Supplementary Information

2-mercaptobenzoxazoles: a class of carbonic anhydrase inhibitors with a novel binding mode to the enzyme active site

Murat Bozdag,[a] Claudiu T. Supuran,[a] Davide Esposito,[b] Andrea Angeli,[a] Fabrizio Carta,[a] Simona Maria Monti,[b] Giuseppina De Simone[b*] and Vincenzo Alterio[b*]

[a] Neurofarba Department, Section of Pharmaceutical and Nutriceutical Sciences, Università degli Studi di Firenze, Sesto Fiorentino, Florence, Italy.

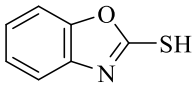
[b] Istituto di Biostrutture e Bioimmagini-CNR, Naples, Italy.

*Corresponding authors. Phone: +39-081-678674, e-mail: vincenzo.alterio@cnr.it (VA) or Phone: +39-081-2534579, e-mail: giuseppina.desimone@cnr.it (GDS)

Experimental methods

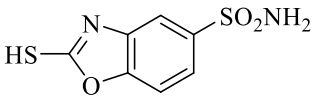
General procedure for the synthesis of benzo[d]oxazole-2-thiol compounds:

A solution of 2-aminophenol derivative (1.0 equv) in dry THF was cooled to 0˚C and thiophosgene (1.2 equv) was added drop-wise. The reaction continued overnight and then was quenched with water, and extracted with EtOAc (3 x 5.0 mL). The combined organic layers were washed with H_2O (3 x 5.0 mL), dried over Na_2SO_4, filtered and concentrated to obtain a residue that was crystallized from IPA (benzo[d]oxazole-2-thiol, **1**) or purified by silica gel column chromatography eluting with MeOH/DCM 10% *v*/*v* (2-mercaptobenzo[d]oxazole-5-sulfonamide, **2**) to obtain desired product as a pale brown solid.



benzo[d]oxazole-2-thiol (1): 45% yield; δ_H (400 MHz, DMSO-d_6) 7.27-7.36 (3H, m), 7.54 (1H, d, *J* 7.8), 13.91 (1H, s, exchange with D_2O, S*H*); δ_C (100 MHz, DMSO-d_6) 110.9, 111.4, 124.7, 126.1, 132.1, 149.1, 181.0.

Experimental data is in agreement with previously reported literature.[1]



2-mercaptobenzo[*d*]oxazole-5-sulfonamide (2): 66% yield; silica gel TLC R_f 0.3 (MeOH/DCM 10 % *v*/*v*); δ_H (400 MHz, DMSO-d_6) 7.51 (2H, s, exchange with D_2O, SO_2NH_2), 7.65 (1H, d, *J* 1.7), 7.71 (1H, d, *J* 8.5), 7.76 (1H, dd, *J* 1.7, 8.5), 14.22 (1H, s, exchange with D_2O, S*H*); δ_C (100 MHz, DMSO-d_6) 108.9, 111.2, 122.7, 132.5, 142.1, 150.8, 181.9.

Experimental data is in agreement with previously reported literature.[2]

CA inhibition

An applied photophysics stopped-flow instrument was utilized for measuring CA catalyzed CO_2 hydration by a spectrophotometric method for obtaining the inhibition constants of compounds reported here.[3] Phenol red (0.2 mM) was used as indicator with HEPES (20 mM, pH 7.5) as buffer and ionic strength was maintained constant by the use of 20 mM Na_2SO_4. CO_2 hydration rates were followed for 10–100 s, using CO_2 concentrations in the range of 1.7–17 mM. Each inhibitor was tested in triplicate, in serial dilutions starting from 0.01 to 0.1 mM. Compound **1** pre-incubated for 6h and compound **2** for 15 min with enzymes in order to be sure that the enzyme–inhibitor complexes have been formed.[4-6] The inhibition constants were obtained by using the Cheng–Prusoff equation[4-6] and are the mean from three determinations. The employed CA isoforms were recombinant proteins obtained by us as reported in earlier publications.[7-9]

Crystallization, X-ray data collection, and refinement

hCA II protein was prepared as previously described.[4] hCA II crystals were obtained by the hanging drop vapor diffusion technique, using a precipitant solution containing 1.3 M sodium citrate and 100 mM Tris-HCl, pH 8.0. hCA II/inhibitor complexes were obtained by soaking for one day hCA II preformed crystals in the crystallization solution supplemented with an inhibitor amount equal to 18 mM and 10 mM for **1** and **2**, respectively.

