CMLS, Cell. Mol. Life Sci. 61 (2004) 387–392
1420-682X/04/040387-06
DOI 10.1007/s00018-003-3291-5

CMLS Cellular and Molecular Life Sciences

Visions & Reflections

Protein structure to function: insights from computation

D. Ringe [a], Y. Wei [b], K. R. Boino [b] and M. J. Ondrechen [b,*]

[a] Rosenstiel Basic Medical Sciences Research Center and Departments of Chemistry and Biochemistry, Brandeis University, Waltham, Massachusetts 02454 (USA)
[b] Department of Chemistry and Chemical Biology and Institute for Complex Scientific Software, Northeastern University, Boston, Massachusetts 02115 (USA), Fax: +1 617 373 8795, e-mail: mjo@neu.edu

Received 22 July 2003; received after revision 16 September 2003; accepted 20 October 2003

Abstract. Computation plays an important role in functional genomics. THEMATICS is a computational method that predicts chemical and electrostatic properties of residues in enzymes and utilizes information contained in those predictions to identify active sites. The only input required is the three-dimensional structure of the query protein. The identification of residues involved in catalysis and in recognition is discussed. The two serine proteases Kex2 from *Saccharomyces cerevisiae* and subtilisin from *Bacillus subtilis* are used as examples to illustrate how the method finds the catalytic residues for both enzymes. In addition, Kex2 is specific for dibasic sites and THEMATICS finds the recognition residues for both the S1 and S2 sites of Kex2. In contrast, no such recognition sites are found for the non-specific enzyme subtilisin. The ability to identify sites that govern recognition opens the door to better understanding of specificity and to the design of highly specific inhibitors.

Key words. Functional genomics; THEMATICS; protein function; enzyme catalysis; enzyme specificity; recognition.

The genomics revolution has unlocked the potential to characterize all proteins in a cell and to establish the links between them. The genomes of about 1000 organisms have been sequenced, providing the input for such analyses. These genomes represent all types of living species from the simplest to the most complex. However, the actual realization of this potential has been more difficult than originally imagined. To utilize the information from these sequences, the gene products predicted from them must be characterized in terms of their functions. Because of the large number of proteins for which function must be established, computational methods are being developed that can speed up the process.

Sequence comparisons have been valuable in determining relationships between some proteins, especially if the sequence identity is high enough so that the relationship between two proteins can be established with some level of confidence. That relationship can be at the structural level, or the structural and functional level. Ultimately, the functional relationship is the important one. Structure has played an important role in the assignment of function. If the sequences of two proteins from different organisms are similar, their structures and functions are also expected to be similar. To a first approximation, this is true. However, enough exceptions exist such that this simple rule does not hold every time. Two proteins may be structurally related without a functional relationship. In addition, sequences can be so diverse that the structural and functional relationship between the proteins encoded is no longer evident.

Proteins perform many functions within a cell, ranging from biochemical functions such as those performed by enzymes to extracellular functions such as signaling between cells. All of them depend on the binding of a ligand to a protein. Ligands may be as small as a proton or as large as another protein. In all cases, the nature of the in-

* Corresponding author.

teraction between protein and ligand will determine the functional fate of the system. Consequently, determination of the site at which a ligand interacts with a protein, and the nature of the ligand, will help to determine the function of that protein. If we limit ourselves to enzymes, proteins that catalyze a chemical transformation, that site is designed to recognize a specific ligand, usually to the exclusion of others, and to facilitate the chemical transformation. Residues at that site have chemical or electrostatic properties that enable such protein-ligand interactions. A method to distinguish these residues from others of the same type, but that do not have these special properties, is described here.

Theoretical Microscopic Titration Curves, THEMATICS [1–4], is a computational method that predicts chemical and electrostatic properties of residues in enzymes and utilizes information contained in those predictions to identify active sites. THEMATICS is based on Poisson-Boltzmann [5–15] methods for determining the electrostatic potential function of a protein structure. A Monte Carlo procedure calculates the net charge as a function of pH, C(pH), for each ionizable residue in the protein, ensemble averaged over a large number of protein molecules. The only input required is the three-dimensional structure of the subject protein. The sequence and the structure of this protein need not have any similarity to any previously characterized protein.

