RESEARCH COMMUNICATION

Identification of the Mg^{2+}-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY

Ivo S. RIDDER and Bauke W. DIJKSTRA[1]

Laboratory of Biophysical Chemistry, Department of Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

The large HAD (**h**alo**a**cid **d**ehalogenase) superfamily of hydrolases comprises P-type ATPases, phosphatases, epoxide hydrolases and L-2-haloacid dehalogenases. A comparison of the three-dimensional structure of L-2-haloacid dehalogenase with that of the response regulator protein CheY allowed the assignment of a conserved pair of aspartate residues as the Mg^{2+}-binding site in the P-type ATPase and phosphatase members of the superfamily. From the resulting model of the active site, a conserved serine/threonine residue is suggested to be involved in phosphate binding, and a mechanism comprising a phospho-aspartate intermediate is postulated.

Key words: active site, L-2-haloacid dehalogenase, catalytic mechanism.

INTRODUCTION

The HAD (**h**alo**a**cid **d**ehalogenase) superfamily of hydrolases contains enzymes such as L-2-haloacid dehalogenase, epoxide hydrolase and a variety of phosphatases, including phosphoserine phosphatase, phosphomannomutase, phosphoglycolate phosphatase and sucrose-phosphate synthase [1]. Recently, also, the catalytic subunits of the P-type ATPases were included in this superfamily on the basis of additional sequence comparisons [2,3]. These ATPases are essential for the transport of cations across biological membranes, and whereas some P-type ATPases (P-ATPases) exist as two- or multi-subunit complexes, others

```
                    motif I                   motif II                                     motif III

DHLB_Xa      1 IKAVVFDAYGTLFDV  -84- AQCLAELLKRAILSNGAPD  -15- DAVISVDAKRVFKPHPDSYALVEEVLGVTPAEVLFVSSN----GFDVGGAKNFGFSVARV
L-DEX_Ps     3 IKGIAFDLYGTLFDV  -86- RELKRRGLKLAILSNGSPQ  -15- DHLLSVDPVQVYKPDNRVYELAEQALGLDRSAILFVSSN----AWDATGARYFGFPTCWI
HYES_Hs      2 LRGAVFDLDGVLALP  -76- LMLRKKGFTTAILTNTWLD  -20- DFLIESCQVGMVKPEPQIYKFLLDTLKASPSEVVFLDDI----GANLKPARDLGMVTILV
ATHA_Hs    380 TSVICSDKTGTLTQN -219- LKCRTAGIRVIMVTGDHPI  -63- PEMVFARTSPQQKLVI-VESCQRLGAI-----VAVTGDG----VNDSPALKKADIGVAMG
ATKA_Eh    418 LNSIILDKTGTITQG -135- EQLQQKGVDVFMVTGDNQR  -12- SDHIFAEVLPEEKANY-VEKLQKAGKK-----VGMVGDG----INDAPALRLADVGIAMG
PSPASE_Hs 13 ADAVCFDVDSTVIRE  -67- SRLQERNVQVFLISGGFRS  -32- ETQPTAESGGKGKVI--KLLKEKFHFK----KIIMIGDG----ATDMEACPPADAFIGFG
TPS2_Sc    570 RRLFLFDYDGTLTPI  -18- KLCADPHNQIWIISGRDQK -106- KANIEVRPRFVNKGEIVKRLVWHQH -21- FVLCLGDD----FTDEDMFRQLNTI
CBBZP_Ss     2 IKAVLFDFDGTIADT  -79- TLLNQKGYVLGIVTSNSKD  -15- AFVKAGTTLF-GKNRIINRVLKEHKFGT--DEVIYVGDE----TRDISAAKKSRLTMVSV
PMM_Sc      12 ETLVLFDVDGTLTPA  -35- LAKLRNKCCIGFVSDLSKQ -124- QISFDVFPAGWDKTYC-LQHVEKDGFKE----IHFFGDKTMVGGNDYEIFVDERTIGHSV
                                                  ^GG

HAD_cons           UUUUDxxGTU                 UUUUT                       K                         UUUUGD     D
                       8   12                   114                     147                            172    176

CheY_cons         DLxxxDxxMP              PVIxxTA                     DYxxKP                      LxVDD
                      57                       87                      109                          13
```

