Novel Potential Binding Sites for Selective Inhibitor Design of Human Soluble Epoxide Hydrolase

Maria Bzówka

Silesian University of Technology <https://orcid.org/0000-0001-6802-8753>

Karolina Mitusińska

Silesian University of Technology <https://orcid.org/0000-0002-0183-0845>

Artur Góra (✉ a.gora@tunnelinggroup.pl)

Silesian University of Technology <https://orcid.org/0000-0003-2530-6957>

Research article

Keywords: human soluble epoxide hydrolase, enzyme inhibitors, in silico, classical molecular dynamics simulations, mixed-solvent molecular dynamics simulations, ligand tracking approach

Posted Date: May 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-29814/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Human soluble epoxide hydrolase (hsEH) has been known to be involved in the hydrolysis of epoxyeicosatrienoic acids (EETs), which are anti-inflammatory and cardioprotective signalling molecules. Since the conversion of EETs into the corresponding diols generates non-bioactive molecules, the enzyme's inhibition would be beneficial. hsEH quickly became the molecular target, however, the proposed inhibitors directed toward this enzyme have not undergone clinical trials, mostly due to their low solubility in water.

Our goal is to indicate new regions, distinct from those already investigated, that could be considered during the development of novel and perhaps more soluble inhibitors of hsEH. To fulfil this goal, we used a combination of classical and mixed-solvent molecular dynamics (cMD and MixMD) simulations and a ligand tracking approach. Such a strategy, considers conformational changes in the interior of the enzyme, extends conformational space of binding cavities, and provides insight into different sets of potential binding modes. By analysis of the local distribution of different molecular probes (water and organic cosolvent molecules), we were able to identify potential binding sites (hot-spots), which can indicate the location of the most important functional groups for the binding of designed inhibitors. The uniqueness of the binding spots was examined by superpositioning the indicated hot-spots with a map of chemical interactions between the known inhibitors and hsEH. We analysed all available crystal structures of hsEHs with inhibitors deposited in the PDB database to gain insight into the binding modes of the known inhibitors and to find similarities and dissimilarities between them. By comparing the results obtained based on the map of chemical interactions, with the regions identified by using cMD and MixMD, we were able to identify regions that have not been targeted by ligands until now and which are promising possible binding sites for polar groups that would be beneficial for inhibitors solubility.

Introduction

The human soluble epoxide hydrolase (hsEH, EC 3.3.2.10) is a bifunctional homodimeric enzyme weighing 62 kDa. It is a member of the epoxide hydrolase superfamily and is composed of two independently folded domains: N-terminal domain (NTD) with phosphatase activity and C-terminal domain (CTD) with hydrolase activity. hsEH is encoded by the EPHX2 gene and can be found in both cytosol and peroxisomes[1–3]. The enzyme shows the highest activity in the liver; however, it is distributed in the majority of tissues, including the kidney, intestine, adipocytes, and vessels[4–6]. The clinical relevance of hsEH is linked with a C-terminus motif which performs the hydrolysis of arachidonic acid epoxides and other natural epoxy-fatty acids, especially the transformation of four regioisomers of epoxyeicosatrienoic acids (EETs): 5,6-, 8,9-, 11,12-, and 14,15-EETs to the corresponding dihydroxyeicosatrienoic acids (DHETs)[7, 8]. The conversion of EETs into the corresponding diols performed by hsEH generates non-bioactive molecules, thus the enzyme inhibition would be expected to enhance the EETs bioavailability and their beneficial properties[9, 10]. EETs have been found to modulate the activities of many molecular targets and signalling pathways. They have been shown to possess vasodilator and anti-inflammation properties, as well as act as mediators of autocrine and paracrine[9,

11, 12]. These characteristics are why many studies implicate their significant meaning in several cardiovascular disorders, e.g. blood pressure, hypertension, and atherosclerosis[13, 14]. EETs have also been found in the brain, where they regulate cerebral blood flow and protect against ischemic brain injury[15, 16]. Several studies were carried out to investigate the effect of the sEH gene deletion[17–19]. Samokhvalov *et al.* reported that such a deletion provides a protective effect against lipopolysaccharide (LPS)-induced cardiotoxicity by maintaining mitochondrial function[20]. In their research associated with Parkinson's disease, Ren *et al.* reported that sEH gene deletion or enzyme inhibition may protect against MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) induced neurotoxicity[21]. EETs from the cytochrome P450/soluble epoxide hydrolase pathway, together with the prostaglandins derived from the cyclooxygenase (COX) pathway, have also been found to be important eicosanoids that may regulate angiogenesis and tumorigenesis[22, 23]. Therefore, a great effort has been made to develop inhibition strategies towards sEH.

Some of the earliest sEH inhibitors were chalcone oxide, which can form a covalent intermediate with the enzyme. However, they were slowly hydrolysed and then released from the active site due to time-dependent inhibition and they were rather unstable, making *in vivo* usage limited[9, 24–26]. Urea, carbamate, and amide compounds with the hydrophobic group substitution were the first successful generation of hsEH inhibitors. They were identified as significantly more potent inhibitors and still represent the most popular class of hsEH inhibitors[26, 27]. These compounds can mimic epoxide binding and are not released from the active site. However, they have rather poor physical properties which include, low solubilities and high melting points, resulting in inappropriate bioavailability[26, 28, 29]. To improve their properties, while maintaining a high potency, a polar functional group at a proper distance from the urea or amide moiety was added. Soon after, the second generation of inhibitors, based on the 12-(3-adamantan-1-ylureido)dodecanoic acid (AUDA) and 1-adamantan-1-yl-3-{5-[2-(ethoxyethoxy)ethoxy]pentyl} urea (AEPU), were developed[9, 26, 28, 30–32]. But, these compounds remain flawed, due to their rapid *in vivo* metabolism and limited pharmacokinetics parameters.