Diffraction data were collected at a resolution of 1.48 Å and 1.50 Å for hCA II/**1** and hCA II/**2** complexes, respectively, in-house at 100 K, using a Rigaku MicroMax-007 HF generator producing Cu Kα radiation and equipped with a Saturn 944 CCD detector. Prior to data collection, crystals were transferred into the soaking solution supplemented with 10% glycerol to prevent the formation of crystalline ice during the flash-freezing under a nitrogen stream.

Diffraction data were indexed, integrated and scaled using the HKL2000 software package.[10] Crystal parameters and relevant X-ray data collection statistics are reported in Table S1. Initial phases were calculated using as starting model for both complexes the hCA II structure solved in the $P2_1$ space

group (PDB code 6EQU)[11] after deletion of non-protein atoms. An initial round of rigid body refinement followed by simulated annealing and individual B factor refinement was performed using the program CNS.[12, 13] Several rounds of manual rebuilding were performed using the program O[14] with careful inspection of the |2Fo-Fc| and |Fo - Fc| electron density maps. In both structures, the inhibitor molecules were clearly visible in the difference |Fo - Fc| map from the early stages of refinement.

2-mercaptobenzoxazole can exist in two tautomeric forms: thiol (a) and thione (b) (Fig. S8).[15] According to the electron density maps the planar (a) conformation was modelled in both crystal structures. Restraints on bond angles and distances for the two inhibitors were taken from the Cambridge Structural Database[16] while standard restraints were used on protein bond angles and distances throughout refinement. Topology files for the inhibitors were generated using the PRODRG2 server.[17]

Water molecules were built into peaks >3σ in |Fo - Fc| maps that demonstrated appropriate hydrogen-bonding geometry. Several alternate cycles of refinement and manual model building were performed to reduce the R_{work} and R_{free} to the final values of 0.174 and 0.193 for the hCA II/**1** complex and 0.178 and 0.206 for the hCA II/**2** complex. The geometric restraints of the final models were analysed using the programs PROCHECK[18] and WHATCHECK.[19] Relevant refinement statistics can be found in Table S1. The atomic coordinates of the complexes have been deposited in the Protein Data Bank with accession codes 6YQT and 6YQU.

Table S1. Data collection and refinement statistics for hCA II/**1** and hCA II/**2** complexes.

	6YQU	**6YQT**
Crystal parameters		
Space group	$P2_1$	$P2_1$
a (Å)	42.3	42.4
b (Å)	41.3	41.5
c (Å)	72.2	72.1
α (°)	90	90
β (°)	104.4	104.3
γ (°)	90	90
Data collection statistics		
Resolution (Å)	35.6-1.48 (1.51-1.48)	35.7-1.50 (1.53-1.50)
Temperature (K)	100	100
Total reflections	153608	181761
Unique reflections	39321	38111
Completeness (%)	96.6 (80.6)	97.2 (79.0)
$\langle I \rangle / \langle \sigma(I) \rangle$	19.4 (2.6)	19.9 (2.0)
Redundancy (%)	3.9 (2.8)	4.8 (2.3)
R_{merge}[a]	0.055 (0.368)	0.063 (0.431)
R_{meas}[a]	0.063 (0.442)	0.070 (0.532)
R_{pim}[a]	0.030 (0.239)	0.029 (0.304)
Refinement statistics		
Resolution (Å)	35.6-1.48	35.7-1.50
R_{work}[b] (%)	17.4	17.8
R_{free}[b] (%)	19.3	20.6
r.m.s.d. from ideal geometry:		
Bond lengths (Å)	0.008	0.009
Bond angles (°)	1.5	1.6
Number of protein atoms	2048	2058
Number of inhibitor atoms	10	28
Number of water molecules	212	196
Average B factor (Å^2)		
All atoms	10.7	12.7
Protein atoms	9.8	12.0
Inhibitor atoms	7.8	8.8[c] (12.6[d])
Water molecules	19.1	20.8

