

ARTICLE

Electrostatic complementarity in pseudoreceptor modeling based on drug molecule crystal structures: the case of loxistatin acid (E64c)

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,

Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Ming W. Shi,^a Alexandre N. Sobolev,^a Tanja Schirmeister,^b Bernd Engels,^c Thomas C. Schmidt,^c Peter Luger,^d Stefan Mebs,^d Birger Dittrich,^e Yu-Sheng Chen,^f Joanna M. Bąk,^a Dylan Jayatilaka,^a Charles S. Bond,^a Michael J. Turner,^a Scott G. Stewart,^a Mark A. Spackman,^a and Simon Grabowsky^{a,g,*}

After a long history of use as a prototype cysteine protease inhibitor, the crystal structure of loxistatin acid (E64c) is finally determined experimentally using intense synchrotron radiation, providing insight into how the inherent electronic nature of this protease inhibitor molecule determines its biochemical activity. Based on the striking similarity of its intermolecular interactions with those observed in a biological environment, the electrostatic potential of crystalline E64c is used to map the characteristics of a pseudo-enzyme pocket.

Introduction

Loxistatin acid (E64c, Figure 1a) and its ethyl ester loxistatin (E64d) derive from the natural product E64, an epoxysuccinylpeptide.¹ Both are potent cysteine protease inhibitors which served as leads for various derivatives and which are widely used in in-vitro and in-vivo studies.² In both cases, E64c is the active component (as hydrolysis product of E64d) that irreversibly inhibits the protease through nucleophilic ring opening at carbon atom C2 and formation of a covalent S-C2 bond (Figure 1b).^{3,4} Several crystal structures of E64c-enzyme complexes have been determined³ - of which we use the E64c-cathepsin B structure of ref. 3c in this study (Figure 1d) – but the crystal structure of pure E64c has not been obtained until now. Substituent and environmental effects on the activity of E64c and the inhibition mechanism have been analyzed theoretically.⁴

The binding of a drug to its receptor and the way in which a molecule packs in a crystal are governed by the same interactions, principally electrostatic forces, hydrogen bonding and van-der-Waals interactions. In biological as well as crystalline environments, a molecule will adapt its shape according to the prevailing binding requests so that the observed conformation is a measure of both its inherent molecular flexibility and the surrounding binding forces. Correlations between geometrical properties derived from small-molecule crystal structures and the observed binding modes of the same molecule in complexes with enzymes are

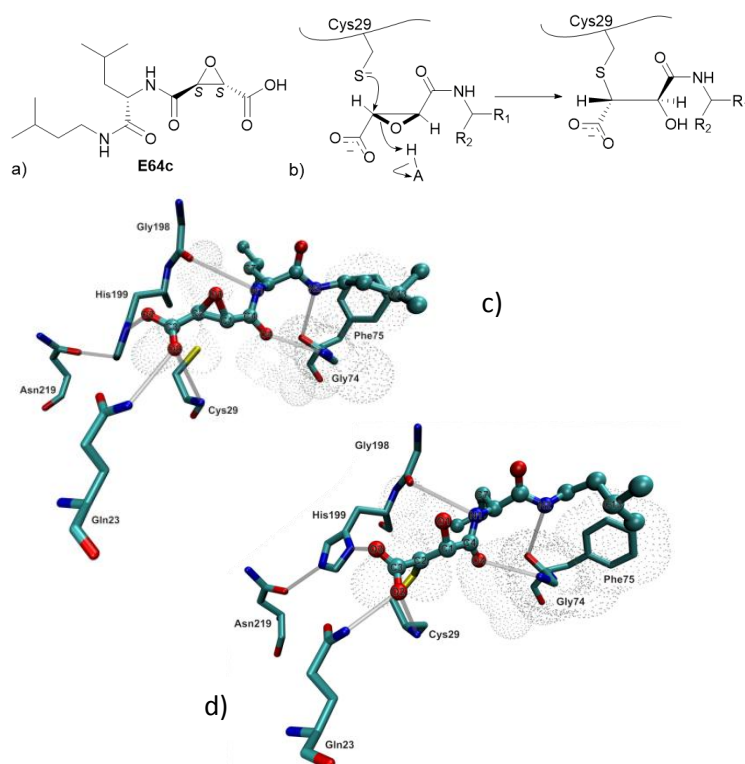


Fig. 1 a) structure; b) sketch of irreversible cysteine inhibition; c) simulated structure^{4c} of the reversible E64c – cathepsin B complex before covalent inhibition; d) crystal structure of the E64c – cathepsin B complex (irreversible complex after covalent inhibition).

numerous and promising for further application,⁵ and these have led to the development of the idea of composite crystal-field environments⁶ and development of software such as SuperStar⁷ connected to the Cambridge Structural Database.⁸ This enables rational drug design based on pharmacophores.⁹ However, the pharmacophore approach requires the knowledge of the three-dimensional structure of a specific biological target. If this is experimentally unavailable, generalized structural correlations of the kind discussed above are utilized to simulate a receptor or enzyme-pocket geometry based on arrays of crystal structures of corresponding small biologically active compounds as probes (pseudoreceptor modeling¹⁰ or reverse pharmacophore mapping¹¹).

