using an energy function of the form

$$E_{cryst} = w_A E_{Xray} + E_{MM}, \tag{1}$$

where E_{Xray} is a penalty function, describing how well the model agrees with the experimental data (we used the maximum-likelihood refinement target using amplitudes, MLF) [13,14]), E_{MM} is a MM energy function with bond, angles, dihedral, and non-bonded terms, and w_A is a weight factor, which is necessary because E_{MM} is in energy units, whereas E_{Xray} is in arbitrary units [15].

Quantum chemistry can be introduced in this function by replacing the MM potential for a small (but interesting) part of the protein (system 1) by a quantum mechanics (QM) calculation, yielding a QM energy for system 1, E_{QM1}. To avoid double counting we must then subtract the MM energy of system 1, E_{MM1}:

$$E_{tot} = w_A E_{Xray} + E_{MM} + w_{QM} E_{QM1} - E_{MM1}, \tag{2}$$

Thereby, we introduce an accurate energy function for the system of interest. Such a penalty function is implemented in the software ComQum–X [10], which is a combination of the softwares Turbomole [16] and Crystallography and NMR system (CNS) [17]. The factor w_{QM} in Eqn. 2 is another weight, which is needed because the CNS MM force field is based on a statistical analysis of crystal structures [3]. Therefore, the force constants are not energy-derived, as is the QM energy, but they are in arbitrary statistical units. Experience has shown that the CNS force constants are typically three times larger than energy-based force constants [3], and $w_{QM} = 3$ was therefore used throughout this work [10].

Special attention is needed if there is a covalent bond between system 1 and the surrounding protein. This is a well-known problem in the popular com-

bined QM and MM methods (QM/MM) [18–20] (ComQum–X, can be seen as a QM/MM method with restraints to crystallographic raw data), and a simple and robust solution [21] is to truncate the QM system with hydrogen atoms, the positions of which are linearly related to the corresponding carbon atom in the protein [10]. Of course, E_{MM1} is also calculated with these hydrogen atoms, so that artefacts introduced by the truncation may cancel out. Following crystallographic custom, hydrogen atoms and electrostatic interactions are ignored in the refinement (hydrogen atoms are of course present in the quantum chemical calculations).

ComQum–X has been tested by re-refining the structure of *N*-methylmesoporphyrin bound to ferrochelatase [10]. The results showed that we may improve the structure locally in terms of the R_{free} factor. Moreover, we have shown [9] that refinement with ComQum–X of a medium-resolution (170 pm) crystal structure of cytochrome c_{553} brings the geometry of the haem group and its ligands closer to that observed in an atomic-resolution structure (97 pm) of the same protein [22]. For example, the errors in the Fe–ligand distances are reduced from 3–9, 12, and 32 pm to 1, 0, and 2 pm (for the porphyrin, histidine, and methionine ligands, respectively).

2.2 Computational details

The protonation status was studied in two systems, viz. a trifluoroethanol or a water molecule bound to the active-site zinc ion in the complex of horse-liver alcohol dehydrogenase with NAD^+ (PDB protein databank entries 1axe and 1ju9, both at 200 pm resolution) [23,24]. The two structures were collected at a temperature of 277 and 100 K, respectively.

Coordinates, occupancies, B factors, and structure factors were downloaded from the Brookhaven protein databank. From these files, we also obtained the space group, unit-cell parameters, resolution limits, R factors, and the test set used for the evaluation of the R_{free} factor.

The QM system was $Zn(SCH_3)_2$(imidazole) with H_2O/OH^- or $CF_3CH_2OH/CF_3CH_2O^-$ bound, corresponding to the catalytic zinc ion, its ligands Cys–46, Cys–174, and His–67, and the solvent or inhibitor.

All QM calculations were performed with the density functional Becke–Perdew-86 method (BP86) [25,26], treating the Coulomb operators with the resolution-of-identity (RI) approximation [27,28]. The DZP basis set of Schäfer et al. [29] was used for metals and for all other atoms the 6-31G* basis set was used [30]. Since the interest of the present article is protonation states, it is conceivable that polarising functions on hydrogen atoms are important. Therefore, we repeated the calculations on the trifluoroethanol complex of alcohol dehydrogenase also with the 6-31G** basis set (but still with DZP on Zn). However, this led to changes in the geometries, energies, and R_{free} factor of less than 0.2 pm, 0.8 kJ/mole, and 0.00002, respectively. Thus, enlargement of the basis set would not affect the conclusions.