Figure 1 Alignment of members of the HAD superfamily

The following sequences are used (SWISS-PROT code in parentheses): DhlB-Xa, L-2-haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 (Q60099); L-DEX-Ps, L-2-haloacid dehalogenase from *Pseudomonas* sp. YL (Q53464); HYES-Hs, epoxide hydrolase from *Homo sapiens* (P34913); ATHA-Hs, gastric H^{+},K^{+}-ATPase from *Homo sapiens* (P20648); ATHA-Eh, Cu^{+}-ATPase from *Enterococcus hirae* (P32113); PSPASE-Hs, L-3-phosphoserine phosphatase from *Homo sapiens* (P78330); TPS2-Sc, trehalose-6-phosphate phosphatase from *Saccharomyces cerevisiae* (P31688); CBBZP-Ss, phosphoglycolate phosphatase from *Synechocystis* sp. (P73525); PMM-Sc, phosphomannomutase from *Saccharomyces cerevisiae* (P07283). The PMM-Sc sequence contains an insertion of two glycine residues upstream of the conserved motif II serine/threonine residue. Conserved HAD residues discussed in the text are in **bold** with the equivalent residue numbers in DhlB indicated; 'U' denotes a bulky hydrophobic residue; 'x' is any residue. The sequence alignment is largely based on [3]; for comparison the structurally equivalent residues in CheY are given under 'CheY-cons', the consensus sequence of the CheY superfamily [33].

Abbreviations used: DhlB, L-2-haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10; L-DEX YL, L-2-haloacid dehalogenase from *Pseudomonas* sp. YL; HAD superfamily, **h**alo**a**cid **d**ehalogenase superfamily; P-ATPase, P-type ATPase.

[1] To whom correspondence should be addressed (e-mail bauke@chem.rug.nl).

consist of only one subunit [4]. Their catalytic subunit contains the binding sites for ATP and Mg^{2+} and a phosphorylation site [5]. Structural information on the HAD hydrolases is beginning to emerge with recent electronmicroscopy studies, which yielded 0.8 nm (8 Å) resolution maps of two P-type ATPases [6,7]. Furthermore, high-resolution X-ray structures have been solved for two other members of the HAD superfamily, namely the L-2-haloacid dehalogenases from *Xanthobacter autotrophicus* GJ10 (DhlB) [8] and *Pseudomonas* sp. YL (L-DEX YL) [9].

In a multiple sequence alignment of members of the HAD superfamily, three motifs have been identified [3]. Motif I contains an absolutely conserved aspartate residue and a highly conserved threonine residue, which are Asp^{8} and Thr^{12} in DhlB. Motif II comprises a conserved Ser/Thr (Ser^{114}) at the end of a β-strand, and motif III includes a fully conserved lysine (Lys^{147}) and a pair of aspartate residues of which the first one is a serine residue only in the dehalogenases (Ser^{172} and Asp^{176}) (Figure 1). On the basis of the three-dimensional structure of L-DEX YL, Aravind et al. [3] have presented a model for the P-ATPase catalytic subunit from which they suggest that the absolutely conserved aspartate residue in motif I is the nucleophile forming an acyl-phosphate intermediate in the proposed reaction mechanism for the P-ATPases. In contrast with L-2-haloacid dehalogenase, the P-ATPases and phosphatases require an Mg^{2+} ion for maximum activity [2,10–14]. It has been demonstrated that the ion is important in the P-ATPase pump cycle [15,16]. However, not much attention has been paid to the role of Mg^{2+} in recent reviews or proposed mechanisms for the ATPase reaction [3,4,11].

Here we present a role for the conserved residues in motifs II and III, the identification of the Mg^{2+}-binding site, and a proposal for the reaction mechanism of the P-type ATPases and phosphatases which includes the Mg^{2+} ion, based on our structural work on DhlB.