Bruce Hammock's group proposed the so-called "Pharmacophore Model", which has been widely used to develop new sEH inhibitors[33]. The "Pharmacophore Model" is as follows: the "primary pharmacophore" (P1) may consist of carbamate, amide, or urea (preferably) and bear a bulky, hydrophobic substituent, whereas the "secondary pharmacophore" (P2) should consist of a polar group, located at a distance of five or six carbons (7.5 Å) from the urea carbonyl group. The "secondary pharmacophore" could be any group that is capable of serving as a proton acceptor (e.g., amide, ketone, alcohol, etc.); however, moieties that are too polar offered less potent inhibition. The model may also show an optional "tertiary pharmacophore" (P3), located at least 12 Å from the urea carbonyl group. The hydrophobic linkers, L1 and L2, should connect the polar moieties from the primary, secondary and tertiary pharmacophores[33–35].

Several pharmaceutical companies and academic institutions and pharmaceutical companies, such as the University of California-Davis, Columbia University, Chinese Academy of Sciences, Arête, Boehringer Ingelheim, Taisho, Takeda, Merck, Roche, etc., have entered the arena of developing sEH inhibitors.

Thousands of various hSEH inhibitors were synthesised over the last three decades, but only individual compounds have been qualified into clinical trials. Such an example is the 1-(1-acetyl-piperidin-4-yl)-3-adamantan-1-yl-urea (AR9281) inhibitor from Arête Therapeutics[36]. The compound has completed phase I of the clinical trial successfully[37] but has failed the IIa clinical trial due to lack of efficiency[38]. Esters and salts of AUDA, commonly used as an experimental sEH reference inhibitor, are good inhibitors of sEH but their clinical use has been restricted due to metabolic instability and limited solubility in water and many organic solvents[33, 34, 39]. Also, GSK2256294 ((1R,3S)-N-(4-cyano-2-(trifluoromethyl)benzyl)-3-((4-methyl-6-(methylamino)-1,3,5-triazin-2-yl)amino)cyclohexanecarboxamide) inhibitor has shown promising effects in two phase I studies, however, there is no information on the results of phase II clinical trials[40].

Various research groups began using the fragment-based drug discovery with biophysical or biochemical assays and a high-concentration of fragments to speed up the research of new inhibitors. For instance, Amano *et al.* conducted fragment screening against sEH using X-ray crystallography together with biochemical assays[41, 42]. They were able to identify numerous fragments that inhibit sEH and reveal their binding modes. Using the optimisation techniques, they determined structures of fragment-enzyme complexes and discovered a novel scaffold for sEH inhibitors. In 2015, Öster *et al.* also presented a fragment-based crystallisation approach, where a fragment-based lead generation was executed with a limited high-throughput screening against the sEH[43]. The obtained complexes extended the structural knowledge about possible dual conformations or multiple binding sites. In 2016, Xue *et al.* used a similar approach to completely map the active site pocket to gain a deeper knowledge of ligand interactions and to guide the rational structure-based drug design[44].

In the last decade, research groups also applied various approaches to discover new hSEH inhibitors based primarily on *in silico* studies. In 2011, Tanaka *et al.* identified almost 70 new sEH inhibitors using a docking-based virtual screening[45]. They analysed different physicochemical properties of the hit compounds, e.g., size and lipophilicity, to find a compound with high ligand efficiency. Xing *et al.* used a similar approach to construct combinatorial libraries of potential inhibitors based on a benzoxazole template[46]. The libraries were further evaluated by docking and by analysing the capacity to make hydrogen bonds. Research was then expanded into exploring the structure-activity relationship by the amide bond formation. In 2012, Moser *et al.* developed a pharmacophore model based on co-crystallised inhibitors of hSEH[47]. They used a combination of Protein Ligand Interaction Fingerprints with the QueryGenerator tool to model the shape of the binding pocket indicated as excluded volumes. Then, they proposed a new pharmacophore model (with a special focus on the catalytic triad, as described by Morisseau and Hammock[26]) by examination of the clusters of receptor pharmacophore annotation points within the binding pocket. This pharmacophore model was later used by Moser *et al.* to find dual 5-lipoxygenase (5-LO)/soluble epoxide hydrolase inhibitors[48]. In 2016, Waltenberger *et al.* created multiple structure- and ligand-based pharmacophore models to cover different putative binding modes of the ligands and to cover the active space more thoroughly[49]. In 2018, Tripathi *et al.* applied a 3D-QSAR approach using sEH inhibitors from literature (with a wide range of activity and structural diversity) to design a pharmacophore model[50]. A new pharmacophore hypothesis was presented by Bhagwati *et al.*

in recent research, based on the molecular modelling studies, which include ligand-based pharmacophore modelling, virtual screening, molecular docking, and all-atom molecular dynamics simulations[51].