The choice of QM method is based on our previous experience that the BP86 method gives excellent metal–ligand distances [9,12]. For small organic molecules, other density functional methods are known to give better results, but the difference is small: For the G2 test set, the B3LYP method gives an average absolute error in the bond lengths of 1.3 pm, whereas BP86 gave 2.2 pm [31]. This is more than compensated for by the metal–ligand bond lengths. For example, for cytochrome models, B3LYP gives errors in the Fe–ligand

distances of 2–3, 4-5, and 6 pm (for porphyrin, histidine, and methionine, respectively), whereas the errors for BP86 are 1–3, 0–2, and 1–3 pm [9]. Similar results have been obtained for other metals [12]. We have optimised the $Zn(SCH_3)_2$(imidazole)(H_2O/OH^-) complexes also with the B3LYP method. For the optimum vacuum geometries, the bond lengths obtained with the two methods differ by less than 1.0 pm for the ligands (except for a single H–O bond of 1.5 pm), whereas the Zn–ligand distances differ by up to 2.8 pm. Therefore, we have preferred the BP86 method, which also is $\sim$5 times faster than B3LYP, owing to the RI approximation.

The whole protein was used in all calculations, including all crystal water molecules in the PDB files. The full geometry of the proteins was optimised until the change in E_{tot} was below 10^{-6} Hartree (2.6 J/mole) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u. In each cycle of the geometry optimisation, the surrounding protein was allowed to relax by one cycle of crystallographic minimisation and one cycle of individual B-factor refinement. However, the new coordinates and B factors were accepted only if the R factor was reduced. For the protein, we used the standard CNS force field (protein_rep.param, water_rep.param, and ion.param). For the other program parameters, we used data form the PDB files or the default choices. Residue (real-space) R factors [32] were calculated from σ_A-weighted maps using CNS. The presented values are the average of the factor for the zinc ion and its four ligands (full residues).

Finally, the w_A factor in Eqn. 2 need to be specified. In standard crystallographic refinements (e.g. in CNS), it is determined so that the MM and crystallographic forces have a similar magnitude [15]. However, this is a quite arbitrary choice and there is no warranty that it gives an optimum structure.

A better solution is select the value of w_A that gives a refined structure with the lowest R_{free} factor. We have used such an approach. Unfortunately, it turns out that the various protonation state sometimes have different optimum values of w_A. Then, it is important to compare only results obtained with the same value of w_A (except for R_{free}).

3 Results and Discussion

Alcohol dehydrogenase (EC 1.1.1.1) catalyses the reversible oxidation of alcohols to aldehydes or ketones using NAD^+ as a coenzyme [11]. The active site of the enzyme contains a catalytic zinc ion, which is bound by two cysteines, a histidine, and a substrate or a solvent molecule. From kinetic measurements on the horse-liver enzyme [11] the pK_a of the zinc-bound water molecule is known to be 9.2 when no coenzyme is bound, 7.6 in the complex with NAD^+ and 11.2 in the complex with NADH. The pK_a of alcohols is 1–2 units lower [11]. Therefore, we can use alcohol dehydrogenase to calibrate our method and see if we can predict the correct protonation status of metal-bound solvent molecules with quantum refinement (program ComQum–X). We have employed two different structures, one with a deprotonated alcohol, and the other with a neutral water molecule.

3.1 A structure with deprotonated alcohol

We started to study the complex between horse-liver alcohol dehydrogenase, NAD^+, and trifluoroethanol at 200 pm resolution [23]. In this complex, the alcohol should have a pK_a of ∼6 [11], which is well below the pH at which

the crystal was grown, 8.4. Thus, the complex should contain a deprotonated alkoxide ion. We have calculated the ComQum–X structures of this complex with both an alkoxide or an alcohol. The results obtained with three different values of the w_A factor (Eqn. 1) are shown in Table 1, viz. the default CNS value (1.77), a common value of 3, and the optimum value in terms of the R_{free} factor, which is 10 for the alcohol, but 0.1 for the alkoxide ion. **Tab 1**

We can see that ComQum–X improves the structure in terms of the R_{free} factor, which decreases from 0.239 to 0.228 in all structures. The standard R factor (not shown) indicates slightly smaller improvement, from 0.191 to 0.190. The residue (real-space) R factor [32] (for the residues included in the quantum system) shows an improvement from 0.087 to 0.084–0.086 in all structures, except those with the smallest value of w_A. This illustrates that the residue R factor strongly depends on the w_A (the smaller w_A is, the stronger are the empirical restraints and the lower is the weight of the crystallographic data, giving a higher value of the R factors) and can only be compared for calculations performed with the same value of w_A.