Beyond the level of geometrical and force-field applications discussed in the previous paragraph,^{5,6,10} experimental electron-density (ED) properties of crystal structures of pure low-molecular weight enzyme ligands give detailed insight into intermolecular interactions that are believed to be biologically relevant.^{13,14b} Quantitative studies have revealed the similarity between the ED of a molecule in the biological and crystal environments using E64c as a test case.^{4c,14} The results indeed imply that the environments inside the molecular crystal and in the target enzyme generally influence the ED of the active compound in a closely similar manner.^{4c,14} However, a serious drawback of these studies was that comparisons had to be made between the crystal structures of E64c-cathepsin B and pure E64d,¹⁵ since no crystal structure of E64c was available at that time. With the knowledge of the E64c crystal structure, these assumptions based on the ED can now be checked and comparisons can be extended to include the electrostatic potential. It is increasingly recognised that electrostatic complementarity is a crucial property for crystal packing^{12a,b} as well as for the biological ligand-enzyme recognition process in general upon forming a reversible complex^{12c-g}. Hence, in this study we combine the concepts of pseudoreceptor modeling and electrostatic complementarity.

Results and discussion

In the E64c-cathepsin B complex, covalently-bonded E64c is present after ring opening (irreversible complex, Figure 1d), whereas in the crystal structure of the pure inhibitor (E64d in previous studies, E64c in this study) the epoxide ring is closed and hence more closely resembles the state of the molecule in the reversible complex during the recognition process, and before covalent inhibition. Therefore, this reversible complex has been modelled with QM/MM means (Figure 1c, details in ref. 4c). In the present study we use the new crystal structure of E64c and the modelled E64c-cathepsin B reversible complex for a direct comparison of their geometry, hydrogen-bonding environment and electrostatic properties.

Comparison of Figures 1c and 1d reveals that the conformational changes upon ring opening are minimal, including residues involved in the hydrogen bonding network,

and suggesting that the properties discussed below may even be an adequate approximation for the irreversible complex.

E64c was purchased as a finely crystalline solid from Cayman Chemicals. Once dissolved in any solvent, normally only glass-like material forms. We managed to crystallize E64c from a few milliliters of pure ethanol by extremely slow evaporation over three months. Very thin (less than 20 μm) plate-like crystals were obtained that readily formed twins. As a diffraction pattern could not be observed at atomic resolution with conventional X-ray sources, single-crystal diffraction data were measured with intense synchrotron radiation and at the lowest possible temperature (12K, open-flow helium gas stream) at the Advanced Photon Source of the Argonne National Laboratories in Illinois, U.S.A, beamline 15-ID-B. A maximum resolution of $d = 0.7 \text{ \AA}$ was achieved. Problems persisted since absent reflections suggested the orthorhombic space groups C222 or P2₁2₁2. After solving the structure in the triclinic space group P1, pseudo-symmetry relationships became apparent between four independent molecules in the asymmetric unit (Figure 2a, atom numbering in 2b) with disorder in the iso-butyl/-pentyl groups for two of them. Invarions were used for an improvement of the structural parameters and the assignment of the disorder components. This technique is based on tabulated transferable multipole parameters and hence allows the use of aspherical atomic scattering factors in the least-squares refinement.¹⁶ Pertinent crystallographic and refinement details are listed in the experi-

