Karin Hauser
Institut für Biophysik,
Johann Wolfgang
Goethe-Universität Frankfurt,
Max-von-Laue-Strasse 1,
60438 Frankfurt am Main,
Germany

Electrostatic Calculations for Assignment of Infrared Difference Bands to Carboxyl Groups Getting Protonated During Protein Reactions

Abstract: Fourier transform infrared (FTIR) difference spectroscopy is predestinated to monitor the protonation of carboxyl groups during protein reactions, making glutamic and aspartic amino acids unique to follow proton pathways. The absorption of the corresponding vibrations are clearly distinguishable from the absorption of other amino acids. However, the assignment to specific groups within the protein needs additional information, e.g., from induced spectral changes due to isotopic labeling or mutation. Here, the capability of electrostatic calculations to assign IR difference bands to specific carboxyl groups getting protonated is demonstrated by the ion pump mechanism of the sarcoplasmic reticulum Ca^{2+} ATPase. Active Ca^{2+} transport is coupled to the hydrolysis of ATP. Two Ca^{2+} ions are transported per ATP hydrolysed and two or three H^+ ions are counter transported. FTIR difference spectra show that during the Ca^{2+} release step, carboxyl groups become protonated. Multiconformation continuum electrostatic calculations (MCCE) have been carried out to determine the equilibrium distribution of residue ionization and side chain conformation in dependence of pH. Available structural X ray data from the calcium bound and the calcium free state allows us to simulate the transition between the two states monitored in the IR difference spectra. Exemplarily for Asp 800, ligand of both calcium ions, it is shown that MCCE calculations can identify this specific Asp to contribute to the IR bands and therefore to take part in the proton countertransport of the Ca^{2+} ATPase. In addition, an energy analysis can be performed to understand what interactions shift the pK_a .

Keywords: multiconformation continuum electrostatics; infrared difference spectroscopy; protonation; carboxyl groups; Ca^{2+} ATPase

INTRODUCTION

Many protein reactions have been successfully studied by infrared difference spectroscopy. Difference spectra are built between two protein states of the reaction and have the advantage of containing contri-

butions just of those molecular groups that undergo changes during the transition from one protein state to the other. Often, there is only a small part of the protein actually involved in the reaction. Therefore, the strong absorbance background of nonchanging groups is subtracted, which leads to an enormous

reduction of infrared bands in the spectra and enables high sensitivity. Monitoring the protonation of carboxyl groups during protein reactions is especially feasible for infrared difference spectroscopy: the C=O stretching vibration of carboxylic amino acid residues, R-COOH, absorbs at $1700\text{--}1780\text{ cm}^{-1}$, a frequency region where no other protein bands are detected other than either protonation or environmental changes of the carboxyl group.^{1,2} The corresponding deprotonated carboxylate, R-COO⁻, absorbs in much lower frequency ranges. Aspartic and glutamic acids are therefore unique for studying proton pathways during protein activity by infrared spectroscopy. However, band assignment to specific residues often requires additional information, e.g., from studies with mutations, labeling, model compounds, or calculations.

In this report, the capability of electrostatic calculations to provide complementary information for interpretation of infrared data shall be demonstrated by the enzyme reaction of the sarcoplasmic reticulum Ca²⁺-ATPase (for reviews, see, e.g., Refs. 3 and 4). The Ca²⁺-ATPase transports calcium in the lumen of the sarcoplasmic reticulum, thereby causing the relaxation of muscle cells. Recently, high-resolution X-ray structures from different enzyme states of the Ca²⁺-ATPase have been determined.^{4–8} These structures allow its accessibility for computational analysis and push the Ca²⁺-ATPase to become the representative among the P-type ATPases for understanding the ion transport reaction in molecular details. P-type ATPases use the energy of ATP hydrolysis to drive active ion transport across the membrane. The reaction cycle is commonly interpreted according to the classical E1/E2 model assuming two major conformational enzyme states, E1 and E2. Ion binding and enzyme phosphorylation alternate with ion transfer and enzyme dephosphorylation. In the case of the Ca²⁺-ATPase, 2 Ca²⁺ ions are transported per ATP hydrolyzed and 2–3 protons are counter-transported (Figure 1). The proton countertransport is

FIGURE 1 Basic reaction cycle of the ion transport mechanism of the Ca²⁺ ATPase with important intermediate states (forward direction only).