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl)-\langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$; $R_{meas} = \Sigma_{hkl}\{n(hkl)/[n(hkl)-1]\}^{1/2}\Sigma_i|I_i(hkl)-\langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$; $R_{pim} = \Sigma_{hkl}\{1/[n(hkl)-1]\}^{1/2}\Sigma_i|I_i(hkl)-\langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for its unique reflection; summations are over all "n" reflections.

[b] $R_{factor} = \Sigma_h||F_o(h)|-|F_c(h)|| / \Sigma_h|F_o(h)|$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated with 2.4% (hCA II/**1**) or 2.6% (hCA II/**2**) of the data excluded from the refinement.

Average B factor of the inhibitor molecule [c]in the active site and [d]on the protein surface.

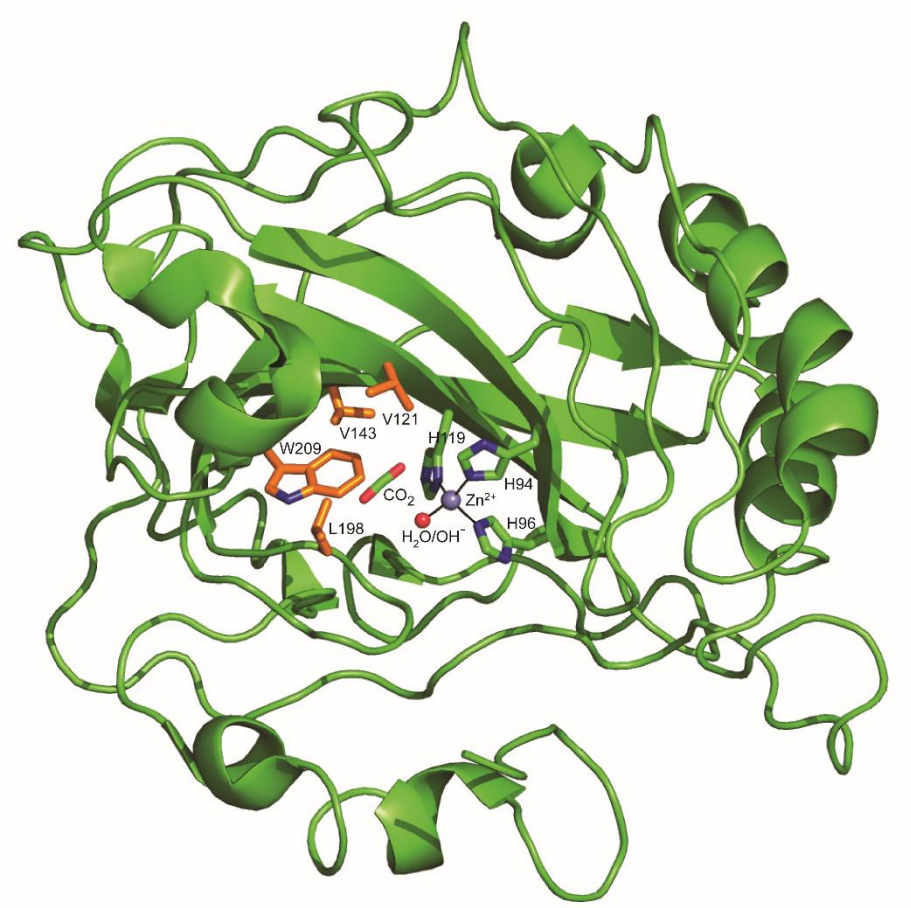


Fig. S1 Ribbon representation of hCA II, which has been chosen as representative CA. The zinc ion and its coordination, as well as residues delimiting the CO_2 binding pocket are shown in stick (PDB code 2VVA)[20].