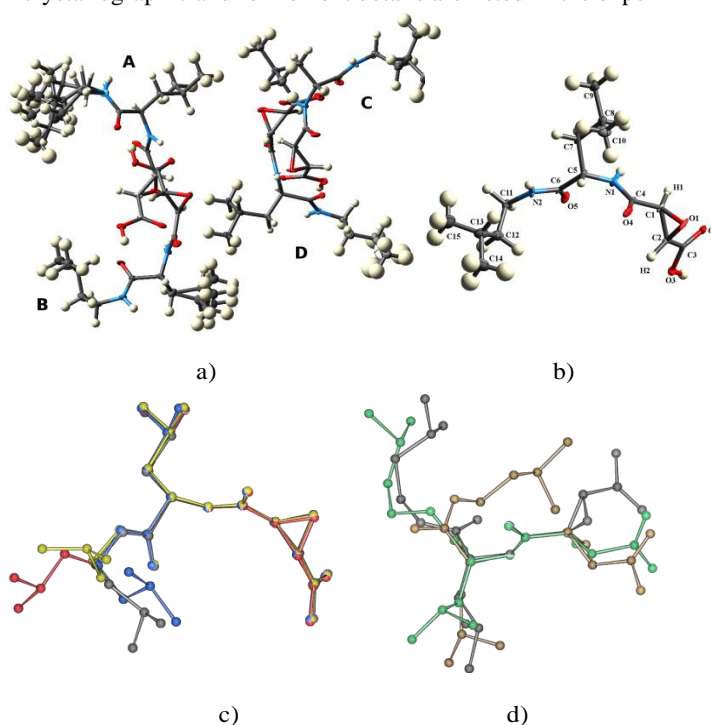


Fig. 2 a) Asymmetric unit of the crystal structure of pure E64c with four symmetry-independent molecules; b) disorder-free molecule D showing atom numbering scheme; c) overlay plot of the four independent E64c molecules in the crystal structure (hydrogens omitted, only major disorder components shown); d) overlay plot of E64c from the crystal structure of the pure compound (molecule D, grey), from the reversible E64c - cathepsin B complex (green), and from isolated-molecule optimization (light brown).

mental section and the ESI. An isolated-molecule geometry optimization was also carried out at the B3LYP/6-311++G(2d,2p) level of theory.

Figure 2c shows that the conformations of the four independent molecules A-D are virtually identical with the exception of the iso-pentyl group (C11-C15). This conformational flexibility and the fact that this group is prone to be heavily disordered (three different disorder components in molecule A) show that it lacks some stabilizing intermolecular interactions in the crystal. In the enzyme, it is this iso-pentyl group that is closest to the hydrophobic pocket, represented by residue Phe75 in Figures 1c,d and 3b. This is a first hint that the behavior of the molecule in the crystal structure of the pure compound is comparable to that in the enzyme environment. Hydrogen bonding interactions for molecules A-D are found to be identical in the crystal (see Table 1 and the Hirshfeld surface¹⁷ fingerprint plots^{17a} in Figure S4 in the ESI), hence the biologically active regions around the epoxide rings lie in the same crystal environment in all cases, leading to identical conformations (Figure 2c). For these reasons we choose only one of them (disorder-free molecule D, Figure 2b) as a basis for further analysis.

Table 1 Donor-acceptor distances for hydrogen bonds in the crystal structure of E64c, the reversible^{4c} and the irreversible E64c – cathepsin B complex^{3c}

Donor-hydrogen...acceptor	Distance donor-acceptor in Å
E64c crystal structure:	
O3A—H...O5B	2.536(7)
O3B—H...O5A	2.522(7)
O3C—H...O5D	2.544(6)
O3D—H...O5C	2.542(7)
N1A—H...O2B	2.842(7)
N1B—H...O2A	2.826(8)
N1C—H...O2D	2.838(8)
N1D—H...O2C	2.826(8)
C1A—H...O2B	3.296(7)
C1B—H...O2A	3.284(8)
C1C—H...O2D	3.274(8)
C1D—H...O2C	3.279(8)
N2A—H...O4C ^a	2.855(7)
N2B—H...O4D ^b	2.928(7)
N2C—H...O4A ^c	2.916(7)
N2D—H...O4B ^d	2.887(7)
Reversible complex:	
His ¹⁹⁹ N—H...O3	2.68
Gln ²³ N—H...O2	3.16
Cys ²⁹ N—H...O2	2.97
Gly ⁷⁴ N—H...O4	2.94
N1—H...OGly ¹⁹⁸	4.02
N2—H...OGly ⁷⁴	2.94
Irreversible complex:	
His ¹⁹⁹ N—H...O3	3.18
Gln ²³ N—H...O2	2.84
Cys ²⁹ N—H...O2	3.07
Gly ⁷⁴ N—H...O4	3.12
N1—H...OGly ¹⁹⁸	3.81
N2—H...OGly ⁷⁴	3.69

^a x, -1+y, z; ^b 1+x, y, -1+z; ^c -1+x, 1+y, z; ^d x, y, 1+z; rest: intra-asu

Figure 2d overlays the molecular conformations of E64c from the crystal of the pure compound, the reversible complex with cathepsin B, and the isolated-molecule optimization. There is a considerable twist in the orientation of the epoxide ring between all three of them if the amide group is used as refe-