The Henderson-Hasselbalch equation predicts that the charge-versus-pH curve C(pH), a titration curve, will be a sigmoidal function with a sharp fall-off in charge when the pH is near the pK_a. We find that typically most of the ionizable residues in a protein follow predicted Henderson-Hasselbalch behavior, but a small fraction of the residues are found to have perturbed behavior. In particular, these perturbed residues are predicted to have elongated titration curves, such that partial protonation persists over a wide pH range. Residues with such unusual curves we label as THEMATICS positives. A cluster of two or more of such positive residues has been found to be a reliable predictor of active-site location [1–4]. The method does give false positives, residues with perturbed titration behavior that are not in or near any known active site, but such residues tend to be isolated in space. Hence the clustering criterion is highly successful. A residue is deemed to be a member of a cluster if its reactive atom is within 6 Å of the reactive atom of any other residue in the cluster. To date, THEMATICS has been applied to about 100 enzymes and it finds the active site for about 90% of known enzymes. THEMATICS has been established to successfully locate catalytic residues in the active site [1–4]. The purpose of the present paper is to show that it finds recognition sites also, and can distinguish between specific and non-specific enzymes.

Our method complements well other methods for protein function prediction. For some classes of proteins, information about function can be inferred from the evolutionary history derived from sequence relationships [16–18], or by other homology and non-homology sequence methods [19–22]. Combined analysis of sequence and structure data gives more revealing clues about function [23]. Methods to locate active sites rely either on analogies to related proteins of known function [24–31], on searches for clefts in the structure [32], or on computational searches for binding sites by docking [33] of selected sets of small molecules [34–36]. Energetics [37] and flexibility [38] can also predict functionally important sites. Our method is different from existing approaches because it takes advantage of the unique chemical properties of active sites in proteins to identify and characterize them.

In an enzyme, the protein environment influences the chemical and electrostatic properties of active site residues in order to facilitate catalysis and recognition. Any active-site residue that acts as a catalytic Brønsted acid(/base) must then act as a Brønsted base(/acid) in order to restore itself to its initial state for the next turnover cycle. A perturbed titration curve increases the range of conditions over which a residue may act as both an acid and a base, as is required by the definition of a catalyst. Similarly, such unusual titration behavior expands the pH range over which a residue may exist in both charged and neutral forms, thus assisting reversible recognition. Finding these perturbed titration properties not only identifies the active site but may also give clues about the type of chemistry that is catalyzed by that active site. Our working hypothesis is that this type of perturbed titration behavior facilitates both catalysis and reversible recognition [2].

There have been only a few experimental titrations of individual residues in proteins because the experiments (generally involving nuclear magnetic resonance on a mutant protein with site-directed mass labeling) are very difficult to perform and the mutations can cause changes in functional properties. Consequently, there is only a little bit of experimental evidence at the present time for non-sigmoidal titration behavior in proteins [39, 40; W. W. Bachovchin, personal communication].

In the present paper, we compare the specific serine protease Kex2 to the structurally related non-specific serine protease subtilisin Carlsberg. These two proteins have identical catalytic residues but one has specificity determinants that the other protein lacks. We show how our method identifies the catalytic residues for both specific and non-specific proteases and also identifies the recognition residues for a specific protease. The ability to identify sites that govern recognition opens the door to better understanding of specificity and to the design of highly specific inhibitors. We discuss the implications for catalysis, recognition, and drug design.

Example – serine proteases

The active site of an enzyme in the serine protease family possesses the 'catalytic triad,' an acid, a histidine, and a serine, where the acid is an aspartate or glutamate side chain, the histidine acts as a catalytic base, and the serine serves as a nucleophile [41, 42]. This catalytic triad performs hydrolysis on peptide bonds. In the case of a specific protease, the protein recognizes the side chain of one or more residues adjacent to the peptide bond to be hydrolyzed. The residue that provides the carbonyl carbon atom of the reactive peptide bond is labeled as P_1. The pocket in the protease that recognizes the side chain of the P_1 residue is called the S_1 pocket. Similarly, the next residue to the N-terminal side of the substrate protein is labeled as P_2 and it is recognized by the S_2 site of the protease, and so on.

Kex2 – a specific protease

Kex2 (kexin, E.C. 3.4.21.61) is a calcium dependent transmembrane protease found in the yeast *Saccharomyces cerevisiae*. The name derives from the killer phenotype of Kex2 mutant cells. In *S. cerevisiae*, Kex2 is required for the production and secretion of mature alpha-mating factor and killer toxin, among other activities. Proteolysis occurs at paired basic residues.