The use of *in silico* methods has undoubtedly helped propose new hSEH inhibitors, however, the research was mainly focused on exploring the main binding pocket and developing new pharmacophore models based on the structures of known inhibitors binding in this area. Exploration of other parts of the hSEH enzyme, e.g., accessible after the conformation changes that could host additional functional groups, was omitted. For the exploration of new, potential binding sites in the protein, molecular dynamics simulation with (co)solvent molecules could be applied. Such an approach, called mixed-solvent molecular dynamics (MixMD) or cosolvent-based molecular dynamics (CMD) simulations, mixes small organic molecules with a water solvent to stimulate dynamic protein motions and induce partial conformational changes of binding pocket residues appropriate for the binding of diverse ligands[52–56]. Based on such abundant information, it is possible to identify the ligand binding modes, preferable binding regions (often called hot-spots), and present structural characteristics of a given place to build a pharmacophore model[57–59]. But, the MixMD has several issues. There is no universal procedure or methodology for performing this type of simulation. Various research groups select different types of cosolvent molecules and the concentration of cosolvent in the systems can vary significantly[52, 53, 67–69, 57, 60–66]. One problem is the selection of cosolvents to properly map all of the possible protein-ligand interactions and regions, such as hydrogen bonding (donors and acceptors), hydrophobic regions, and positive and negative ionic interactions[60, 70, 71]. Moreover, the water:cosolvent ratio could also affect the results, since a higher concentration of particular molecules could disturb the protein's structure and the cosolvent molecules could aggregate[54, 72–74]. MixMD simulations were usually used for relatively small proteins, such as thermolysin[52, 63, 75–77], p53[76, 78, 79], HIV-1 protease[80–82], hen egg-white lysozyme[83], Hsp90[81], and several kinases[72, 84] to identify potential binding regions on the protein's surface exclusively. Therefore, an analysis for proteins with buried active sites (as in the case of hSEH) remains a challenge. For a long time, no bioinformatics tools were available that would allow accurate tracking of the behaviour of solvent molecules inside the protein's core and its further analysis from an intramolecular voids the point of view during the molecular dynamics simulations. The situation changed when new and easy-to-use AQUA-DUCT (AQ) software appeared[85]. This tool was primarily used to track the molecular trajectories of ligands (solvent molecules) throughout molecular dynamics simulations. The new version, released in 2019, also allows the identification of structurally important residues and regions of macromolecules responsible e.g., for ligand trapping, which can vary depending on the tracked ligand[86]. Thus, such functionality can be used for *in silico* drug design and for identifying new potential binding sites for hSEH inhibitors.

In this paper, we have addressed two main questions: i) is it possible to expand the targeted space of the hSEH interior to propose unique binding sites for new classes of inhibitors? and ii) is it possible to identify amino acids that, so far, were not involved with inhibitors' interactions and could accept functional groups to increase the inhibitors solubility? To address these questions, we used the combination of small molecules tracking and classical and mixed-solvent molecular dynamics simulations to analyse the intramolecular voids of the hSEH. Such an approach allows us to analyse the local distribution of

particular solvent molecules employed as molecular probes. The densest regions of water and cosolvent pathways (hot-spots), are used to predict particular regions of protein-ligand interactions and could be used for further pharmacophore design. Since hsEH has a relatively high number of crystallographic protein-inhibitor complexes deposited at the Protein Data Bank (PDB) database[87], we could use this abundant information to gain structural insight into the binding of known inhibitors and compare them with the obtained results. We believe that such an approach can be applied for any molecular target and can work as a universal pipeline in the drug discovery pathway.

Results

Crystal structure of human soluble epoxide hydrolase and entrances to the protein's interior

In the hsEH hydrolase domain, we can distinguish two main parts, the core (also known as main) domain and cap (also known as the lid) domain (Fig. 1a). The main domain consists of the amino acids in the range M235-P369 and M469-M555. The cap domain consists of the amino acids in the range S370-R468 with a highly flexible cap-loop region (F409-T443). The so-called NC-loop (Y348-S370) connects the cap and main domains. Studies have revealed that the active site of hsEH is buried inside the protein core, in an “L”-shaped pocket. The “L”-shaped pocket extends from the bottom of the main domain up to the cap domain. The lengths of the two branches, which make up the internal pocket, are 15 Å (long branch) and 10 Å (short branch)[41]. The entire pocket is basically hydrophobic, even though each branch of the pocket accepts diverse functional groups. The branches are connected by a small bottleneck in which the catalytic triad and two stabilising residues are located (D335, D496, H524, Y383, and Y466 respectively).

The buried active site of hsEH is connected with the environment through tunnels, which act as exchange pathways between a bulk solvent and the active site. The presence of two branches of the “L”-shaped pocket may suggest that the protein only has two entrances to its interior. Using water as a molecular probe and tracking approach implemented in AQUA-DUCT software, we could identify four entrance/exit locations for water molecules penetrating the hsEH structure (Fig. 1b). We named them: Tm1 – used by the vast majority of the identified water molecules (~ 78%), it is permanently open, located in a long branch between two helices in the main domain (α 4: S370-E389 and α 10: V500-I511), Tc/m – used by ~ 20% of water molecules, located at the border of the cap and main domains, in a short branch (between E520-H524 loop of the main domain and the A411-S415 loop of the cap domain), Tg – transient (< 1%) gorge linking Tc/m with Tm1 (regulated by E494-L499 loop), and Tm2 – rarely used (< 1%), separated by two loops, located deep in the main domain, adjacent to the Tm1 tunnel (N359-M369 part of the NC-loop and S479-P488 loop). The nomenclature was adopted according to the publication on *Solanum tuberosum* soluble epoxide hydrolase[88]. Secondary structures motifs of hsEH are presented in **Supplementary Information, Figure S1 and Table S1**. Tracked water molecules depict the internal cavity of hsEH and provide information about maximal accessible volume, which was visited during MD simulations (Fig. 1c).