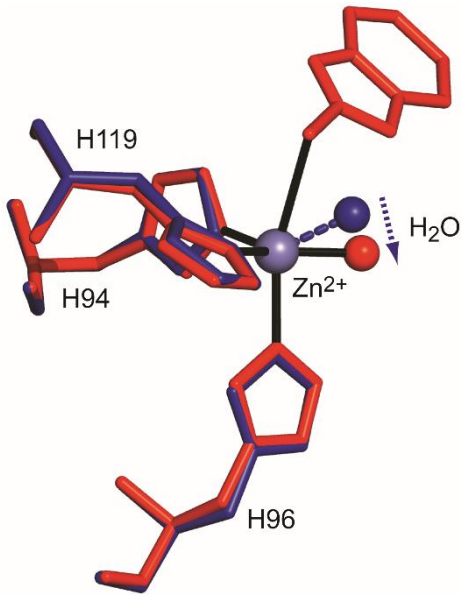


Fig. S2 Structural superposition of hCA II/**1** (red) to the native enzyme (blue, PDB code 1CA2).[21] Histidine residues coordinating the zinc ion and the inhibitor molecule are represented as sticks, while the zinc ion and the coordinating water molecule are represented as spheres. Continuous black lines indicate zinc ion coordination in the hCA II/**1** complex, while the blue dotted line indicates the Zn-O coordination distance in the native enzyme. The blue arrow indicates the displacement of the coordinating water molecule in the two structures.

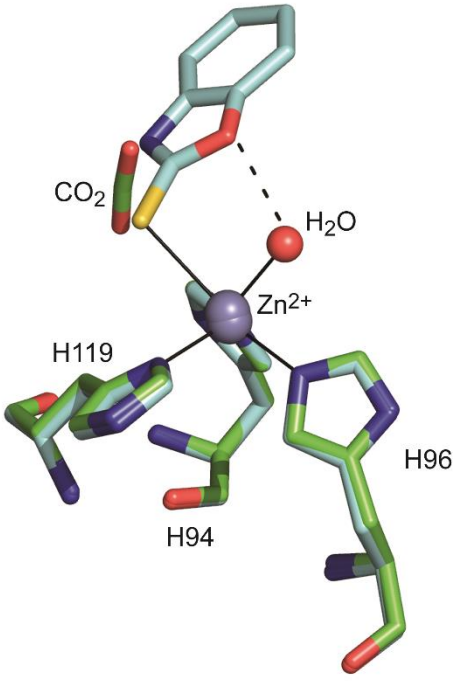


Fig. S3 Structural superposition between compound **1** (cyan) and CO_2 (green, PDB code 2VVA)[20], when bound to hCA II active site. Continuous lines indicate zinc ion coordination, whereas dashed lines indicate hydrogen bond distances.