Kex2 is the prototype of a large family of eukaryotin proprotein-processing proteases that includes furin, the proprotein convertases, and PACE4 in mammals. This family of proteases is responsible for the processing of neuropeptides and peptide hormones, proinsulin, coagulation factors, many growth hormones and their receptors, Alzheimer-related secretases, and cancer-associated extracellular matrix proteinases. In addition, furin is known to function in embryogenesis and homeostasis, and is required for the activation of many bacterial toxin precursors and virus envelope glycoproteins.

Kex2 specifically cleaves peptide bonds where P_1 and P_2 are the basic residues arginine and lysine. (Kex2 cleaves preferentially after RR and RK.) Kex2 belongs to a family of proteases structurally related to the subtilisins. However, subtilisin is a very non-specific protease: the structural similarity does not account for the unusual specificity observed for Kex2. Inhibited structures of the catalytic and P domains of Kex2 [43] and furin [44] have been determined, and show the source of the calcium dependence and the nature of the high specificity.

Subtilisin – a non-specific serine protease

Subtilisin is mechanistically and structurally related to Kex2. However, subtilisin is non-specific and cleaves peptide bonds in proteins with little regard for the nature of the surrounding residues, except for a small preference for hydrophobic residues. Subtilisin is used commercially in laundry detergents to degrade proteinaceous material into small, soluble fragments.

THEMATICS results

THEMATICS calculations were performed on the 2.4-Å resolution structure of Kex2 [43] from yeast, Protein Data Bank (PDB) [45] code 1OT5 and on the 1.2-Å resolution structure [46] of subtilisin Carlsberg from *Bacillus subtilis*, 1CSE. Table 1 shows the results of these calculations for the two featured proteins plus some additional enzymes chosen to highlight different types of recognition. THEMATICS-positive residues for these examples are known to be involved in catalysis and recognition.

THEMATICS-positive residues for Kex2 were found to be [**D175,** *D176, D210, D211*, E220, **H213,** H381, Y212], [*D277, D320, E350*], [D184], and [Y308], where residues that are clustered together in coordinate space are shown together in brackets, known catalytic residues are shown in boldface, and known recognition residues are shown in italics. The first and largest cluster includes the ionizable catalytic residues D175 and H213. This same cluster also includes the S_2 recognition residues, D176, D210, and D211. The second cluster contains the S_1 recognition residues D277, D320, and E350. D184 and Y308 are each isolated in space and we are not aware of any functional significance for them.

THEMATICS-positive residues for subtilisin were found to be [**D32, H64**], the acid and histidine residues of the known catalytic triad, and D41 (presumed to be an isolated false positive). No residues, analogous to those in the first two clusters for Kex2, indicative of specificity were found.

Results for additional enzymes

Table 1 also summarizes THEMATICS results for some additional enzymes. The last two enzymes in the table are active as dimers and the calculations for these were performed on the dimer structure. In all of these enzymes, the method finds at least one residue involved in catalysis and, except for the non-specific subtilisin, at least one residue involved in recognition. Overall, these enzymes represent a variety of different kinds of chemistry and topology and different types of recognition. Kex2 utilizes the ionizable side chains of the acidic residues aspartate and glutamate to recognize arginine in the S_1 pocket and either arginine or lysine in the S_2 pocket; THEMATICS identifies residues in both of these pockets. For arginine

Table 1. THEMATICS results.

PDB ID/name	Chemistry	Topology
1OT5 Kex2 catalytic residues: **D175, H213** basic side chain recognition: *D277, D320, E350, D176, D210, D211* all positives: [**D175**, *D176, D210, D211,* E220, **H213**, H381, Y212], [*D277, D320, E350*], [D184], [Y308]	serine protease	subtilisin + jelly roll
1CSE subtilisin Carlsberg catalytic residues: **D32, H64** all positives: [**D32, H64**], [D41]	serine protease	subtilisin
1BG0 arginine kinase catalytic residues: **E225** phosphate recognition: *R126, R229, R280, R309* all positives: [*R126, R229,* D226, E224, **E225**], [R124, *R280, R309*, R330], [C127], [E335], [H185], [Y134], [Y145]	phosphate transfer	creatine kinase
1CHM creatine amidinohydrolase catalytic residues: **H232** guanidinium recognition: *E262, E358* all positives: [R64′, R335, D83′,*E262, E358*, **H232**, H324, H376], [D217, H331′, Y48], [E124], [Y54], [Y257]	C-N bond hydrolysis	unique
1QQ5 L-2-haloacid dehalogenase catalytic residues: **D8, D176, K147** recognition residues for halide: *R39* all positives: [*R39*, **D8, D176**, **K147**, Y153], [R217′, E44, K41, Y45, Y68], [Y223]	dehalogenation	unique