Additional entrance/exit areas

In the MixMD simulations with water and acetonitrile molecules, as well as in the simulation with water and urea molecules, an additional entrance, Tcap1 located in the cap domain, between the K421-G426 α 6 helix and G427-S432 part of the cap-loop appeared. In the MixMD simulation with water and methanol molecules, the additional entrance between the α 10 helix and β 7 sheet (A488-A493) of the core domain, named Tm3, has been detected (**Supplementary Information, Figure S2**).

Binding residue clustering based on the known crystal structures

The effort to understand the principles of hSEH function and selectivity has resulted in 104 crystal structures deposited in the Protein Data Bank, of which 100 structures are complexes containing ligands relevant to hSEH inhibition. We revisited these structures to map the interactions between the protein and inhibitors, and to summarise targeted parts of the hSEH structures. In some cases, inhibitors had been crystallised at various positions or conformations, which resulted in an analysis of 123 protein-ligand complexes. We analysed the structures of complexes using the LigPlot program and identified amino acids involved in forming hydrogen bonds or non-bonded interactions (**Supplementary Information, Table S2**).

43 amino acids were identified as interacting with ligands in various crystal structures. The amino acids were clustered based on the interaction-pattern similarity between residues and inhibitors. This approach revealed five principal clusters: C-1a, C-1b, C-2, C-3, C-4, and a set of outliers (Fig. 2). C-1a and C-1b clusters cover the long branch of the “L”-shaped pocket and part of the bottleneck, where the active site is located. In total, they consist of 22 amino acids. C-2, C-3, and C-4 clusters surround the short branch of the “L”-shaped pocket and cover the upper part of the bottleneck. In total, 13 amino acids build these clusters. Cluster C-1a (light pink), formed by eleven amino acids, is built by residues covering one of the walls of the long branch from the “L”-shaped pocket. Five out of 11 amino acids formed this cluster (A365, P371, S374, N472, A476) and were identified by NetSurfP server[89] as surface-exposed residues. One of the amino acids belonging to the C-1a cluster was found to be located on the NC-loop. Cluster C-1b (deep pink), also formed by eleven amino acids, covers the opposite wall of the long branch and part of the bottleneck with the active site of hSEH. Within this cluster, two residues from the catalytic triad, D335 and H524, as well as Y383, an amino acid from a pair of tyrosines that stabilise the transition state intermediates during catalysis[90], were identified. Residues from the C1-b cluster were found to interact with the majority of analysed inhibitors. None of the amino acids that make up the C-1b cluster were identified as surface-exposed residue. In this cluster, as in the case of the C-1a cluster, one amino acid was found on the NC-loop, A365. Amino acids from C-1a together with the residues from the C-1b cluster surround the main entrance to the protein's interior – Tm1 tunnel identified by AQUA-DUCT software. Cluster C-2 (cyan), formed by six amino acids, is located between the cap and main domains, close to the active site of the hSEH. Amino acids that build this cluster cover part of the bottleneck linking two branches of the “L”-shaped pocket, as well as the vast majority of the short branch. Within this cluster,

there is one residue from the catalytic triad of hsEH – D496. Even though this residue belongs to the active site, it interacts with known inhibitors less often than in the case of active site residues from the C-1b cluster. It is worth mentioning that F497 residue, which is also a member of the C-2 cluster, potentially works as a gate that separates the Tc/m and Tg tunnels. As described in works by Góra *et al.*, and Marques *et al.*, enzyme gates have multiple molecular functions[91, 92]. One of their main function is to control the access of substances to the enzyme's interior. They can also contribute to enzyme selectivity, prevent solvent access to specific regions of the protein, synchronise processes occurring in distant parts of the protein works. Due to the F497 mobility, it may open the additional access to the interior of the protein through the Tg tunnel and possibly regulate access to the active site of hsEH for more bulky substrates, or can modulate solvent molecules entry. Moreover, F497 was identified as the only surface-exposed residue within the C-2 cluster. Cluster C-3 (green), formed by four amino acids from the cap domain, especially from the cap-loop (R410, A411, and S415), is located exactly above the entrance to the Tc/m tunnel. It covers the part of the short branch of the "L"-shaped pocket. All amino acids from the C-3 cluster (mentioned above and S407) were characterised as surface-exposed residues. Lastly, from the main clusters, cluster C-4 (yellow), formed by three amino acids, is also located in the cap domain, closer to the hinge region which links the cap and main domains and contributes to the structure flexibility. The C-4 cluster also covers the short branch as did the C-2 and C-3 clusters. Within the amino acids that build this cluster, F387 is anchored exactly at the hinge region and its aromatic ring is faced toward the active site. In addition, one of the amino acids from the C-4 cluster, L417, was characterised as a surface-exposed residue.