Most importantly, the results show that the alkoxide fits the experimental data better by at least four criteria. First, the alkoxide gives the lowest value for the R_{free} factor, both when calculated with the same value of w_A (3 or 1.77) or with the optimum value of w_A, 10 for the alcohol and 0.1 for the alkoxide. However, the difference is not very large, 0.0004. This reflects that R_{free} is a global factor, quite insensitive to local changes in the structure [2,10]. This is the reason why we have also studied the residue R factor, which is more sensitive to local changes. It gives a lower value for the alkoxide at $w_A = 1.77$ (0.086 compared to 0.088), but the same value at $w_A = 3$.

Second, the alkoxide gives a lower strain energy (ΔE_{QM1}) than the alcohol for all values of w_A. This energy is the difference in the quantum chemical energy of the quantum system in vacuum, calculated for the ComQum–X structure and the optimum vacuum structure. Thus, it is a measure of how well the quantum system fits into the crystallographic raw data (how much the active site must distort to fit into the density). The results clearly show that the alkoxide fits better into the electron density than the alcohol.

Third, the Zn–O distance in the alkoxide structures (190–191 pm) is close to that found in the vacuum calculation (193 pm) at all values of w_A, whereas for the alcohol, the Zn–O distance (201–207 pm) is far from the vacuum value (229 pm) and actually converges to the vacuum value of the *alkoxide* complex. This clearly indicates the Zn–O bond length preferred by the crystal data is closer to that expected for the alkoxide than that for the alcohol. It is important to note that this could not be decided from the original crystal structure, in which the Zn–O distance is intermediate between the optimum distance of an alcohol and an alkoxide, 205–207 pm.

Fourth, we can compare how well the ComQum–X structures of the alcohol and alkoxide complexes fit to the electron-density maps. In Figure 1, we visualise the structures of the protonated and deprotonated alcohol. It can be seen that the latter gives the best $f_o - f_c$ difference maps (the blue and green volumes are appreciably smaller than the yellow and red volumes), showing an improvement especially around the alcohol. The figure also shows the difference in geometry between the two protonation states of trifluoroethanol: There are considerable changes in the position of all atoms in trifluoroethanol, but also of the atoms in the histidine ligand. **Fig 1**

In conclusion, we see that all four criteria unambiguously point out the deprotonated alkoxide as the correct structure, in accordance with the kinetic data. This shows that the protonation status of a metal-bound solvent molecule in a crystal structure can be determined by quantum refinement.

3.2 A structure with neutral water

In order to check that ComQum–X does not have a bias for deprotonated structures, we also studied a structure with a neutral zinc-bound water molecules. It was quite hard to find such a structure (with deposited structure factors) in the PDB data base. The best candidate was a complex of alcohol dehydrogenase and NAD^+ at 200 pm resolution, which has a water molecule bound to the zinc ion [24]. The crystal was grown at pH 7.0, which is slightly below the experimental pK_a of ∼7.6 for the zinc-bound water molecule in this NAD^+ complex. Thus, the crystal should contain mainly bound H_2O. We have calculated the ComQum–X structures of both a hydroxide ion or a water molecule bound to this crystal at three values of the w_A factor, the optimum value (300 for both structure), 30, and the default CNS value, 1.25 (this value depends on the forces in the system [15]; therefore, it not the same as in the trifluoroethanol complex).

The results in Table 2 are quite similar to those of the trifluoroethanol structure. The R_{free} factor decreases in all ComQum–X calculations (except one with OH^-, the incorrect protonation state), but only marginally (from 0.240 to 0.239). Those with H_2O always have a lower R_{free} than those with OH^- (by 0.0003–0.0011). The same applies to the residue R factor: It is smaller for water than for OH^-, by 0.002–0.005, except for the structure at $w_A = 300$. **Tab 2**

Likewise, water gives a lower strain energy than the hydroxide ion for all values of w_A. At $w_A = 1.25$, the difference is 12 kJ/mole, but at $w_A = 300$, the difference is as much as 124 kJ/mole. This shows that the OH^- structure is strongly unfavourable and that water fits the electron density appreciably better than OH^-.