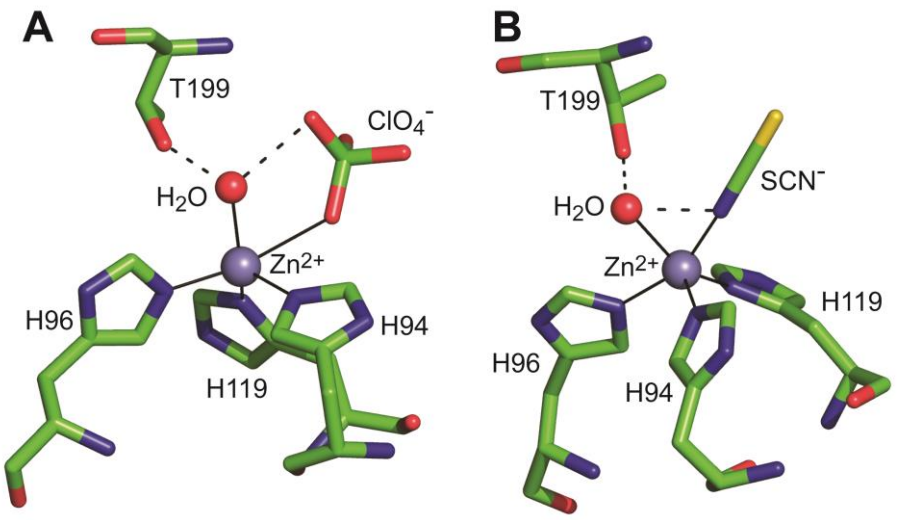


Fig. S4 Details of the interactions of perchlorate (A) (PDB code 4YGL)[22] and thiocyanate (B) (PDB code 2CA2)[23] ions with hCA II active site. Continuous lines indicate zinc ion coordination, whereas dashed lines indicate hydrogen bond distances.

```
hCA II      4 --HWGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVS-YDQA--TSLRILNNGHAF
hCA I       5 ---WGYDDKNGPEQWSKLYPIANGNNQSPVDIKTSETKHDTSLKPISVS-YNPA--TAKEIINVGHSF
hCA IX      2 QSHWRYGGDP---PWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFQLPPLPELRLRNNGHSV
hCA XII     4 --KWTYFGPDGENSWSKKYPSCGGLLQSPIDLHSDILQYDASLTPLEFQGYNLSANKQFLLTNNGHSV
                * *                                                         * **

                                          ▼                   ▼         ▼
hCA II     67 NVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDG-QGSEHTVDKKKYAAELHLVHWN-TKYGDFGK
hCA I      67 HVNFEDNDNRSVLKGGPFSDSYRLFQFHFHWGSTNE-HGSEHTVDGVKYSAELHVAHWNSAKYSSLAE
hCA IX     67 QLTLPP---GLEMALG-PGREYRALQLHLHWGAAGR-PGSEHTVEGHRFPAEIHVVHLS-TAFARVDE
hCA XII    67 KLNLPS---DMHIQG--LQSRYSATQLHLHWGNPNDPHGSEHTVSGQHFAAELHIVHYNSDLYPDAST
                * *                   ** * *         *           * *         **

                                                                          ▼▼
hCA II    134 AVQQPDGLAVLGIFLKVG-SAKPGLQKVVDVLDSIKTKGKSADFTNFDPRGLLPESL-DYWTYPGSLT
hCA I     134 AASKADGLAVIGVLMKVG-EANPKLQKVLDALQAIKTKGKRAPFTNFDPSTLLPSSL-DFWTYPGSLT
hCA IX    134 ALGRPGGLAVLAAFLEEGPEENSAYEQLLSRLEEIAEEGSETQVPGLDISALLPSDFSRYFQYEGSLT
hCA XII   134 ASNKSEGLAVLAVLIEMG-SFNPSYDKIFSHLQHVKYKGQEAFVPGFNIEELLPERTAEYYRYRGSLT
                **    * *                                                   **

              ▼       ▼
hCA II    200 TPPLLECVTWIVLKEPISVSSEQVLKFRK-LNFNGEGE-PEELMVDNWRPAQPLKNRQIKASF-
hCA I     200 HPPLYESVTWIICKESISVSSEQLAQFRS-LLSNVEGD-NAVPMQHNNRPTQPLKGRTVRASF-
hCA IX    200 TPPCAQGVIWTVFNQTVMLSAKQLHTLSDTLWGPGD-----SRLQLNFRATQPLNGRVIEASFP
hCA XII   200 TPPCNPTVLWTVFRNPVQISQEQLLALETALYCTHMDDPSPREMINNFRQVQKFDERLVYTSFS
                *** *  * *
```

Fig. S5 Structure-based sequence alignment of hCA isoforms II (PDB code 1CA2)[21], I (PDB code 2CAB)[24], IX (PDB code 3IAI)[25], XII (PDB code 1JCZ)[26]. Residues delimiting the active site cavity are marked with asterisks. Residues interacting with inhibitor **1** in the hCA II/**1** complex are indicated by triangles, while the corresponding residues that in hCA I, hCA IX and hCA XII are different are highligthed in red.