Almost all remaining amino acids (besides T360) are specific for interactions with individual inhibitors and therefore were clustered as outliers. These residues are located: i) on the NC-loop, near the entrance to the rarely used Tm2 tunnel (N366) and facing one of the walls from the long branch of the "L"-shaped pocket (T360, P361); ii) on the loop between the main and cap domains, near the active site (P268); iii) close to the hinge region (N378); iv) in the cap domain (S412, V416, and L397 – of which the first two amino acids are part of the cap-loop). All amino acids classified as outliers are buried residues.

Inhibitor clustering based on interaction pattern similarity

Clustering of inhibitors was based on the interaction matrix (**Supplementary Information, Table S2**) and indicated four major binding locations in the hsEH protein interior. We named them: C-I, C-II, C-III, and C-IV. C-I is the lower part of the long branch of the "L"-shaped pocket, where nine inhibitors bind and do not reach the active site. C-II is the upper part of the long branch of the "L"-shaped pocket, close to the active site and consists of six inhibitors. C-III is the whole "L"-shaped pocket, spreading from the main domain to the cap domain (inside this area we can distinguish inhibitors that are expanded towards the short branch cap domain, as well as those expanded towards the long branch) – 72 inhibitors. C-IV is the short branch of the "L"-shaped pocket, slightly covering the upper part of the bottleneck, where the active site is located – 28 inhibitors. Inhibitors characterised as outliers bind deep in the main domain in the long branch or at the border of cap and main domains located in the short branch.

Table 1

Division of residues among clusters. Red colour indicates residues from the active site, residues marked bold interact with more than 50% of inhibitors.

CLUSTER ID	I	II	III	IV	OUT
1a	M310, Y343, A365, P371, I375, F381, M469, N472, W473, A476	P371, S374, I375, F381, M469	Y343, S374, I375, F381, M469, N472		Y343, A365, P371, S374, I375, F381, M469, N472, W473
1b	W336, M339, I363, Q384, Y466,	D335, W336, M339, I363, Y383, Q384, Y466, L499, M503	F267, D335, W336, M339, I363, Y383, Q384, Y466, L499, M503, H524	F267, D335, Y383, Y466, L499, H524	F267, W336, M339, I363, Y383, Q384, M503
2			L408, M419, D496, F497, V498, W525	L408, M419, D496, F497, V498, W525	L408, L428
3			R410, A411, S415	L397, R410, S415	
4			F387, L417, L428	F387, L417	F387
OUT		P361	P268, T360, P361, S407	S407, S412, V416	P361, N366, N378

Inhibitors grouped in the C-I cluster interact with 15 amino acids (several amino acids from the C-1a cluster and almost all from the C-1b cluster). Y343 and N472 are two amino acids that interact with all inhibitors within the C-I cluster. Four of the identified amino acids also form hydrogen bonds with inhibitors from the C-I cluster. They are Q384, Y466, M469, and N472. Six inhibitors are part of the C-II cluster. They are located in the upper part of the long branch and they are slightly covering the lower part of the bottleneck, where the active site is located. 15 amino acids (both from the C-1a and C-1b clusters, as well as one amino acid from the outliers) are involved in binding to those inhibitors. Although inhibitors occupy only a small part of the active site bottleneck, they interact with one amino acid from the catalytic triad, D335, as well as with Y383 and Y466, which stabilise the transition state intermediates. I375 and Q384 are involved in interactions with all inhibitors from the C-II cluster. Within this group, only two inhibitors may form hydrogen bonds 3-[4-(benzyloxy)phenyl]propan-1-ol (49Q) with Y383 and Y466, and 3-methyl-4-phenyl-1H-pyrazole (GVG – one of its crystallised position) with D335. The majority of the inhibitors (72 structures) occupy almost the whole “L”-shaped pocket (cluster C-III). They are usually inhibitors with long chains. These inhibitors interact with a huge number of amino acids

from all identified clusters. 33 amino acids are capable of binding to at least one of the inhibitors. Among the amino acids that build the catalytic triad, inhibitors belonging to this cluster interact most often with D335 and H524. As for D496, only a few inhibitors from the C-III cluster interact with this residue. According to the list of the identified interactions, interactions with Y383 and Y466 are of great importance. Of the 33 identified amino acids, nine can form hydrogen bonds with the examined inhibitors. Of these nine amino acids, the most targeted are D335, from the catalytic triad, as well as Y383 and Y466, which stabilise the transition state. F267, Q374, S415, D496, F497, and H524 are also involved in the formation of single hydrogen bonds. Only three out of 72 inhibitors from the C-III cluster do not form any hydrogen bonds. 28 inhibitors bind to the short branch of the “L”-shaped pocket and they are slightly covering the upper part of the bottleneck with the active site (cluster C-IV). They interact with amino acids from C-1a (amino acids forming the upper part of this cluster), C-2 (all amino acids), C-3 (all amino acids, except A411), and C-4 (all amino acids). Since H524 and D496 are two residues from the catalytic triad, the high number of interactions is not surprising. D496 together with F497 are capable of forming hydrogen bonds with nearly half of the inhibitors belonging to the C-IV cluster. In most cases, both amino acids are simultaneously involved in forming hydrogen bonds. Moreover, 2-[(5-chloro-2-pyridyl)sulfanyl]ethanol (YPN) inhibitor from the 5ak6 crystal structure may form three hydrogen bonds with L408, R410, and Y525. Whereas 5-(3-ethylsulfanylphenyl)thiophene-2-carboxamide (DUL) inhibitor from 5alx crystal structure may form two hydrogen bonds with V416 and L416. Out of 28 inhibitors, eight from the C-IV cluster do not indicate the ability to form hydrogen bonds.