The Zn–O distance in the water structure (222–227 pm) is at all values of w_A close to that found in the optimal vacuum structure (227 pm). On the other hand, the Zn–OH^- distance (198–227 pm) is always far from the vacuum value. In particular, it converges to the vacuum distance of the *water* complex for high values of w_A. This clearly indicates the Zn–O bond length preferred by the crystal data is closer to that expected for water than that for OH^-. Once again, this could not be told from the original crystal structure, where the Zn–O distances in the two subunits are 208 and 216 pm.

Finally, electron-density difference maps confirm these results, as is shown in Figure 2: The deprotonated structure (red and yellow) gives rise to larger volumes than the water structure (green and blue), especially between the zinc ion and the solvent molecule.

Fig 2

3.3 Concluding remarks

We have presented a new method to determine the protonation status of important ligands bound to a protein. The applications on alcohol dehydrogenase show that we can reproduce experimentally known protonation states, using four different criteria: the cystallographic R factors, electron-density maps, strain energies, and the difference in the metal–O distance between the op-

timum vacuum structure and the re-refined crystal structure. The latter two criteria seem to be strongest and most easily interpreted, because they give the same results for all values of w_A tested. The R_{free} factor also gives the same results for all structures, but the differences are small in absolute terms. The residue R factor gives slightly larger differences, but the results vary more with w_A. Moreover, they are sensitive to how they are calculated (what electron-density map, which residues are included in the calculation, etc.).

Our method is based on a comparison of structures refined assuming different protonation states in the empirical potential. In principle, any type of empirical potential could have been used, provided that it is accurate enough to describe the structural differences caused by changes in the protonation. Thus, a normal molecular-mechanics potential could also have been used, for example based on accurate small-molecule crystallographic data or extracted from accurate quantum chemical calculations [2]. However, by using density functional calculations, we obtain a fully automatic method.

Of course, such a choice has a strong influence on the size of the system that can be treated. With normal computer resources and proper basis sets, up to $\sim$100 atoms can be included in these calculations. However, as has been shown in this paper, this is normally enough, because changes caused by the protonation are quite local. Still, it should be noted that the present calculations are quite time consuming compared to standard crystallographic refinements; each quantum-refinement calculation took one or two weeks on a single personal computer (but subsequent calculations with different w_A values normally took only a few days, because the changes are then quite restricted). Thus, these calculations are not intended to be used during the early phase of the refinement, but only at the end, when the general structure of the protein

is settled and only details of the active site is of interest.

There have been numerous approaches to estimate the pK_a values of various groups in protein with theoretical methods, even for metal-bound ligands [33–43]. However, the results are varying and often quite poor. In particular, it has been hard to obtain reliable results for groups that are buried in the protein. The present approach is very different from the previous ones, because we study only the geometry of the protonated and deprotonated group and compare it to experimental data (a crystal structure). All the previous methods have tried to directly estimate the acid constant by looking at the proton affinity and how it is modified by the surroundings. Of course, the various methods have their specific advantages and specific areas of application. Our method avoids the problem of describing the electrostatic surroundings of the interesting group in a balanced and accurate way, and it is directly applicable to buried groups. On the other hand, it does not give any estimate of the pK_a values, only of the dominating protonation state at the particular conditions in the crystal structure.

In this paper we have shown that we can determine the protonation status of metal ligands by quantum refinement. Of course, the same method could also be used to systems without metal ions, e.g. for the protonation of amino acids, substrates, or inhibitors. However, for such groups, the change in geometry upon protonation is much smaller than for a metal ligand. Moreover, it is less clear which distances (or angles) will change. Therefore, it becomes harder to decide the protonation state. We have tested the method on the protonation of the inhibitor *N*-methylmesoporphyrin, bound to the enzyme ferrochelatase at 190 pm resolution [44]. Unfortunately, it was hard to decide the protonation status, because various criteria pointed in different directions.