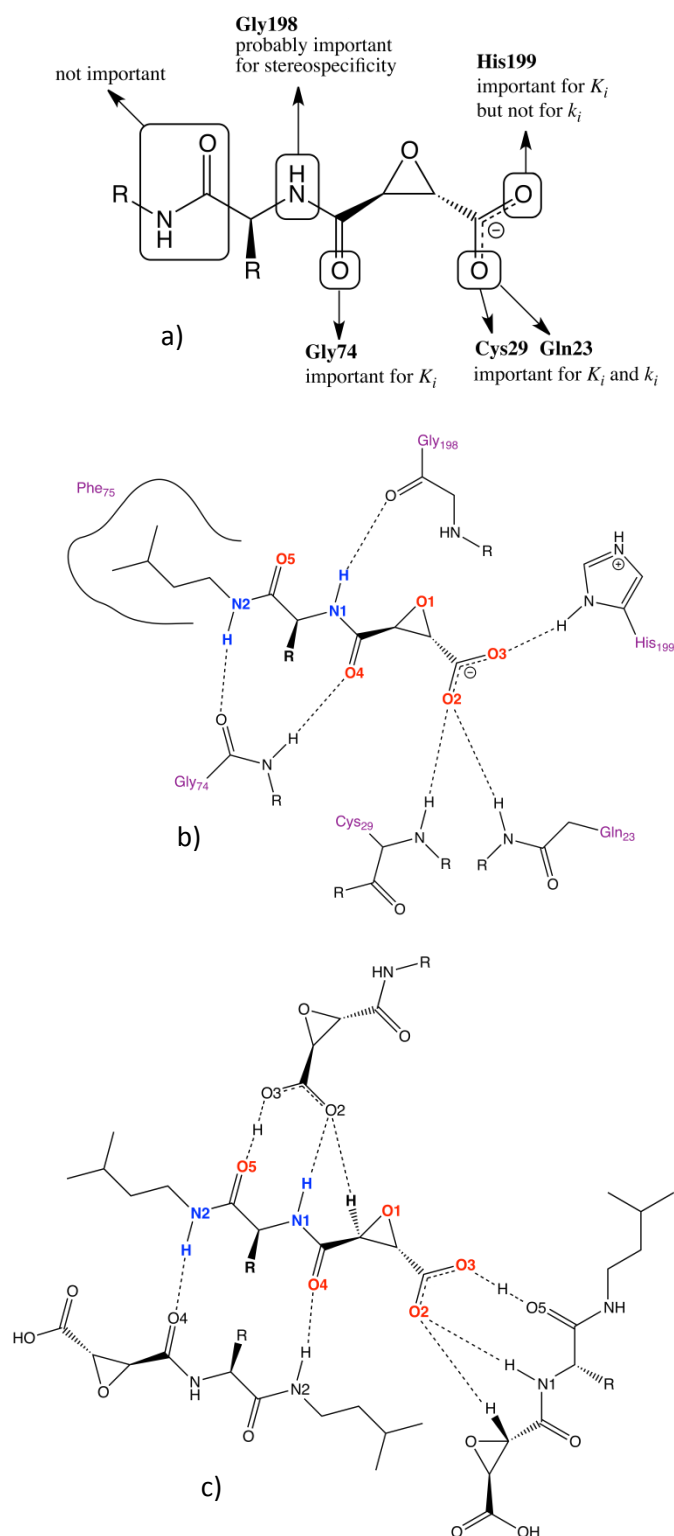


Fig. 3 a) Biochemical importance of various hydrogen bonds of the cathepsin B – E64c complex according to ref. 4a. (For definitions of K_i and k_i see text.) Hydrogen bond patterns of E64c in b) the reversible cathepsin B – E64c complex, c) the crystal structure of pure E64c (central molecule D).

rence, which is important if one considers that C2 is attacked in the inhibition mechanism. However, the overall molecular T-like shape is retained between the molecule in the crystal and

the enzyme environment, whereas in the isolated state the optimum structure is less open with one of the alkyl chains looping back onto the biologically active region and shielding it sterically. It appears that interactions with the environment in the crystal and the enzyme overcome this energy barrier and open the molecule. This must be due to an overall similar electric field acting on the molecule exerted by neighboring groups (other E64c molecules or amino acid residues, respectively) of which hydrogen bonds are an important part.

Assuming a two-step inhibition process where first the reversible E64c–cathepsin B complex is formed (K_i = dissociation constant of the reversible complex) and then irreversible inhibition follows (k_i = first-order rate constant of the inhibition), previous work^{4a} has identified several hydrogen-bonding interactions with the protease to be of crucial importance for the biological activity of E64c (compare Figures 3a,b and Table 1). A bifurcated hydrogen bond of O2 with residues Gln23 and Cys29 is important for both K_i and k_i , whereas the hydrogen bonds between O3 and His199 as well as O4 and Gly74 are important only for K_i . A hydrogen bond between N1-H and Gly198 probably determines the stereospecificity, but interactions with the amide group O5/N2 are less important. Epoxide oxygen O1 is not involved in any interaction. Interestingly, in the crystal structure of the pure compound, the same number and configuration of the biochemically important hydrogen bonds are observed (Figure 3c and Table 1), including a bifurcated hydrogen bond with O2 as acceptor, furthermore involving O3, O4, and N1-H, but not involving O1. However, the second interaction to O2 is a C-H...O hydrogen bond, and O3 is a donor and not a deprotonated acceptor (although the O3...O5 distance is so short at 2.54 Å that the beginning of a proton transfer process may be frozen in the crystal structure, indicated with dotted lines in Figure 3c). Overall, the arrangement of pairs of antiparallel molecules exemplified in Figure 3c is reminiscent of two antiparallel peptides in a beta-pleated sheet. Table 1 shows that not only the number and configuration of the hydrogen bonds are similar between the crystal structure of the pure compound and the reversible complex, but also the donor-acceptor distances are very similar, even reflecting similar trends, e.g. that the hydrogen bond involving O3 is the shortest one in the assembly. Although the same hydrogen bonds persist after the inhibition (cf. Figures 1c,d), in terms of distances the similarity to the reversible complex is higher than to the irreversible one, as expected.