Hot-spot identification

The analysis of water and cosolvent molecule trajectories combined with the local distribution approach made it possible to identify areas of increased density of (co)solvent molecules during the MD simulations (Fig. 4). Such areas (hot-spots) are potential binding/interacting sites and may provide additional information about regions attracting particular types of molecules (functional groups). Using AQUA-DUCT software, the overall distribution of tracked water and cosolvent molecules (so-called pockets) were also defined. The pockets represent the actual regions in the protein interior that are explored by (co)solvent molecules.

Hot-spots for water molecules were found mostly in the long branch of the “L”-shaped pocket, as well as in the bottleneck where the active site is located. No water hot-spots were found in the short branch of the “L”-shaped pocket. Hot-spots with the highest density are located in proximity of the active site and two more water hot-spots, which also show high density, can be found outside the “L”-shaped pocket. One is located between the cap and main domains, whereas the other is located in the main domain. For acetonitrile molecules, hot-spots were found both in the long and short branches of the “L”-shaped internal cavity. However, acetonitrile hot-spots were not found in the bottleneck between the branches. The distribution of hot-spots in both branches is similar, both in terms of their quantity and density. In the case of DMSO molecules, few hot-spots were found. One hot-spot (with the highest density) was found in the long branch internal pocket, and the others were located in the bottleneck connecting the branches and in the vent to the short branch. Hot-spots for methanol were found in all key parts of the internal

pocket, both in the long and short branches, as well as in the bottleneck between them. However, when it comes to the long branch, they were found mostly near the amino acids of the C-1b cluster and do not cover the wall of the long branch built by amino acids from the C-1a cluster. For phenol molecules, as in the case of DMSO, few hot-spots were identified and all of them were located in the long branch of the “L”-shaped cavity. What is more, the density of these hot-spots is rather low. This is because very few phenol molecules were added to the simulation’s system. Almost all hot-spots for urea molecules were found in the long branch of the internal pocket, while a single hot-spot was found in the short branch. No hot-spots were identified in the bottleneck connecting both branches.

The areas of overall distribution of tracked water and cosolvent molecules vary depending on the (co)solvents. (Co)solvents characterised by high polarity easily penetrate the protein’s interior, while low-polarity solvents occupy only small regions of the protein (in the vicinity of the active site). In the case of solvents with relatively high polarity, their pockets also extend beyond the area of the main “L”-shaped binding cavity.

To obtain a global view of the occurrence of hot-spots from different (co)solvents, all were imposed on one structure (Fig. 5). In this way, a global “hot-spots map” was obtained indicating potential binding sites. Some of the hot-spots from different (co)solvents overlap mainly in the core domain in the long branch, but some occur individually in selected regions of the protein for a given (co)solvent. Interestingly, hot-spots representing different physicochemical properties (such as phenol and urea) may overlap in the same region of the protein.

Based on the hot-spots localisation, six different regions could be identified. Three regions of identified hot-spots overlap, with regions in which the known inhibitors of hsEH bind. These regions are located: i) at the border of the cap and core domains – corresponding with the short branch of “L”-shaped internal cavity, ii) in the active site, iii) in the core domain – corresponding with the long branch of the “L”-shaped pocket. Three remaining regions are unique and were not identified as known inhibitors of hsEH. The first one, P1, is located in the cap domain and corresponds to the location of the Tcap1 entrance/exit area (Fig. 5c). The second pocket predicted as a potential binding site, P2, is located between the cap and main domains and can host elongated inhibitors from the C-IV cluster (Fig. 5d). The last pocket, P3, is located at the Tm2 tunnel entrance (Fig. 5e). The detected hot-spots provide the evidence that inhibitors designed to block the Tm1 entrance can be elongated.

Potential interactions with amino acids

Pockets shown as potential binding sites were further analysed for identifying nearby amino acids with which inhibitors could potentially bind. The identification of amino acids was based on the search for amino acid residues located at a given distance (3.9 Å for water hot-spots and 3.35 Å for remaining hot-spots) from the geometric centre of the hot-spots during the molecular dynamics simulations. As a result of the search, a list of amino acids was obtained, along with information on how long the distance criterion was met during the molecular dynamics simulations. In total, 66 amino acids were identified, which met the given criterion during at least 1% of the simulation time. Interestingly, more than half of the

identified amino acids were found not for water hot-spots, but for hot-spots from other cosolvents. This indicates that the use of additional cosolvents, treated as different molecular probes, can provide greater information about the target. Among all identified amino acids, 27 can be characterised as unique and they did not appear in previous research. These amino acids were mainly found in regions that were not previously targeted by known inhibitors and are as follows: G266, Q269, H334, Y342, L358, N359, F362, P364, V380, T404, S418, H420, V422, A425, G427, F429, V430, M441, F459, K495, S504, T510, C522, G523, T526, G527, and D529. The majority of the unique amino acids were identified for a particular (co)solvent and only a few were identified for more than one (co)solvent, e.g., Y342 (acetonitrile, methanol, and water), F362 (acetonitrile, methanol, urea, and water), R410, S412 (acetonitrile and methanol), and A411 (acetonitrile, methanol, and water). The small cavity detected by acetonitrile molecules, described above as P1, is mostly surrounded by short and aliphatic residues, L417, V422, A425, G427, L428, and S418 (Fig. 5c). Among them, only leucines were classified as interacting with known inhibitors. The second unique region, P2, identified by a large-density hot-spot for water molecules, is surrounded by M441, C522, G523, and D529 and could provide polar groups for potential inhibitors. The W525 residue located between the short branch of the “L”-shaped pocket and detected cavity can provide additional constraints during the inhibitors design (Fig. 5d). The last unique region, P3, was detected by both water and acetonitrile hot-spots. Their hot-spots are surrounded by M339, W342, Y343, F362, and I363 (Fig. 5e).