Finally, it should also be noted that our method is not restricted to the determination of protonation states. It can also be used to locally improve the structure of the active site [9]. Moreover, similar methods could be used to determine the oxidation state of metal ions. They are often not known, because they may change during preparation and data collection. Thus, in general, quantum refinement can be used to interpret what exactly is seen in the crystal structure, by comparing how well various interpretations fit the crystallographic raw data. Thus, we may conclude that quantum refinement is a powerful tool to determine the detailed structure of the active site of a protein. We predict that it may become a standard tool, once the computers have become faster and quantum chemical methods have become more widely spread in the chemical community. We currently work with the application of this method on a number of interesting proteins, e.g. myoglobin, peroxidase, superoxide dismutase, Ni–Fe hydrogenase, and aminopeptidase P.

4 Acknowledgements

This investigation has been supported by grants from the Crafoord foundation and the Swedish research council (VR). It has also been supported by computer resources of Lunarc at Lund University. We thank Prof. G. J. Kleywegt for fruitful discussions.

5 Supplementary material

The final re-refined structures of alcohol dehydrogenase with deprotonated trifluoroethanoland with water are included as supplementary material.

References

[1] G. J. Kleywegt, T. A. Jones, Databases in protein crystallography, Acta Crystallogr. D54 (1998) 1119–1131.

[2] K. Nilsson, D. Lecerof, E. Sigfridsson, U. Ryde, An automatic method to generate force-fild parameters for hetero-compounds, Acta Crystallogr. D59 (2003) 274–289.

[3] R. A. Engh, R. Huber, Accurate bond and angle parameters for x-ray protein structure refinement, Acta Crystallogr. A47 (1991) 392–400.

[4] D. W. J. Cruickshank, Remarks about protein structure precision, Acta Crystallogr. D55 (1999) 583–601.

[5] B. A. Fields, H. H. Bartsch, H. D. Bartunik, F. Cordes, J. M. Guss, H. C. Freeman, Accuracy and precision in protein crystal structure analysis: two independent refinements of the structure of poplar plastocyanin at 173 K, Acta Crystallogr. D50 (1994) 709–730.

[6] F. Jensen, Introduction to computational chemistry, John Wiley & Sons, Chichester, 1999.

[7] E. Sigfridsson, M. H. M. Olsson, U. Ryde, A comparison of the inner-sphere reorganisation energies of cytochromes, iron–sulphur clusters, and blue copper proteins, J. Phys. Chem. B 105 (2001) 5546–5552.

[8] M. H. M. Olsson, U. Ryde, Geometry, reduction potential, and reorganisation energy of the binuclear cu_a site, studied by theoretical methods, J. Am. Chem. Soc. 123 (2001) 7866–7876.

[9] U. Ryde, K. Nilsson, Quantum chemistry can improve protein crystal structures locally, J. Am. Chem. Soc. 125 (2003) 14232–14233.

[10] U. Ryde, L. Olsen, K. Nilsson, Quantum chemical geometry optimisations in proteins using crystallographic raw data, J. Comput. Chem. 23 (2002) 1058–1070.

[11] G. Pettersson, Liver alcohol dehydrogenase, CRC Crit. Rev. Biochem. 21 (1987) 349–389.

[12] U. Ryde, K. Nilsson, Quantum refinement – a combination of quantum chemistry and protein crystallography, J. Mol. Struc. (Theochem) 632 (2003) 259–275.

[13] N. S. Pannu, R. J. Read, Improved structure refinement through maximum likelihood, Acta Crystallogr. A52 (1996) 659–668.

[14] P. D. Adams, N. S. Pannu, R. J. Read, A. T. Brunger, Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement, Proc. Nat. Acad. Sci. 94 (1997) 5018–5023.

[15] A. T. Brünger, Crystallographic refinement by simulated annealing: methods and applications, Meth. Enzym. 277–269 (1997) 243.

[16] R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Haase, M. Häser, H. Horn, C. Huber, C. Kölmel, M. Kollwitz, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, H. Weiss, TURBOMOLE Version 5.6, Universität Karlsruhe, Germany (2000).

[17] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse.Kunstleve, J.-S. Jiang, J. Kuszewski, M. Niges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, Crystallography & NMR System CNS, Version 1.1, Yale University (2000).

[18] G. Monard, K. M. Merz, Combined quantum mechanical/molecular mechanical

methodologies applied to biomolecular systems, Acc. Chem. Res. 32 (1999) 904–911.

[19] A. J. Mulholland, The QM/MM approach to enzymatic reactions, in: L. A. Eriksson (Ed.), Theoretical biochemistry – Processes and properties of biological systems (Theoretical and computational chmistry, Vol. 9), Elsevier Science, Amsterdam, 2001, pp. 597–505.