FIGURE 2 Calcium binding site of the Ca²⁺ ATPase. M4, M5, M6 and M8 are transmembrane helices. Side chain oxygen ligands are represented by stick models. Underlined are the labels of those ligands that provide carboxyl groups.

well accepted; nevertheless the participating residues as well as the proton pathway remain unclear. Proton countertransport stops when the luminal pH is higher than 8.0.⁹ Ca²⁺ and H⁺ ions have been indicated to compete for the same sites of the Ca²⁺-ATPase¹⁰ and IR difference spectroscopy could show that carboxyl groups get protonated during the calcium release step.¹¹ Thus, calcium ligands providing carboxyl groups (Figure 2) become good candidates for being involved in both calcium binding and proton countertransport. Multiconformation continuum electrostatic (MCCE) calculations have been carried out in order to assign the measured infrared (IR) protonation signals to specific carboxylic amino acid residues. There are various methods that have been developed and are applied to calculate pK_as of residues in proteins. Among them are hybrid methods like MCCE that combine molecular mechanics and continuum electrostatics.^{12–16} These methods keep the protein dielectric constant low and allow the consideration of multiple positions of side chains. The aim is to incorporate some of the reorganization of the protein to changes in charge into explicit conformational changes rather than averaging them in the dielectric constant. Here, MCCE feasibility for IR band assignment is shown exemplarily for the calcium ligand Asp 800 of the Ca²⁺-ATPase.

METHODS

Multiconformation continuum electrostatics has been used to calculate the equilibrium distribution of residue ioniza-

tion and side chain conformation in dependence of pH. Traditional continuum electrostatics considers protein microstates that differ only in the residue ionization. Other protein and solvent responses to charges are implicitly included in the continuum dielectric constant. In MCCE, the response of side chains is explicitly taken into account by considering different side chain orientations (rotamers) that are constructed systematically by rotating rotatable bonds. Thus, when a residue undergoes a change in charge, the surrounding dipoles are allowed to change orientation, which leads to a more accurate calculated pK_a . Each residue is represented by a set of conformers that differ in their ionization state, side chain rotamer, and/or hydrogen position. Protein microstates are created by choosing one conformer for each residue. These are subjected to Monte Carlo sampling yielding the occupancy of each conformer according to a Boltzmann distribution as a function of pH. All electrostatic interactions are calculated with the program Delphi.¹⁷ For the protein dielectric constant $\epsilon = 4$ is chosen, while 80 is used for the solvent. PARSE parameters provide atomic charges and radii.¹⁸ Torsion and Lennard Jones parameters were previously reported.¹² For a detailed description of the methodology, see Ref. 19. Recent MCCE calculations on proteins can be found, e.g., in Ref. 20–25.

Calculations were performed with the calcium bound E1 state (pdb entry 1SU4.pdb; resolution 2.4 Å) and the calcium free E2 state (pdb entry 1IWO.pdb; resolution 3.1 Å) of the Ca^{2+} -ATPase. The transmembrane domain, where the calcium binding is located, was defined with the structural classification program SCOP.²⁶ These residues (1–124, 240–343, and 751–994) were subjected to more extended rotamer sampling. The inhibitor thapsigargin ($C_{34}O_{12}H_{50}$) was kept in the structure file of E2 and treated as neutral cofactor. 1IWO.pdb does not contain crystal water whereas 1SU4.pdb contains 26 crystal water molecules. For calculations with extra added water molecules in the 1IWO.pdb structure, the IPECE program was used.²⁴

RESULTS AND DISCUSSION

Infrared difference measurements of the Ca^{2+} -ATPase at pH 7 show that carboxyl groups get protonated upon calcium release and have been tentatively assigned to Ca^{2+} ligands that become protonated when calcium leaves the binding site.¹¹ To simulate this measured $E1P \cdot 2Ca^{2+} \rightarrow E2P$ transition, calculations have been carried out with the coordinate files of the unphosphorylated states, $E1 \cdot 2Ca^{2+}$ and E2. This seems to be justified since IR signals are similar in amplitude and frequency positions for the phosphorylated and unphosphorylated Ca^{2+} -ATPase.¹¹ The calculations provide the ionization of each residue of both enzyme states in dependence of pH. Changes in residue side chain orientation due to changes in the charge distribution are thereby included. Water molecules that are available in the $E1 \cdot 2Ca^{2+}$ structure were at first

deleted for more equivalent comparison with the low-resolution E2 structure in which crystal water is lacking. Water-filled cavities were accounted for by continuum water.