Discussion

A combination of experimental and computational approaches has accelerated the discovery of numerous compounds inhibiting the human soluble epoxide hydrolase. However, none of the identified compounds which have qualified for clinical trials have passed the required tests allowing it to be marked as a drug. One of the biggest problems is low solubility of the proposed inhibitors that comes from the predominantly hydrophobic characteristics of the binding site of hSEH[41, 94, 95]. To shed light on new possibilities in drug design, we performed a comprehensive review of all deposited structures of hSEH with inhibitors and crash these analyses with the molecular probe tracking results in classical and MixMD simulations.

Based on studies by other researchers a relatively large number of crystal structures are available. In 2015, Amano *et al.* used X-ray crystallographic fragment screening to identify fragment hits and describe their binding modes[42]. They identified eight fragment hits and determined their co-crystal structures. They presented N-ethylmethylamine as a promising scaffold for inhibitors of soluble epoxide hydrolase and they selected compounds containing this scaffold for biochemical assays. Although the starting fragment had a weak inhibitory activity (IC₅₀: 800 uM), they were able to identify compounds with more than a 1500-fold increase in the inhibitory activity of IC₅₀: 0.51 uM. They characterised the binding site for each fragment in their work[42]. In the same year, Öster *et al.* also presented a fragment-based crystallisation approach, where a fragment-based lead generation was executed with a limited high-throughput screening against the sEH[43]. They obtained 55 structure-ligand complexes, with a wide range of size and affinity. Superimposition of all structures indicated that the entire binding pocket is

covered by ligands. They chose to deposit 52 structures within the hsEH dataset, including those containing some ambiguities, such as dual conformations or multiple binding sites, to provide the additional structural information[43]. In 2016, Xue *et al.* also used a similar approach to completely map the active-site pocket to gain a deeper knowledge of ligand interactions and to guide the rational structure-based drug design[44]. They obtained more than 50 structures of ligand complexes and through the analysis, they characterised preferred binding sites with some consensus hot-spots, in addition to the central channel. They identified two scaffolds: oxoindoline and 2-phenylbenzidazole-4-sulfonamide as a new series for potential hsEH inhibitors[44].

Most of the scientific research conducted so far for hsEH inhibitors design highlights the importance of the side chain residues composing the active site, D335 from the catalytic triad, as well as the stabilizing residues Y383 and Y466. The hydrogen-bonding network between the known inhibitors and these residues is a highly conserved feature across the reported hsEH inhibitors. This was demonstrated in numerous research studies. For instance, Gomez *et al.* presented that the binding of N-cyclohexyl-N'-(iodophenyl)urea (CIU) inhibitor is stabilised by a hydrogen bond interaction between the urea oxygen and T383 and Y466, as well as between the urea NH groups and D335[96]. In other work, Gomez *et al.* again stressed the importance of the hydrogen bond interaction of the above-mentioned residues with 4-(3-cyclohexylureido)-butyric acid (CU4) and 4-(3-cyclohexylureido)-hexanoic acid (CU6)[94]. Furthermore, X-ray crystallography research conducted by Tanaka *et al.* revealed the binding mode of inhibitors in which the amide group creates a hydrogen bonding network with the sEH active site residues[45]. Other studies have also been conducted using fragment-based drug discovery or *in silico* methods and have also pointed out the importance of D335, T383, and Y466[44, 97, 98]. Our studies confirm these findings, since the largest number of interactions (both non-bonded and hydrogen bonds) were found for these amino acids. Other important places, e.g. the hydrophobic region described in various studies and defined by F267, Y383, L408, M419, V498, and W525 (also known as F267 pocket), as well as the hydrophobic surface defined by W336, M339, and L499 (also known as W336 niche) are also reflected in our research by a high number of identified interactions between these residues and known inhibitors (**Supplementary Information, Table S2**).

However, the interaction with active site residues is not imperative. All deposited crystal structures underline the location of the binding site inside the protein's interior, a common feature in enzymes. According to the study by Pravda *et al.*, over 64% of enzymes contain the buried active site[99]. Although such a location adds additional constraints for an inhibitors design, it may also provide new opportunities. As described in the latest review by Marques *et al.*, inhibitors can target not only the active site itself but also the tunnels providing access to it[92]. Indeed, our analysis of the deposited structures of hsEH with bound inhibitors indicates such compounds. Inhibitors belonging to the C-III cluster may be described as the ones that fill the internal pockets of the proteins, but as can be seen, their ends are placed within the Tc/m and Tm1 entry/egress areas. Inhibitors from C-I, C-II, and C-IV clusters are located exactly within the entry/egress areas with most near the Tm1 and Tg tunnels. In turn, inhibitors that have not been classified into any major cluster are located deeper in the internal pocket.