[20] U. Ryde, Combined quantum and molecular mechanics calculations on metalloproteins, Curr. Opin. Chem. Biol. 7 (2003) 136–142.

[21] R. M. Nicoll, S. A. Hindle, G. MacKenzie, I. H. Hillier, N. A. Burton, Quantum mechanical/molecular mechanical methods and the study of kinetic isotope effects: modelling the covalent junction region and application to the enzyme xylose isomerase, Theor. Chim. Acta 106 (2001) 105–112.

[22] S. Benini, A. González, W. R. Rypniewski, K. S. Wilson, J. J. Van Beeumen, S. Ciurli, Crystal structure of oxidized bacillus pasteurii cytochrome c_{553} at 0.97-å resolution, Biochemistry 39 (2000) 13115–13126.

[23] B. J. Bahnson, T. D. Colby, J. K. Chin, B. M. Goldstein, J. P. Klinman, A link between protein structure and enzyme catalyzed hydrogen tunneling, Proc. Nat. Acad. Sci. 94 (2000) 12797–12802.

[24] J. K. Rubach, S. Ramaswamy, B. V. Plapp, Contributions of valine-292 in the nicotinamide binding site of liver alcohol dehydrogenase and dynamics to catalysis, Biochemistry 40 (2001) 12686–12694.

[25] A. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys. Rev. A 38 (1988) 3098–3100.

[26] J. P. Perdew, Density-functional approximation for the correlation energy of the inhomogeneous electron gas, Phys. Rev. B 33 (1986) 8822–8824.

[27] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Chem. Phys. Lett. 240 (1995) 283–290.

[28] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, Theor. Chim. Acta 97 (1997) 119–124.

[29] A. Schäfer, H. Horn, R. Ahlrichs, Fully optimized contracted gaussian basis set for atoms Li to Kr., J. Chem. Phys. 97 (1992) 2571–2577.

[30] W. J. Hehre, L. Radom, P. v. R. Schleyer, J. A. Pople, Ab initio molecular orbital theory, Wiley-Interscience, New York, 1986.

[31] C. W. Bauschlicher, A. Ricca, H. Partridge, S. R. Langhoff, Chemistry by density functional theory, in: D. P. Chong (Ed.), Recent advances in density functional methods (Part II), World Scientific Publishing Co, London, 1997, pp. 165–227.

[32] T. A. Jones, J.-Y. Zou, W. E. Cowand, M. Kjeldgaard, Acta Crystallogr. A47 (1991) 110–119.

[33] G. Pettersson, H. Eklund, Electrostatic effects of bound NADH and NAD+ on ionizing groups in liver alcohol dehydrogenase, Eur. J. Biochem. 165 (1987) 157–161.

[34] M. K. Gilson, B. Honig, Proteins: Struc. Funct. Genet. 4 (1988) 7–18.

[35] A. Warshel, J. Åqvist, Annu. Rev. Biophys. Biophys. Chem. 20 (1991) 267–298.

[36] D. Bashford, Curr. Opin. Struct. Biol. 1 (1991) 175–184.

[37] B. Honig, A. Nicholls, Science 268 (1995) 1144–1149.

[38] C. L. Fisher, J.-L. Chen, J. Li, D. Bashford, L. Noodleman, Density-functional and electrostatic calculations for a model of a manganese superoxide dismutase active site in aqueous solution, J. Phys. Chem. 100 (1996) 13498–13505.

[39] T. Kesvatera, B. Jönsson, E. Thulin, S. Linse, Ionization behaviour of acidic residues in calbinding d9k, Proteins: Struc. Funct. Genet. 37 (1999) 106–115.

[40] J. E. Nielsen, K. V. Andersen, B. Honig, R. W. W. Hooft, G. Klebe, G. Vriend, R. C. Wade, Improving macromolecular electrostatics calculations, Protein Engineering 12 (1999) 657–662.

[41] T. Simonson, Macromolecular electrostatics: continuum models and their growing pains, Curr. Opin. Struct. Biol. 11 (2001) 243–252.

[42] C. N. Schutz, A. Warshel, Proteins: Struc. Funct. Genet. 44 (2001) 400–417.

[43] C. A. Fitch, D. A. Karp, K. K. Lee, W. E. Stites, E. E. Lattman, B. Garcia-Moreno, Experimental pka values of buried residues: Analysis with continuum methods and role of water penetration, Biophys. J. 82 (2002) 3289–3304.