THEMATICS results are summarized for some enzymes with known recognition residues, plus subtilisin. We include THEMATICS-positive residues that are known catalytic residues (bold) or known recognition residues (italicized). Prime indicates a residue from another subunit in a dimer. Then we list all of the THEMATICS-positive residues for each enzyme, with residues in physical proximity in the protein shown together in brackets.

kinase [47], THEMATICS identifies a set of arginine residues that recognize a phosphate group. Creatine amidinohydrolase [48] has two glutamate side chains that recognize a guanidinium group of the substrate, and these glutamate residues are found to be positive by THEMATICS. THEMATICS also identifies as positive the residues in the site where the halogen is recognized in L-2-haloacid dehalogenase [49, 50].

Discussion

Since catalysis and recognition are governed by geometry, chemistry, and electrostatics, computational methods that predict chemical and electrostatic properties from the three-dimensional structure should, at least in principle, be able to predict protein function at the atomic and molecular level. We have shown how THEMATICS can predict residues involved in both catalysis and recognition. One can therefore distinguish between a specific and a non-specific protease, for example, from the structure alone with just a simple calculation.

The ability to predict precisely where the specificity occurs in an enzyme, in addition to providing insight into enzyme function, could prove to be very valuable in the design of highly specific inhibitors. Genomics and proteomics have produced manifold new potential drug targets and simple computational tools can point to the source of specificity in these targets. We are in the process of analyzing hundreds of protein structures to establish relationships between computed chemical and electrostatic properties and specific types of recognition**,** and also specific types of chemistry.

Since this computational tool is based on a chemical property of the residues found, simple computational methods can now rapidly identify active-site residues and assign at least some functional information to them for the new protein structures emerging from structural genomics efforts. Computed properties of such residues, and their spatial relationship within the context of an active site, may be able to predict even more detailed functional information in the near future.

Finding our way on this structure-to-function path takes us enticingly closer to the realization of the hope for revolutionary post-genomic innovations in medicine. The discovery of specific inhibitors for gene products associated with diseases involves the identification of active sites and the design of molecules that will recognize and interact with them specifically. THEMATICS is a powerful tool that provides information about both reactivity and specificity in a protein, especially about recognition sites that distinguish one enzyme from another, even when they come from the same family.

Acknowledgements. This work was supported by the National Science Foundation under grant MCB-0135303 and by the Institute for Complex Scientific Software at Northeastern University.