RESULTS AND DISCUSSION

DhlB and the response regulator protein CheY are structurally similar

A DALI-search [17] identified several Mg^{2+}-binding proteins that are structurally related to DhlB. Among them are phosphofructokinase, which catalyses the phosphorylation of fructose 6-phosphate by ATP [18], the H-*ras* oncogene protein p21, which hydrolyses GTP [19], members of the integrin family of plasma-membrane proteins that bind collagen [20,21], and CheY, a Mg^{2+}-dependent response regulator protein that acts as a phosphorylation-activated switch [22,23]. The former two structures show an active-site organization that is entirely different from that of DhlB, and no phosphorylation of an aspartate residue is involved in their reaction mechanisms. The integrins contain a conserved aspartate residue in a position topologically similar to that of Asp^{8} in DhlB, which is part of the DXSXS motif that binds the Mg^{2+} ion. However, these proteins bind collagen and have no hydrolytic activity. Moreover, the environment of the conserved aspartate residue is completely different from the active-site aspartate residue in DhlB.

In contrast, a structural alignment of CheY and DhlB revealed unexpected similarity in their active-site architecture, although these proteins do not share significant overall sequence identity. In CheY, the phosphorylation takes place at Asp^{57} [24]. Alignment of Asp^{57} of CheY with Asp^{8} of DhlB, and superposition of the central β-sheets of which they are part, shows that the cores of the two structures superimpose with a root-mean-square difference of 0.22 nm (2.2 Å) for 68 C^{α}-atoms. Thr^{87} of CheY is then the structural equivalent of the motif II Ser/Thr of the HAD superfamily (Ser^{114} of DhlB), and Lys^{109} of CheY is equivalent to the motif III Lys^{147} of DhlB (Figures 2A and 2B). In DhlB, this lysine residue makes a salt bridge to the active-site aspartate residue, and we have suggested that it activates the water molecule that hydrolyses the acyl-intermediate [8]. In CheY, a lysine residue is present at a similar position, but its role has not unambiguously been established. Stock et al. [23] found that the catalytic aspartate–lysine salt bridge that exists in the

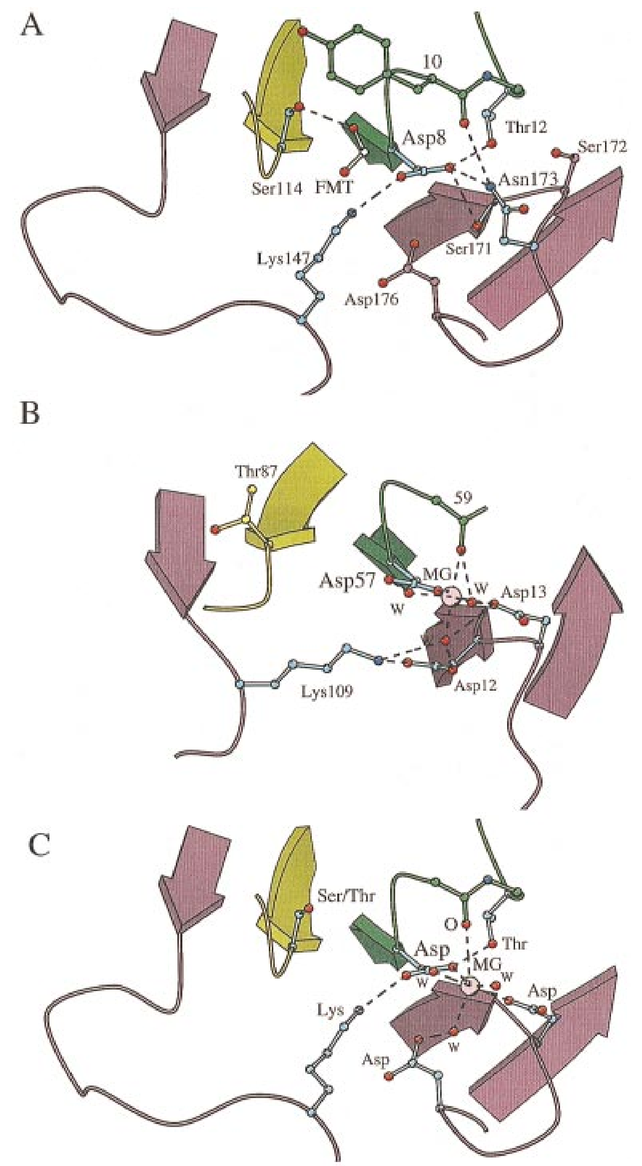
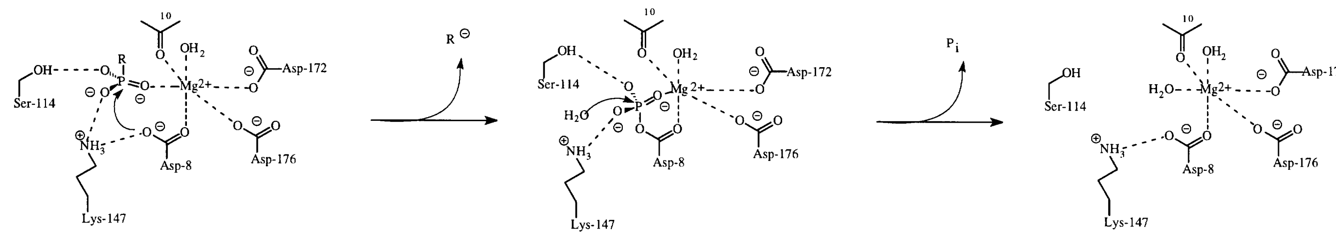