[44] K. Nilsson, Quantum chemical interpretation of protein crystal structures, Ph.D. Thesis, Lund University.

Table 1

The ComQum–X results for the catalytic zinc ion in the complex of alcohol dehydrogenase with NAD^+ and trifluoroethanol. Zn–ligand distances, strain energies (ΔE_{QM1}), and the R factors are calculated with trifluoroethanol modelled either by a protonated alcohol (ROH) or a deprotonated alkoxide (RO^-). For comparison, the original crystal structure [23] and the results of vacuum optimisations are also included.

Ligand	w_A	ΔE_{QM1}	Distance to Zn (pm)[a]				R_{free}	residue R
		kJ/mole	N	Cys–46	Cys–174	O		
	Crystal		213–220	222–229	213–229	205–207	0.2390	0.087
ROH	Vacuum	0.0	209	229	225	229		
RO^-	Vacuum	0.0	224	233	233	193		
ROH	10	118.8	8.9	8.0	8.9	–28.3	0.2283	0.084
ROH	3	92.9	5.9	4.4	6.4	–24.5	0.2283	0.086
RO^-	3	61.6	3.5	2.1	3.1	–2.2	0.2280	0.086
ROH	1.77	86.7	4.9	2.7	5.4	–21.8	0.2284	0.088
RO^-	1.77	59.8	3.9	0.9	2.7	–2.2	0.2280	0.086
RO^-	0.1	53.7	0.7	–2.2	–0.3	–1.0	0.2278	0.103

[a] For the crystal and vacuum structures, the actual bond lengths are given. For the ComQum–X structures, the deviation from the corresponding vacuum structure is listed instead.

Table 2

The ComQum–X results for the catalytic zinc ion in the complex of alcohol dehydrogenase and NAD^+. Zn–ligand distances, strain energies (ΔE_{QM1}), and R factors are calculated with the zinc-bound solvent molecule modelled either by H_2O or OH^-. For comparison, the original crystal structure [24] and the results of vacuum optimisations are also included.

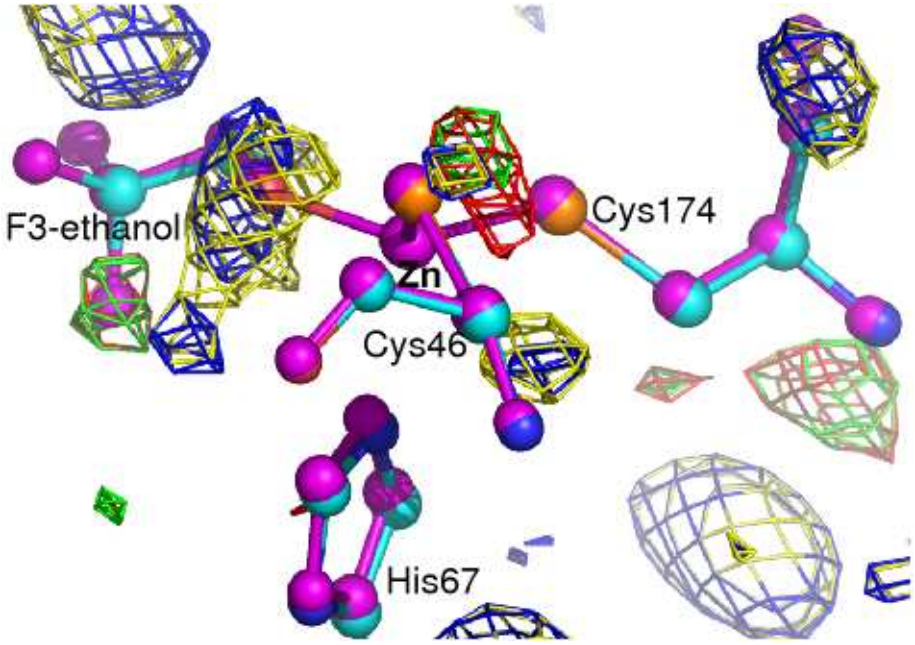
Ligand	w_A	ΔE_{QM1}	Distance to Zn (pm)[a]				R_{free}	residue R
		kJ/mole	N	Cys–46	Cys–174	O		
	Crystal		205–210	241–245	232–240	208–216	0.2404	0.081
H_2O	Vacuum	0.0	207	226	227	227		
OH^-	Vacuum	0.0	229	236	237	189		
H_2O	300	192.0	16.2	11.8	3.1	–0.4	0.2387	0.072
OH^-	300	315.8	–5.2	3.0	–6.8	38.2	0.2390	0.071
H_2O	30	73.5	10.0	10.5	3.5	–0.8	0.2389	0.073
OH^-	30	126.7	–11.5	1.6	–6.0	36.3	0.2400	0.075
H_2O	1.25	60.5	2.3	4.5	1.3	–5.4	0.2390	0.077
OH^-	1.25	72.5	–17.2	–2.4	–5.6	9.0	0.2404	0.082