According to the IR measurements, a carboxylic amino acid residue that contributes to the IR difference bands should be a Ca^{2+} ligand, should be unprotonated in the calcium-bound E1 state, and becomes protonated in the calcium-free E2 state at pH 7. At higher pH values, it should deprotonate in E2 since the proton countertransport ceases.⁹ There are four calcium ligands in the calcium-binding site that provide carboxyl groups (Figure 2). The calculations could identify Asp 800, calcium ligand to both calcium ions, to exhibit the above-mentioned properties. Figure 3 shows the calculated titration curves of Asp 800 for the calcium-bound $E1 \cdot 2Ca^{2+}$ state and the calcium-free E2 state. A fractional protonation of 1 is equivalent to a fully protonated carboxyl group whereas 0 corresponds to a fully ionized carboxylate. For the calcium-bound E1 state, the calculated pK_a is below 0, corresponding well to other calculations²⁷ and revealing that Asp 800 is unprotonated in the IR ground state ($E1P \cdot 2Ca^{2+}$) at each pH. A protonation signal will occur in the IR difference spectrum when Asp 800 becomes protonated in the transition state (E2P). The calculated pK_a for Asp 800 is 7.9 in E2, which is in good agreement with experimental studies, indicating that the pK_a of the luminal binding sites is approximately 7.7.⁹ According to Figure 3, the infrared ν_{COOH} signal that originates from Asp 800 should be maximum in the pH region below pH 6, decreasing with increasing pH and disappearing approximately at pH 10 and above. An error within

FIGURE 3 Calculated titration curves of Asp 800 in the $E1 \cdot 2Ca^{2+}$ state (squares) and E2 state (triangles) of the Ca^{2+} -ATPase. A fractional protonation of 1 is equivalent to a fully protonated carboxylic acid, a fractional protonation of 0 corresponds to a fully ionized carboxylate.

1 pK_a unit has to be taken into account for the pK_a calculations. It should be noted that the E2 structure 1IWO.pdb provides only a resolution of 3.1 Å that may have errors in atomic position accuracy and therefore may lead to errors in the pK_a calculation. Rotamer sampling in MCCE reduces but does not eliminate the dependence on the initial protein structure.

The better-resolved calcium-bound structure reveals that water molecules are in a small cavity close to Asp 800. Additional calculations have been performed with water molecules in this cavity close to Asp 800 (data not shown). Available crystal water molecules have been used for the E1 structure, whereas for the E2 structure, water molecules have been added by the IPECE program.²⁴ The water molecules have been treated explicitly with water conformers differing in proton positions. For the E1 state, the pK_a remains below 0 and for the E2 state, the pK_a shifts only slightly about 0.5 ΔpK units (1 ΔpK unit = 1.36 kcal/mol = energy needed to change a pK_a by 1 pH unit at 25°C).

To get a deeper understanding in what are the factors shifting the pK_a of a residue in the protein compared to the pK_a in solution, an additional energy analysis can be performed. The reference value pK_{sol} for Asp and Glu is 4.75, corresponding to a simple carboxyl acid. In numerous other electrostatic calculations on proteins, the values used for pK_{sol} of Asp and Glu are 4.0 and 4.4, respectively. Those values are from experimental studies of peptides.²⁸ In MCCE calculations, the protein backbone is explicitly treated, and using a simple carboxylic acid as reference pK_{sol} avoids double counting the impact of the peptide backbone.¹⁹ Studies have found that the backbone lowers the pK_a of these residues even in peptides.²⁹ Asp 800 is deeply buried inside the protein and loses about 4.7 ΔpK units (6.4 kcal/mol) of reaction field (desolvation) energy in both enzyme states, which would shift the pK_{sol} up in the absence of any other interaction with the protein. In the E2 state, the electrostatic interaction of Asp 800 with the protein backbone, its van der Waals interactions, and entropy terms are less than 0.5 ΔpK units in absolute values and thus of minor impact for shifting the pK_a. The interaction between Asp 800 and all other residues, which is approximated by a mean field approach, contributes less than 1 ΔpK unit at pH 7. Hence, in E2, the upshift of the pK_a is mainly induced by the desolvation energy. In the calcium-bound E1 state, the interaction with the backbone shifts the pK_a up about 2.6 ΔK units; van der Waals interactions and entropy terms are also less than 0.5 ΔpK units in absolute values. However, in total, the free ionization energy remains extremely negative at each pH and

strongly favors the ionized form of Asp 800. Main reason for the low pK_a are the 2 calcium ions stabilizing the ionization of Asp 800 by more than 10 ΔpK units.