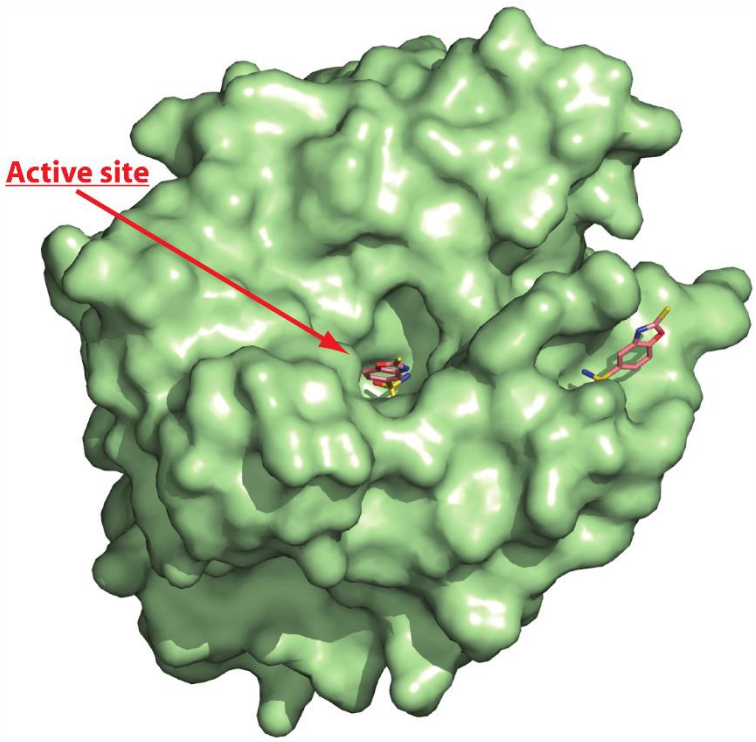


Fig. S6 Solvent accessible surface of hCA II in its complex with **2**. The two inhibitor molecules bound within the active site cavity and on the protein surface are shown in stick representation. The active site position is indicated by an arrow.

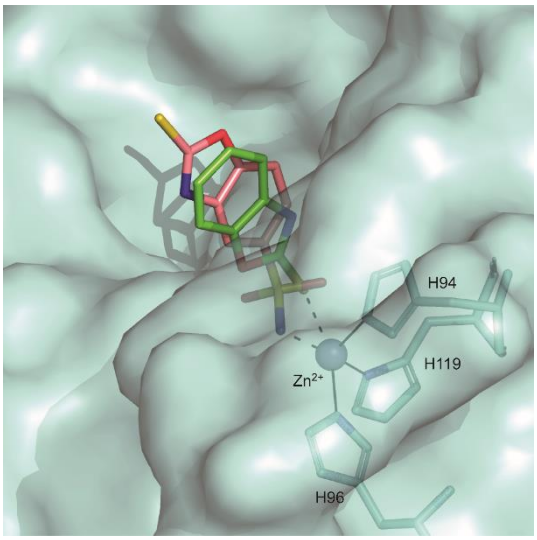


Fig. S7 Structural superposition between hCA II/**1** (green) and hCA II/**2** (pink). Continuous lines indicate His-Zn coordination, while dotted lines indicates Inhibitor-Zn coordination distances.

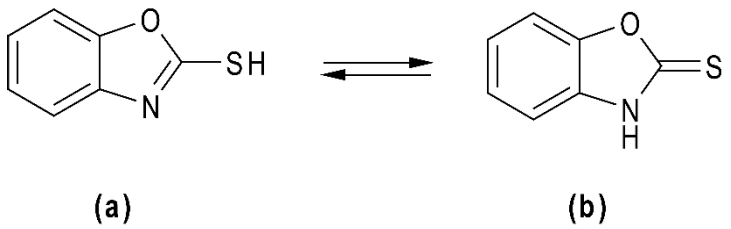


Fig. S8 Thiol (a) and thione (b) tautomeric forms of 2-mercaptobenzoxazole.