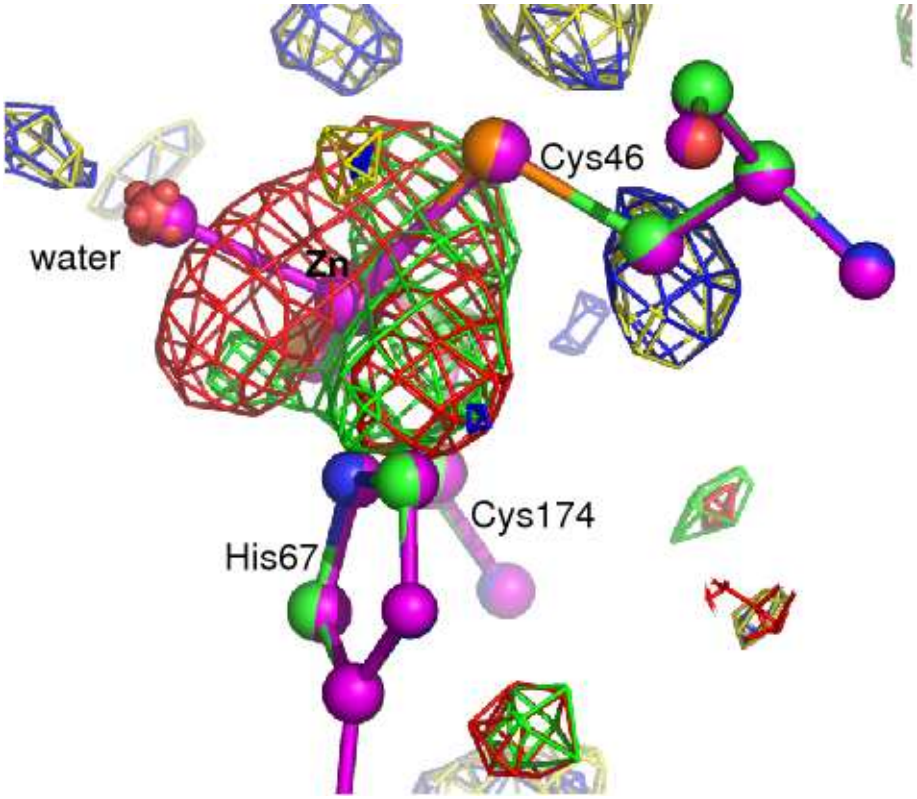
[a] For the crystal and vacuum structures, the actual bond lengths are given. For the ComQum–X structures, the deviation from the corresponding vacuum structure is listed instead.

6 Figure Legends

Figure 1. The structure of the catalytic Zn ion in alcohol dehydrogenase in complex with NAD^+ and trifluoroethanol. Structures with a protonated (magenta) and a deprotonated alcohol are compared, together with the corresponding electron-density $f_o - f_c$ difference maps at the $\pm 2.7\sigma$ level (yellow and red for the protonated alcohol; blue and green for the deprotonated alkoxide). The figure is based on structures obtained with $w_A = 1.77$; maps obtained at $w_A = 3$ show the same trends.

Figure 2. The structure of the catalytic Zn ion in the alcohol dehydrogenase–NAD^+ complex. Structures with H_2O and OH^- (magenta) are compared, together with the corresponding electron-density $f_o - f_c$ difference maps at the $\pm 2.5\sigma$ level (yellow and red for OH^-; blue and green for H_2O). The figure is based on structures obtained with $w_A = 1.25$; maps obtained at $w_A = 30$ and 300 show the same trends.


F3-ethanol
Zn
Cys174
Cys46
His67


Cys46
water
Zn
Cys174
His67