In summary, the similarity between the hydrogen bonding patterns of E64c in its crystal structure and in the enzyme environment is striking, and is complemented by the similar hydrophobic interactions discussed above. Considering these findings, we conclude as a first assumption that inherent electronic and electrostatic properties of E64c are both important in the inhibition process and reflected in the properties of the crystal of the pure compound. We know from previous work^{4,14} that electron-density redistributions upon contact with the protease do occur in E64c and that they are not negligible, however, those results were based on a comparison

with E64d and not E64c in its native form. We therefore proceed with an analysis of the electrostatic potential (ESP) of E64c that is easily, quickly, but accurately derived from the crystal structure of the pure compound and outline how information about the electrostatics of the enzyme–ligand interaction is readily simulated as a result.

The ESP has been generated for E64c in its crystal structure in two different ways: (i) using the theoretical multipole parameters imposed on the experimental geometry of the molecule through the invariom approach (Figure 4a, showing a single molecule), and (ii) calculated from an isolated-molecule wavefunction for the experimental geometry (Figure 4b, showing two neighboring molecules). In both cases the ESP has been mapped on the Hirshfeld surface (HS) of E64c in its crystal packing,¹² and a more direct comparison between the two approaches is shown in the ESI (Figure S5). The separation between areas of negative (red) and positive (blue) ESP on the HS is more pronounced using the invariom approach (Figure 4a), but generally the epoxysuccinylpeptide region is of highly negative ESP around the hetero atoms, of slightly positive ESP around the epoxide carbon atoms (back side of Figure 4a) and of highly positive ESP around both alkyl chains. These broad features in the ESP determine the crystal packing in that HS patches of complementary ESP are in contact with each other (Figure 4b), leading to a continuity of the zero-potential (white) lines (nodal lines), and minimizing the lattice energy. This remarkable continuity of the nodal lines across neighboring

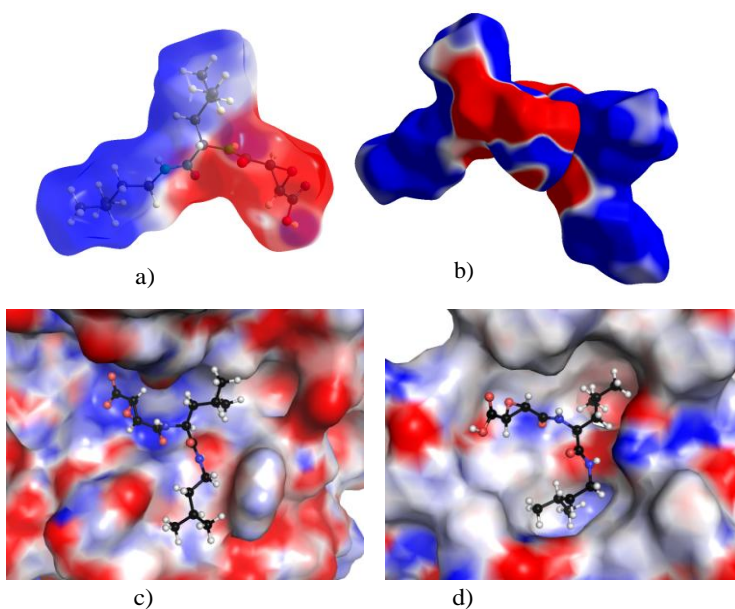


Fig. 4 The electrostatic potential (ESP) mapped onto molecular surfaces of E64c. a) ESP generated from the invariom approach and mapped onto a Hirshfeld surface; b) ESP calculated within CrystalExplorer (isolated molecules) showing electrostatic complementarity of touching Hirshfeld surface patches in the crystal packing; c) ESP of cathepsin B estimated with the APBS method mapped onto a Connolly surface; d) ESP of a cluster of E64c molecules in the crystal structure of the pure compound estimated with the APBS method mapped onto a Connolly surface simulating a native pocket for E64c. (Color code: ± 0.01 au in a) and b); ± 20 kT/e in c); ± 5 kT/e in d); red = negative, blue = positive, white = zero).

molecules supports our assertion that the electronic/electrostatic nature of unperturbed molecules largely governs the crystal packing - since the ESP in both approaches has been derived for the isolated molecules.