Figure 2 Active-site organization of DhlB, CheY and the P-ATPase/phosphatase model

(**A**) Active site and central β-sheet of DhlB; colours according to sequence conservation in the HAD superfamily: motif I, dark green; motif II, yellow-green; motif III, purple; catalytic residues, light blue. Active-site residues are shown in a ball-and-stick format; FMT, formate ion. (**B**) Active site of CheY with Mg^{2+} ion (PDB code 1CHN [22]) in a view similar to that in (**A**); the ion (pink) is co-ordinated by Asp^{57}, Asn^{59} (main chain), a pair of aspartate residues (Asp^{12}/Asp^{13}), one of which is co-ordinating via a water molecule, and two more water molecules (W). Upon phosphate binding a water molecule can be replaced by a phosphate oxygen atom. (**C**) Model for active site of P-ATPase and phosphatase members of the HAD superfamily; the Mg^{2+} ion is co-ordinated by the motif I aspartate residue, a motif III aspartate residue, an oxygen atom of a main-chain threonine/aspartate residue and three water molecules (W), one of which is positioned by the second motif III aspartate residue. Upon phosphate binding a water molecule can be replaced by a phosphate oxygen atom. The Figure was prepared with Molscript [34].



Scheme 1 Proposed catalytic mechanism of ATPase and phosphatase activity

After substrate binding (R = ADP, sugar, or amino acid), mediated by Ser^{114} and the Mg^{2+} ion, Asp^{8} performs a nucleophilic attack on the phosphorus atom. The R group is cleaved off and an acyl-phosphate intermediate is formed. The Mg^{2+} ion is co-ordinated by oxygen atoms from Asp^{8}, Asp^{172}, Asp^{176} (possibly via a water molecule), phosphate, the carbonyl main chain oxygen atom of residue 10, and a water molecule. In the next step, the acyl-phosphate intermediate is hydrolysed by nucleophilic attack of a water molecule on the phosphorus atom to yield phosphate and free enzyme.

Mg^{2+}-free structure is disrupted when an Mg^{2+} ion is bound. They proposed that the lysine residue in CheY might be involved in an interaction with an oxygen of the acyl-phosphate. The lysine residue in the P-ATPases and phosphatase members of the HAD superfamily could likewise be involved in phosphate or ATP binding or stabilization of the phosphorylated state.

CheY has an absolute requirement for Mg^{2+} for phosphorylation. This Mg^{2+} is co-ordinated by oxygen atoms from the nucleophile Asp^{57} (DhlB-equivalent Asp^{8}), Asp^{13} (Ser^{172}), Asn^{59} main chain carbonyl (Tyr^{10}), and three water molecules, one of which is held in position by Asp^{12} (Asp^{176}) [22,23] (Figure 2B). The conservation of Asp^{172} and Asp^{176} thus clearly suggests that the function of these motif III aspartate residues is to bind the Mg^{2+}. The conserved aspartate residues are present in all HAD-superfamily enzymes, with the exception of L-2-haloacid dehalogenase and epoxide hydrolase. The activity of the latter enzymes is independent of Mg^{2+}, and the serine residue at position 172 is not essential for catalytic activity [25].