This study shows the capability of electrostatic calculations to assign IR difference bands to specific carboxyl groups becoming protonated during the Ca²⁺-ATPase activity. Asp 800, ligand to both calcium ions, could be identified to get protonated upon calcium release and thus to contribute to the proton countertransport. An energy analysis facilitates the understanding of the energy terms determining the pK_a of Asp 800 in different enzyme states. A more comprehensive study including pH-dependent IR measurements together with MCCE analysis of all carboxyl ligands is in preparation.

MCCE code is developed in Marilyn Gunners's lab (City College of New York, USA). I thank Marilyn Gunner, Junjun Mao, and Yifan Song for helpful discussions, and Julia Andersson and Andreas Barth (Stockholm University, Sweden) for infrared measurements and collaboration in this project.

REFERENCES

1. Barth, A. *Progr Biophys Mol Biol* 2000, 74, 141–173.
2. Dioumaev, A. K. *Biochemistry (Moscow)* 66(11), 2001, 1269–1276.
3. Stokes, D. L.; Green, N. M. *Annu Rev Biophys Biomol Struct* 2003, 32, 446–468.
4. Toyoshima, C.; Nomura, H. *Nature* 2002, 418, 605–611.
5. Toyoshima, C.; Mizutani, T. *Nature* 2004, 430, 529–535.
6. Toyoshima, C.; Nomura, H.; Tsuda, T. *Nature* 2004, 432, 361–368.
7. Sorensen, T. L. M.; Moller, J. V.; Nissen, P. *Science* 2004, 304, 1672–1675.
8. Toyoshima, C.; Nakasako, M.; Nomura, H.; Oka, T. *Nature* 2000, 405, 647–655.
9. Yu, X.; Hao, L.; Inesi, G. *J Biol Chem* 1994, 269, 16656–16661.
10. Peinelt, C.; Apell, H. J. *Biophys J* 2002, 82, 170–181.
11. Barth, A.; Kreutz, W.; Mäntele, W. *J Biol Chem* 1997, 272, 25507–25510.
12. Alexov, E. G.; Gunner, M. R. *Biophys J* 1997, 72, 2075–2093.
13. Beroza, P.; Case, D. *J Phys Chem* 1996, 100, 20156–20163.
14. Spassov, V. Z.; Luecke, H.; Gerwert, K.; Bashford, D. *J Mol Biol* 2002, 312, 203–219.
15. You, T. J.; Bashford, D. *Biophys J* 1995, 69, 1721–1733.
16. Nielsen, J. E.; Vriend, G. *Proteins Struct Funct Genet* 2001, 46, 403–412.
17. Nicholls, A.; Honig, B. *J Comp Chem* 1991, 12, 435–445.
18. Sitkoff, D.; Sharp, K. A.; Honig, B. *J Phys Chem* 1994, 98, 1978–1988.

19. Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. *Biophys J* 2002, 83 (4), 1731–1748.
20. Haas, A. H.; Lancaster, C. R. D. *Biophys J* 2004, 87 (6), 4298–4315.
21. Hauser, K.; Mao, J.; Gunner, M. R. *Biopolymers* 2004, 74, 51–54.
22. Kim, J.; Mao, J.; Gunner, M. R. *J Mol Biol* 2005, 348, 1283–1298.
23. Mao, J.; Hauser, K.; Gunner, M. R. *Biochemistry* 2003, 42(33), 9829–9840.
24. Song, Y.; Mao, J.; Gunner, M. R. *Biochemistry* 2003, 42, 9875–9888.
25. Zhu, Z. Y.; Gunner, M. R. *Biochemistry* 2005, 44 (1), 82–96.
26. Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. *J Mol Biol* 1995, 247, 536–540.
27. Sugita, Y.; Miyashita, N.; Ikeguchi, M.; Kidera, A.; Toyoshima, C. *J Am Chem Soc* 2005, 127, 6150–6151.
28. Richarz, R.; Wüthrich, K. *Biopolymers* 1978, 17, 2133–2141.
29. Gunner, M. R.; Saleh, M. A.; Cross, E.; Ud Doula, A.; Wise, M. *Biophys J* 2000, 78(3), 1126–1144.