Pictures similar to the one in Figure 4c are frequently used in protein crystallography in order to graphically display results, where the ESP of the active pocket is estimated by solving the Poisson-Boltzmann equation in aqueous media through the Adaptive Poisson-Boltzmann Solver (APBS) method¹⁸ and mapped onto Connolly surfaces¹⁹. This, of course, is a considerable simplification in order to deal with large structures. However, in order to determine how far the present comparison between crystal and enzyme environments can be taken, we have generated a cluster of symmetry-related E64c molecules at their experimental geometries in the crystal of the pure compound, calculated atomic charges for every symmetry-independent molecule using the ChelpG approach²⁰ and plotted the APBS ESP on a Connolly surface in the same way as routinely done for enzymes (Figure 4d, see also Figures S6 to S8). This serves as a simulation of a native pocket for the E64c molecule mirroring its most important inherent electronic features in the ESP distribution of the environment.

The ESP of the enzyme pocket in Figure 4c is generated without any knowledge of the inhibitor molecule; only atomic charges of the amino acid residues are accounted for. In contrast, the information about the pocket shape and ESP distribution in the pocket in Figure 4d stem exclusively from the molecular shape and atomic charges of the central small biologically active compound itself (if simplified to the more usual case of $Z'=1$ small-molecule crystal structures). In other words, the environment is mapped as a complement to the ESP distribution of the small molecule, a combination of the principles of pseudoreceptor modeling and electrostatic complementarity.

Differences between the shape and ESP distribution of the enzyme pocket (Figure 4c) and the pocket in the crystal structure of pure E64c (Figure 4d) include more localized ESP patches of high absolute values in 4d (different color scale, other comparisons shown in the ESI, Figure S8), the different conformation of the iso-pentyl groups (cf. Figure 2d) and the protonation state of the carboxylate group. The latter point rationalizes the positive ESP on the pocket below the carboxylate in Figure 4c in response to the negative charge, but the negative ESP in 4d in response to the highly acidic proton (cf. Figure 3). The shape similarity of the pockets cannot be totally dismissed. In both representations, there is an opening to the upper left and invaginations at the top, the right and the lower left. In both environments the ESP is close to neutral around the alkyl chains, highly negative below the N-H bond of the amide in the side chain and highly positive below the amide group next to the epoxide. These similarities pertain to features very important for the biological activity and reflect the similarities in the hydrogen bonding as well as hydrophobic interactions.

Conclusions

In this work we propose to use the characteristics of the molecular crystal environment as a pseudo-enzyme pocket involving electrostatic complementarity considerations. This can become a helpful tool for the design of new inhibitor molecules. However, its success will be limited by the presence of a ligand-specific induced-fit mechanism and by significant solvation phenomena. Moreover, for generalized conclusions relevant to biological scenarios a more extensive exploration of many more cases will certainly be needed, where both the small-molecule and enzyme-complex crystal structures are known. For the present case study of E64c, similarities between both situations in terms of the conformation, hydrogen-bonding environment and ESP of E64c show that further work in this direction is likely to be fruitful. This is especially true since the method is an enhancement of conventional reverse pharmacophore mapping or pseudoreceptor modeling techniques that only employ geometric properties.

Experimental Part

The following details pertain to the measurement of the data at the synchrotron beamline 15-ID-B of the APS: $\lambda = 0.3936\text{\AA}$, $T = 12\text{K}$, $d = 0.70\text{\AA}$, no. collected reflections = 80246, no. unique reflections = 9541, no. observed reflections used in the refinement = 8764 ($F > 4\sigma(F)$), completeness = 87%, $R_{\text{int}} = 8.09\%$. Crystal data are as follows: space group = P1, $Z = 4$, $a = 10.036(2)\text{\AA}$, $b = 11.838(2)\text{\AA}$, $c = 15.884\text{\AA}$, $\alpha = 79.36(3)^\circ$, $\beta = 78.98(3)^\circ$, $\gamma = 74.71(3)^\circ$, $V = 1768.7(6)\text{\AA}^3$. The final invariom refinement on F^2 magnitudes led to the following figures of merit: $R(F) = 10.23\%$, $wR(F^2) = 19.74\%$, min/max residual density = $-0.45/0.68\text{ e/\AA}^3$. The cif file including all experimental details has been deposited with the Cambridge Structural Database under deposition number CCDC-977799, and can be downloaded free of charge under www.ccdc.cam.ac.uk/data_request/cif.

The crystal structure was solved with SHELXS²¹, refined with InvariomTool²² and XD2006²³, and analyzed with CrystalExplorer²⁴ as well as MOLISO²⁵. Within CrystalExplorer, the electrostatic potential (ESP) was calculated using HF/6-31G(d). Gaussian03²⁶ was used for the isolated-molecule optimization, and for generating ChelpG¹⁵ charges for the four different experimental geometries of E64c (HF/6-31G(d)). Charges for the residues of cathepsin B were transferred from a library of AMBER²⁷ charges. The program PyMOL (Schrödinger, LLC) was used for the APBS¹⁸ estimation of the ESP based on those charges.