1 Ondrechen M. J., Clifton J. G. and Ringe D. (2001) THEMATICS: a simple computational predictor of enzyme function from structure. Proc. Natl. Acad. Sci. USA **98:** 12473–12478
2 Shehadi I. A., Yang H. and Ondrechen M. J. (2002) Future directions in protein function prediction. Mol. Biol. Rep. **29:** 329–335
3 Ondrechen M. J. (2002) THEMATICS as a tool for functional genomics. Genome Inform. **13:** 563–564
4 Ondrechen M. J., Murga L. F., Clifton J. G. and Ringe D. (2003) Prediction of protein function with THEMATICS. Curr. Comput. Mol. Biol. 21–22
5 Warwicker J. and Watson H. C. (1982) Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. J. Mol. Biol. **157:** 671–679
6 Bashford D. and Karplus M. (1991) Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation. J. Phys. Chem. **95:** 9556–9561
7 Gilson M. K. (1993) Multiple-site titration and molecular modeling: two rapid methods for computing energies and forces for ionizable groups in proteins. Proteins **15:** 266–282
8 Yang A. S., Gunner M. R., Sampogna R., Sharp K. and Honig B. (1993) On the calculation of pKas in proteins. Proteins **15:** 252–265
9 Madura J. D., Briggs J. M., Wade R. C., Davis M. E., Luty B. A. and Ilin A. et al. (1995) Electrostatics and diffusion of molecules in solution – simulations with the University of Houston Brownian Dynamics program. Comp. Phys. Commun. **91:** 57–95
10 Karshikoff A. (1995) A simple algorithm for the calculation of multiple site titration curves. Protein Eng. **8:** 243–248
11 Antosiewicz J., Briggs J. M., Elcock A. H., Gilson M. K. and McCammon J. A. (1996) Computing the ionization states of proteins with a detailed charge model. J. Comp. Chem. **17:** 1633–1644
12 Antosiewicz J., McCammon J. A. and Gilson M. K. (1996) The determinants of pKa's in proteins. Biochemistry **35:** 7819–7833
13 Alexov E. G. and Gunner M. R. (1997) Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. Biophys. J. **72:** 2075–2093
14 Nielsen J. E., Andersen K. V., Honig B., Hooft R. W. W., Klebe G., Vriend G. et al. (1999) Improving macromolecular electrostatics calculations. Protein Eng. **12:** 657–662
15 Mehler E. L. and. Guarnieri F. (1999) A self-consistent, microenvironment modulated screened Coulomb potential approximation to calculate pH-dependent electrostatic effects in proteins. Biophys. J. **77:** 3–22
16 Lichtarge O., Bourne H. R. and Cohen F. E. (1996) An evolutionary trace method defines binding surfaces common to protein families. J. Mol. Biol. **257:** 342–358
17 Sjolander K. (1998) Phylogenetic inference in protein superfamilies: analysis of SH2 domains. Proc. Int. Conf. Intell. Syst. Mol. Biol. **6:** 165–174
18 Yao H., Kristensen D. M., Mihalek I., Sowa M. E., Shaw C., Kimmel M. et al. (2003) An accurate, sensitive, and scalable method to identify functional sites in proteins. J. Mol. Biol. **326:** 255–261
19 Marcotte E. M., Pellegrini M., Ng H. L., Rice D. W., Yeates T. O. and Eisenberg D. (1999) Detecting protein function and protein-protein interactions from genome sequences. Science **285:** 751–753
20 Marcotte E. M., Pellegrini M., Thompson M. J., Yeates T. O. and Eisenberg D. (1999) A combined algorithm for genome-wide prediction of protein function. Nature **402:** 83–86
21 Marcotte E. M. (2000) Computational genetics: finding protein function by nonhomology methods. Curr. Opin. Struct. Biol. **10:** 359–365
22 Ramani A. K. and Marcotte E. M. (2003) Exploiting the coevolution of interacting proteins to discover interaction specificity. J. Mol. Biol. **327:** 273–284
23 Carter C. W. Jr, LeFebvre B. C., Cammer S. A., Tropsha A. and Edgell M. H. (2001) Four-body potentials reveal protein-specific correlations to stability changes caused by hydrophobic core mutations. J. Mol. Biol. **311:** 625–638
24 Babbitt P. C. and Klein T. E. (1998) Superfamily analysis: understanding protein function from structure and sequence. In: Encyclopedia of Computational Chemistry, 2859–2870, Schleyer P. v. R., (ed.), Wiley, Chichester
25 Fetrow J. S., Siew N., Di Gennaro J. A., Martinez-Yamout M., Dyson H. J. and Skolnick J. (2001) Genomic-scale comparison of sequence- and structure-based methods of function prediction: does structure provide additional insight? Protein Sci. **10:** 1005–1014
26 Fetrow J. S., Siew N. and Skolnick J. (1999) Structure-based functional motif identifies a potential disulfide oxidoreductase active site in the serine/threonine protein phosphatase-1 subfamily. FASEB J. **13:** 1866–1874
27 Fetrow J. S. and Skolnick J. (1998) Method for prediction of protein function from sequence using the sequence-to-structure-to-function paradigm with application to glutaredoxins/thioredoxins and T1 ribonucleases. J. Mol. Biol. **281:** 949–968
28 Hegyi H. and Gerstein M. (1999) The relationship between protein structure and function: a comprehensive survey with application to the yeast genome. J. Mol. Biol. **288:** 147–164
29 Skolnick J. and Fetrow J. S. (2000) From genes to protein structure and function: novel applications of computational approaches in the genomic era. Trends Biotechnol. **18:** 34–39
30 Teichmann S. A., Murzin A. G. and Chothia C. (2001) Determination of protein function, evolution and interactions by structural genomics. Curr. Opin. Struct. Biol. **11:** 354–363
31 Wallace A. C., Borkakoti N. and Thornton J. M. (1997) TESS: a geometric hashing algorithm for deriving 3D coordinate templates for searching structural databases: application to enzyme active sites. Protein Sci. **6:** 2308–2323
32 Laskowski R. A., Luscombe N. M., Swindells M. B. and Thornton J. M. (1996) Protein clefts in molecular recognition and function. Protein Sci. **5:** 2438–2452
33 Oshiro C. M., Kuntz I. D. and Knegtel R. M. A. (1998) Molecular docking and structure-based design. In: Encyclopedia of Computational Chemistry, 1606–1613, Schleyer P. V. R., (ed.), Wiley, Chichester
34 Chen Y. Z. (2001) Computer search of putative protein targets of a small molecule. Biophys. J. **80:** 497A
35 Chen Y. Z. and Zhi D. G. (2001) Ligand-protein inverse docking and its potential use in the computer search of protein targets of a small molecule. Proteins **43:** 217–226
36 Dennis S., Kortvelyesi T. and Vajda S. (2002) Computational mapping identifies the binding sites of organic solvents on proteins. Proc. Natl. Acad. Sci. USA **99:** 4290–4295
37 Elcock A. H. (2001) Prediction of functionally important residues based solely on the computed energetics of protein structure. J. Mol. Biol. **312:** 885–896
38 Ma B., Wolfson H. J. and Nussinov R. (2001) Protein functional epitopes: hot spots, dynamics and combinatorial libraries. Curr. Opin. Struct. Biol. **11:** 364–369
39 Joshi M. D., Sidhu G., Nielsen J. E., Brayer G. D., Withers S. G. and McIntosh L. P. (2001) Dissecting the electrostatic interactions and pH-dependent activity of a family 11 glycosidase. Biochemistry. **40:** 10115–10139
40 Farr-Jones S., Wong W. Y. L., Gutheil W. and Bachovchin W. W. (1993) Direct observation of the tautomeric forms of histidine in 15N NMR spectra at low temperatures. J. Am. Chem. Soc. **115:** 6813–6819
41 Holmquist M. (2000) Alpha/beta hydrolase fold enzymes: structures, functions and mechanisms. Curr. Protein Pept. Sci. **1:** 209–235