The structure of the DhlB contained a formate ion in the active site that enabled us to construct a model for substrate binding [8]. The model concurred with biochemical evidence that Asp^{8} is the nucleophile in the first step of the reaction where an enzyme-ester intermediate is formed which is subsequently hydrolysed [26]. Thr^{12}, Ser^{171} and Asn^{173} form the oxyanion hole that stabilizes the negative charge of Asp^{8}. It has been shown that the corresponding aspartate residue is the phosphorylation site in the P-ATPases and phosphatases [27,28]. Furthermore, in DhlB the conserved motif II serine residue binds the negatively charged carboxylate group of the 2-haloalkanoate substrate. In our model of the active site of the P-ATPases and phosphatases (Figure 2C), this motif II serine/threonine residue is at a distance of about 0.35–0.45 nm (3.5–4.5 Å) from one of the phosphate oxygen atoms of the acyl-phosphate intermediate, and therefore this residue could have a similar function in phosphate binding.

The structural similarity between DhlB and CheY implies that the sequences of these enzyme families are related by a circular permutation as the two N-terminal β-strands of CheY (Figure 2B, left) correspond to the C-terminal strands of DhlB (Figure 2A, left). Although it is a rare phenomenon, *in vivo* circular permutations have been identified in the β-1,3-1,4-glucanase family [29] and among α-amylases [30].

Model for the catalytic mechanism of phosphatase and ATPase activity in enzymes of the HAD superfamily

On the basis of the structural similarities described above, a possible reaction mechanism for the P-type ATPases and phosphatases of the HAD superfamily can be proposed. The reaction starts with the binding of the phosphorylated substrate by, among others, the motif II serine/threonine residue. This is most likely followed by a nucleophilic attack of one of the carboxylate oxygen atoms of the motif I aspartate residue to form an acyl-phosphate intermediate (Scheme 1). The other carboxylate oxygen could be held in position by the motif I threonine residue and the Mg^{2+}. The positive charge of the bound Mg^{2+} may bind the phosphate moiety of the substrate and supply charge shielding of the negatively charged phosphate group. In addition, it might also be involved in the activation of the attacking nucleophile, the stabilization of the leaving group, or the enhancement of the electrophilicity of the phosphorus atom by polarizing the P–O bond. Also, positively charged active-site residues like the motif III lysine residue may provide one or more of these functions. In the phosphorylated state, one phosphate oxygen atom can also be involved in co-ordination of the Mg^{2+} in addition to the carboxylate oxygen atoms from the motif III pair of aspartate residues.

In L-2-haloacid dehalogenase, subsequent hydrolysis is achieved by attack of a water molecule on the aspartate γ-carbon [31], but it has been demonstrated that, in P-ATPases, the acyl-phosphate ester is hydrolysed by attack on the phosphorus atom [32]. In contrast with DhlB, such an hydrolytic mechanism does not include the formation of an oxyanion intermediate on the motif I aspartate residue. In agreement with this, the residues that stabilize the oxyanion in DhlB, namely Ser^{171} and Asn^{173}, are not conserved in the P-ATPases and phosphatases. Instead, the negative charge will develop on the phosphorus atom, where it is counterbalanced by the Mg^{2+}.

Sequence-alignment methods have been shown to provide a useful tool in relating families of proteins with different functions, and such methods help in constructing a crude model of enzymes for which no structural information is available. However, if structural alignments can also be employed, extra information is to be gained even from enzymes that do not have an obvious evolutionary relationship.

This work was supported by the Netherlands Foundation for Chemical Research (SON), with financial aid from the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

1 Koonin, E. V. and Tatusov, R. L. (1994) J. Mol. Biol. **244**, 125–132
2 Collet, J.-F., Gerin, I., Rider, M. H., Veiga-da-Cunha, M. and Van Schaftingen, E. (1997) FEBS Lett. **408**, 281–284
3 Aravind, L., Galperin, M. Y. and Koonin, E. V. (1998) Trends Biochem. Sci. **23**, 127–129