References

1. C. Ding, S. Wang, Y. Sheng, Q. Dai, Y. Zhao, G. Liang and Z. Song, *RSC Advances*, 2019, **9**, 5.
2. C. Mesangeau, S. Narayanan, A. M. Green, J. Shaikh, N. Kaushal, E. Viard, Y. T. Xu, J. A. Fishback, J. H. Poupaert, R. R. Matsumoto and C. R. McCurdy, *J. Med. Chem.*, 2008, **51**, 1482-1486.
3. R. G. Khalifah, *J. Biol. Chem.*, 1971, **246**, 2561-2573.
4. G. De Simone, A. Angeli, M. Bozdag, C. T. Supuran, J. Y. Winum, S. M. Monti and V. Alterio, *Chem. Commun.*, 2018, **54**, 10312-10315.
5. M. Bozdag, M. Ferraroni, C. Ward, F. Carta, S. Bua, A. Angeli, S. P. Langdon, I. H. Kunkler, A. S. Al-Tamimi and C. T. Supuran, *Eur. J. Med. Chem.*, 2019, **182**, 111600.
6. M. Abdoli, A. Angeli, M. Bozdag, F. Carta, A. Kakanejadifard, H. Saeidian and C. T. Supuran, *J. Enzyme Inhib. Med. Chem.*, 2017, **32**, 1071-1078.
7. M. Bozdag, M. Pinard, F. Carta, E. Masini, A. Scozzafava, R. McKenna and C. T. Supuran, *J. Med. Chem.*, 2014, **57**, 9673-9686.
8. F. Carta, D. Vullo, A. Maresca, A. Scozzafava and C. T. Supuran, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 2182-2185.
9. E. Andreucci, S. Peppicelli, F. Carta, G. Brisotto, E. Biscontin, J. Ruzzolini, F. Bianchini, A. Biagioni, C. T. Supuran and L. Calorini, *J. Mol. Med.*, 2017, **95**, 1341-1353.
10. Z. Otwinowski and W. Minor, *Methods Enzymol.*, 1997, **276**, 307-326.
11. V. Alterio, D. Esposito, S. M. Monti, C. T. Supuran and G. De Simone, *J. Enzyme Inhib. Med. Chem.*, 2018, **33**, 151-157.
12. A. T. Brunger, *Nat. Protoc.*, 2007, **2**, 2728-2733.
13. A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr. D Biol. Crystallogr.*, 1998, **54**, 905-921.
14. T. A. Jones, J. Y. Zou, S. W. Cowan and M. Kjeldgaard, *Acta Crystallogr. Sect. A: Found. Crystallogr.*, 1991, **47 (Pt 2)**, 110-119.
15. J. Susperregui, M. Bayle, J. M. Léger and G. Déléris, *J. Organomet. Chem.*, 1998, **556**, 6.
16. F. H. Allen, *Acta Crystallogr. B*, 2002, **58**, 380-388.
17. A. W. Schuttelkopf and D. M. van Aalten, *Acta Crystallogr. D Biol. Crystallogr.*, 2004, **60**, 1355-1363.
18. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283-291.
19. R. W. Hooft, G. Vriend, C. Sander and E. E. Abola, *Nature*, 1996, **381**, 272.
20. B. Sjoblom, M. Polentarutti and K. Djinovic-Carugo, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10609-10613.
21. A. E. Eriksson, T. A. Jones and A. Liljas, *Proteins*, 1988, **4**, 274-282.
22. J. M. Fox, K. Kang, W. Sherman, A. Heroux, G. M. Sastry, M. Baghbanzadeh, M. R. Lockett and G. M. Whitesides, *J. Am. Chem. Soc.*, 2015, **137**, 3859-3866.
23. A. E. Eriksson, P. M. Kylsten, T. A. Jones and A. Liljas, *Proteins*, 1988, **4**, 283-293.
24. K. K. Kannan, M. Ramanadham and T. A. Jones, *Ann. N. Y. Acad. Sci.*, 1984, **429**, 49-60.
25. V. Alterio, M. Hilvo, A. Di Fiore, C. T. Supuran, P. Pan, S. Parkkila, A. Scaloni, J. Pastorek, S. Pastorekova, C. Pedone, A. Scozzafava, S. M. Monti and G. De Simone, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 16233-16238.
26. D. A. Whittington, A. Waheed, B. Ulmasov, G. N. Shah, J. H. Grubb, W. S. Sly and D. W. Christianson, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 9545-9550.