42 Di Cera E., and Krem M. M. (in press) Serine proteases. In: Encyclopedia of the Human Genome, Cooper D. N. (ed.), Macmillan, London
43 Holyoak T., Wilson M. A., Fenn T. D., Kettner C. A., Petsko G. A., Fuller R. S. et al. (2003) 2.4 Å resolution crystal structure of the prototypical hormone-processing protease Kex2 in complex with an Ala-Lys-Arg boronic acid inhibitor. Biochemistry **42:** 6709–6718
44 Henrich S., Cameron A., Bourenkov G. P., Kiefersauer R., Huber R., Lindberg I. et al. (2003) The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. Nat. Struct. Biol. **10:** 520–526
45 Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H. et al. (2000) The Protein Data Bank. Nucleic Acids Res. **28:** 235–242
46 Bode W., Papamokos E. and Musil D. (1987) The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech *Hirudo medicinalis*. structural analysis, subtilisin structure and interface geometry. Eur. J. Biochem. **166:** 673
47 Zhou G., Somasundaram T., Blanc E., Parthasarathy G., Ellington W. R. and Chapman M. S. (1998) Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. Proc. Natl. Acad. Sci. USA **95:** 8449–8454
48 Coll M., Knof S. H., Ohga Y., Messerschmidt A., Huber R., Moellering H. et al. (1990) Enzymatic mechanism of creatine amidinohydrolase as deduced from crystal structures. J. Mol. Biol. **214:** 597–610
49 Li Y. F., Hata Y., Fujii T., Hisano T., Nishihara M., Kurihara T. et al. (1998) Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism. J. Biol. Chem. **273:** 15035–15044
50 Ridder I. S., Rozeboom H. J., Kalk K. H. and Dijkstra B. W. (1999) Crystal structures of intermediates in the dehalogenation of haloalkanoates by L-2-haloacid dehalogenase. J. Biol. Chem. **274:** 30672–30678