4 Lutsenko, S. and Kaplan, J. H. (1995) Biochemistry **34**, 15607–15613
5 Munson, K. B., Gutierrez, C., Balaji, V. N., Ramnarayan, K. and Sachs, G. (1991) J. Biol. Chem. **266**, 18976–18988
6 Auer, M., Scarborough, G. A. and Kühlbrandt, W. (1998) Nature (London) **392**, 840–843
7 Zhang, P. J., Toyoshima, C., Yonekura, K., Green, N. M. and Stokes, D. L. (1998) Nature (London) **392**, 835–839
8 Ridder, I. S., Rozeboom, H. J., Kalk, K. H., Janssen, D. B. and Dijkstra, B. W. (1997) J. Biol. Chem. **272**, 33015–33022
9 Hisano, T., Hata, Y., Fujii, T., Liu, J. Q., Kurihara, T., Esaki, N. and Soda, K. (1996) J. Biol. Chem. **271**, 20322–20330
10 Skou, J. C. (1957) Biochim. Biophys. Acta **23**, 394–401
11 Moeller, J. V., Juul, B. and Le Maire, M. (1996) Biochim. Biophys. Acta **1286**, 1–51
12 Vandercammen, A., François, J. and Hers, H.-G. (1989) Eur. J. Biochem. **182**, 613–620
13 Norman, E. G. and Colman, B. (1991) Plant Physiol. **95**, 693–698
14 Guha, S. K. and Rose, Z. B. (1985) Arch. Biochem. Biophys. **243**, 168–173
15 Shigekawa, M., Wakabayashi, S. and Nakamura, H. (1983) J. Biol. Chem. **258**, 14157–14161
16 Wakabayashi, S. and Shigekawa, M. (1987) J. Biol. Chem. **262**, 11524–11531
17 Holm, L. and Sander, C. (1993) J. Mol. Biol. **233**, 123–138
18 Shirakihara, Y. and Evans, P. R. (1988) J. Mol. Biol. **204**, 973–994
19 Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W. and Wittinghofer, A. (1990) EMBO J. **9**, 2351–2359
20 Lee, J.-O., Rieu, P., Arnaout, M. A. and Liddington, R. (1995) Cell **80**, 631–638
21 Emsley, J., King, S. L., Bergelson, J. M. and Liddington, R. C. (1997) J. Biol. Chem. **272**, 28512–28517
22 Bellsolell, L., Prieto, J., Serrano, L. and Coll, M. (1994) J. Mol. Biol. **238**, 489–495
23 Stock, A. M., Martinez-Hackert, E., Rasmussen, B. F., West, A. H., Stock, J. B., Ringe, D. and Petsko, G. A. (1993) Biochemistry **32**, 13375–13380
24 Stock, A. M. and Mowbray, S. L. (1995) Curr. Opin. Struct. Biol. **5**, 744–751
25 Kurihara, T., Liu, J. Q., Nardi-Dei, V., Koshikawa, H., Esaki, N. and Soda, K. (1995) J. Biochem. (Tokyo) **117**, 1317–1322
26 Liu, J. Q., Kurihara, T., Miyagi, M., Tsunasawa, S., Nishihara, M., Esaki, N. and Soda, K. (1997) J. Biol. Chem. **272**, 3363–3368
27 Asano, S., Tega, Y., Konishi, K., Fujioka, M. and Takeguchi, N. (1996) J. Biol. Chem. **271**, 2740–2745
28 Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G. and Van Schaftingen, E. (1998) J. Biol. Chem. **273**, 14107–14112
29 Schimming, S., Schwarz, W. H. and Staudenbauer, W. L. (1992) Eur. J. Biochem. **204**, 13–19
30 MacGregor, E. A., Jespersen, H. M. and Svensson, B. (1996) FEBS Lett. **378**, 263–266
31 Liu, J. Q., Kurihara, T., Miyagi, M., Esaki, N. and Soda, K. (1995) J. Biol. Chem. **270**, 18309–18312
32 Dahms, A. S., Kanazawa, T. and Boyer, P. D. (1973) J. Biol. Chem. **248**, 6592–6595
33 Volz, K. (1993) Biochemistry **32**, 11741–11753
34 Kraulis, P. (1991) J. Appl. Crystallogr. **24**, 946–950

Received 22 January 1999/15 February 1999; accepted 18 February 1999