Acknowledgements

We acknowledge financial support from the Australian Research Council within projects DP110105347,

DP130103488, and DE140101330 as well as from the Deutsche Forschungsgemeinschaft within the special priority program SPP1178 and the Emmy Noether grant GR 4451/1-1. ChemMatCARS is principally supported by the Divisions of Chemistry and Materials Research, National Science Foundation, under grant number NSF/CHE-1346572. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

Notes and references

- ^a The University of Western Australia, School of Chemistry and Biochemistry, M310, 35 Stirling Highway, Crawley WA 6009, Australia.
- ^b Johannes-Gutenberg-Universität Mainz, Institut für Pharmazie und Biochemie, Staudinger Weg 5, 55128 Mainz, Germany.
- ^c Julius-Maximilians-Universität Würzburg, Institut für Physikalische und Theoretische Chemie, Emil-Fischer-Str. 42, 97074 Würzburg, Germany.
- ^d Freie Universität Berlin, Institut für Chemie und Biochemie, Fabeckstr. 36a, 14195 Berlin, Germany.
- ^e Universität Hamburg, Institut für Anorganische und Angewandte Chemie, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany.
- ^f ChemMatCARS, Center for Advanced Radiation Sources, c/o Advanced Photon Source/ ANL, The University of Chicago, 9700 South Cass Avenue, Bldg. 434D, Argonne, Illinois 60439, U.S.A.
- ^g Universität Bremen, Fachbereich 2 – Biologie/Chemie, Leobener Str. NW2, 28359 Bremen, Germany.
- *E-Mail: simon.grabowsky@uni-bremen.de
- Electronic Supplementary Information (ESI) available: Full experimental details, geometry tables, HS fingerprint plots, additional ESP representations, invariom electron-density maps and corresponding topological analyses. See DOI: 10.1039/c000000x/
- 1 K. Hanada, M. Tamai, M. Yamagishi, S. Ohmura, J. Sawada, I. Tanaka, *Agric. Biol. Chem.*, 1978, **42**, 523-528.
 - 2 (a) S. Hashida, T. Towatari, E. Koninami and N. Katunuma, *J. Biochem.*, 1980, **88**, 1805-1811; (b) T. Noda, K. Kazuhide, N. Katunuma, Y. Tarumoto and M. Ohezeki, *J. Biochem.*, 1981, **90**, 893-896; (c) K. Suzuki, *J. Biochem.*, 1983, **93**, 1305-1312; (d) D. H. Hu, S. Kimura, S. Kawashima, K. Maruyama, *Zool. Sci.*, 1989, **6**, 797-901; (e) Z. Huang, E. B. McGowan and T. C. Detwiler, *J. Med. Chem.*, 1992, **35**, 2048-2054; (f) H. H. Otto, T. Schirmeister, *Chem. Rev.* 1997, **97**, 133-171; (g) J. C. Powers, J. L. Asgian, Ö. D. Ekici, K. E. James, *Chem. Rev.* 2002, **102**, 4639-4750.
 - 3 (a) K. Matsumoto, D. Yamamoto, H. Ohishi, K. Tomoo, T. Ishida, M. Inoue, T. Sadatome, K. Kitamura and H. Mizuno, *FEBS Lett.*, 1989, **245**, 177-180; (b) M. J. Kim, D. Yamamoto, K. Matsumoto, M. Inoue, T. Ishida, H. Mizuno, S. Sumiya and K. Kitamura, *Biochem. J.*, 1992, **287**, 797-803; (c) A. Yamamoto, K. Tomoo, K. Matsugi, T. Hara, Y. In, M. Murata, K. Kitamura, T. Ishida, *Biochim. Biophys. Acta*, 2002, **1597**, 244-251.
 - 4 (a) M. Mladenovic, K. Junold, R. F. Fink, W. Thiel, T. Schirmeister, B. Engels, *J. Phys. Chem. B.*, 2008, **112**, 5458-5469; (b) M. Mladenovic, K. Ansorg, R. F. Fink, W. Thiel, T. Schirmeister, B. Engels, *J. Phys. Chem. B.*, 2008, **112**, 11798-11808; (c) M. Mladenovic, M. Arnone, R. F. Fink, B. Engels, *J. Phys. Chem. B.*, 2009, **113**, 5072-5082.
 - 5 Summaries in (a) C. Pascard, *Acta Cryst. D*, 1995, **51**, 407-417; (b) G. Klebe, *Structure Correlation and Ligand/Receptor Interactions in Structure Correlation*, vol. 1, (Eds. H.-B. Bürgi, J. D. Dunitz), chapter 13, Wiley-VCH, Weinheim, 2008.
 - 6 (a) G. Klebe, *J. Mol. Biol.*, 1994, **237**, 212-235; (b) M. Böhm, G. Klebe, *J. Med. Chem.*, 2002, **45**, 1585-1597.
 - 7 M. L. Verdonk, J. C. Cole, P. Watson, V. Gillet, P. Willett, *J. Mol. Biol.*, 2001, **307**, 841-859.
 - 8 F. H. Allen, *Acta Cryst. B*, 2002, **58**, 380-388.
 - 9 J. H. Van Drie, *Internet Electron. J. Mol. Des.*, 2007, **6**, 271-279.
 - 10 (a) Y. Tanrikulu, G. Schneider, *Nat. Rev. Drug Discov.*, 2008, **7**, 667-677; (b) A. Vedani, P. Zbinden, J. P. Snyder, P. A. Greenidge, *J. Am. Chem. Soc.*, 1995, **117**, 4987-4994.
 - 11 (a) X. Liu, S. Ouyang, B. Yu, Y. Liu, K. Huang, J. Gong, S. Zheng, Z. Li, H. Li, H. Jiang, *Nucleic Acids Research*, 2010, **38**, W609-W614; (b) P. Thipnate, J. Liu, S. Hannongbua, A. J. Hopfinger, *J. Chem. Inf. Model.*, 2009, **49**, 2312-2322.
 - 12 (a) M. A. Spackman, J. J. McKinnon, D. Jayatilaka, *Cryst. Eng. Comm.*, 2008, **10**, 377-388; (b) M. A. Spackman, *Charge Densities and Crystal Engineering*, in: *Modern Charge-Density Analysis*, (Eds. C. Gatti, P. Macchi), Springer, 2012, pp. 553-572; (c) G. Naray-Szabo, G. G. Ferenczy, *Chem. Rev.*, 1995, **95**, 829-847; (d) J. S. Murray, P. Politzer, *WIREs Comput. Mol. Sci.*, 2011, **1**, 153-163, and references therein; (e) S. Mebs, A. Lüth, P. Luger, *Bioorg. Med. Chem.*, 2010, **18**, 5965-5974; (f) C. F. Matta, A. A. Arabi, D. F. Weaver, *Eur. J. Med. Chem.*, 2010, **45**, 1868-1872; (g) C. F. Matta, *J. Comput. Chem.*, 2014, **35**, 1165-1198, and references therein.
 - 13 (a) P. Luger, *Org. Biomol. Chem.*, 2007, **5**, 2529-2540; (b) P. Luger, M. Weber, B. Dittrich, *Future Med. Chem.*, 2012, **4**, 1399-1407.
 - 14 (a) B. Engels, T. C. Schmidt, C. Gatti, T. Schirmeister, R. F. Fink, *Struct. Bond.*, 2012, **147**, 47-98; (b) S. Grabowsky, D. Jayatilaka, R. F. Fink, T. Schirmeister, B. Engels, *Z. Anorg. Allg. Chem.*, 2013, **639**, 1905-1921.
 - 15 T. Ishida, M. Sakaguchi, D. Yamamoto, M. Inoue, K. Kitamura, K. Hanada and T. Sadatome, *J. Chem. Soc. Perkin Trans.*, 1988, **6**, 851-857.
 - 16 (a) B. Dittrich, C. B. Hübschle, K. Pröpper, F. Dietrich, T. Stolper, J. J. Holstein, *Acta Cryst. B*, 2013, **69**, 91-104; (b) instructions for invariom applications to disordered compounds are to be published (C. B. Hübschle, C. Schuermann, B. Dittrich).
 - 17 (a) M. A. Spackman, J. J. McKinnon, *Cryst. Eng. Comm.*, 2002, **4**, 378-392; (b) M. A. Spackman, D. Jayatilaka, *Cryst. Eng. Comm.*, 2009, **11**, 19-32.
 - 18 N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**, 10037-10041.
 - 19 M. L. Connolly, *Science*, 1983, **221**, 709-713.
 - 20 C. M. Breneman, K. B. Wiberg, *J. Comp. Chem.*, 1990, **11**, 361-373.
 - 21 G. M. Sheldrick, *Acta Cryst. A*, 2008, **64**, 112-122.
 - 22 C. B. Hübschle, P. Luger, B. Dittrich, *J. Appl. Cryst.*, 2007, **40**, 623-627.
 - 23 A. Volkov, P. Macchi, L. J. Farrugia, C. Gatti, P. R. Mallinson, T. Richter, T. S. Koritsanszky, *XD2006*, User Manual, 2006.

- 24 S. K. Wolff, D. J. Grimwood, J. J. McKinnon, M. J. Turner, D. Jayatilaka, M. A. Spackman, CrystalExplorer, version 3.1, University of Western Australia, 2012.
- 25 C. B. Hübschle, P. Luger, *J. Appl. Cryst.*, 2006, **39**, 901-904.
- 26 Gaussian 03, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
- 27 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179-5197.