One limitation of all research conducted so far is the fact that they focus only on the main pocket in which the natural substrate binds, without the exploration of the potential conformational changes of the protein's interior. This omits the possibility of designing compounds that would deeper penetrate the enzyme's interior and would bind to the unexplored regions of hSEH. In a case of selective inhibitors design for nitric oxide synthase, even rotation of the side chain of a single amino acid can provide access to the additional pocket and provide a new anchor stabilising the ligand binding[100, 101]. Therefore, we aimed to identify new pockets that extend the interior of the well-known "L"-shaped binding cavity of hSEH. The identification of new, unique regions, where the binding of new compounds may potentially occur, was achieved by applying the combination of tracking small molecules used as specific molecular probes (water molecules during cMD and cosolvent molecules during MixMD simulations) together with the local distribution approach. Such a combination provides information about an overall distribution of solvent in the protein's interior (pockets) as well as identifies areas with significant density (hot-spots) of selected molecules, which can be considered as regions of special interest. The uniqueness of this approach lies in the fact that the identification of such sites can take place in the protein's interior and not only on its surface, as done in previous studies[54, 57, 102–105]. This is particularly important for enzymes that have a buried active site. The detected additional pockets filled with particular (co)solvent hot-spots may provide an indicator of cavities which can host additional (polar) functional groups of potential inhibitors and could increase the solubility of the newly discovered compounds, thus overcoming the limitations that make inhibitors fail the clinical trials.

The combination of various chemical probes also addresses one more important issue of drug design. Mostly due to their low solubility in water and organic solvents, the chemical properties of amino acids affect the nature of the internal pocket (hydrophobic), leading to problems. The solution of such a problem could be the insertion of some polar fragments into the existing inhibitors that would improve solubility. The analysis of known inhibitors indicates that they do not fully occupy the available internal pocket (this is the case for both branches of "L"-shaped pocket). This suggests that there is an additional potency, which may be gained by filling these places with additional substituents. Admittedly, in their work, Öster *et al.* as well as Xue *et al.* claim that the fragment binders sample the entire binding pocket[43, 44]. However, our analysis shows that there is still some unused space at the ends of both short and long branches of the "L"-shaped pocket. Nevertheless, their research has greatly contributed to understanding the internal binding pocket of hSEH and resulted in obtaining numerous structure-ligand complexes, with a wide range of size and affinity. It is also worth noticing that Amano *et al.* presented an interesting structural insight into the binding of inhibitors by fragment screening and X-ray crystallography[41, 42]. They identified 126 hSEH-fragment complexes, among which almost 50 do not strictly interact with the amino acids from the active site. Instead, they extend their occurrence to long and short branches, which further indicates that it is reasonable to design novel inhibitors outside the active site. These fragments could be expanded toward further protein exploration.

The computational tools were dedicated to tackle the issues e.g., by application of the Lipinski's rule of five. In recent years, several *in silico* approaches have been applied to discover new hSEH inhibitors. In 2011, Tanaka *et al.* identified 68 new inhibitors (out of a 735 compound library) using a docking-based

virtual screening[45]. They analysed different physicochemical properties of the hit compounds, e.g., size and lipophilicity, to find a compound with high ligand efficiency. Then, the lead compound was optimised to yield a more potent compound, characterised by good ADMET profile. Xing *et al.* used a similar approach to construct combinatorial libraries based on a benzoxazole template[46]. The library was evaluated by docking and by analysing the capacity to make hydrogen bonds. The first library focused on the coupling of 591 acids with the 5-chlorobenzoxazoleamine. The results showed that 343 out of 383 synthesised compounds were active in a biological assay. IC₅₀ determination confirmed 23 compounds as single-digit nanomolar hSEH inhibition. The second library focused on exploring the structure-activity relationship by the amide bond formation between a subset of the acids and eight (aza)benzoxazole amines. 322 out of 840 compounds were active against hSEH[46]. The large study by Moser *et al.* involved a massive virtual screening of the “merged libraries” database (Asinex, > 430,000 molecules) [47]. Using this database, they performed a pharmacophore search. In the first step of analysis 3,191 virtual hits were found. Hits were further limited to 89 by applying the modified Lipinski’s rule of five and Oprea’s test for lead-likeness. Within this set, nine compounds were tested experimentally, of which two of the most potent hits showed an inhibition of over 50% at a single concentration of 30 μ M. The pharmacophore model was later used by Moser *et al.* to find dual 5-lipoxygenase (5-LO)/soluble epoxide hydrolase inhibitors[48]. In 2016, Waltenberger *et al.* created multiple structure- and ligand-based pharmacophore models to cover different putative binding modes of the ligands and to cover the active space more thoroughly[49]. They used nine crystal structures and 68 actives from the literature to model the structure-based pharmacophore using LigandScout (LS). Based on the PDB entries, 22 structure-based pharmacophores were generated and the resulting model set was able to recognise 55 out of 68 actives. Because part of the active compounds were not covered, additional three ligand-based pharmacophore models were generated using LS. 25 combined pharmacophore models were able to find 58 out of 86 active compounds. The remaining structures (containing aliphatic chains and ring systems) were not recognised as hydrophobic by LS version 3.03. For this reason, Discovery Studio software was used to create an additional three structure-based and two ligand-based models. Finally, 30 pharmacophore models were evaluated and 67 active compounds were found to interact with them. Further elimination of models with low enrichment factor (EF) qualities was performed. In the end, eight models (five structure-based and three ligand-based) were able to cover 65 active molecules. Models were used to perform the prospective virtual screening of Specs database. Six compounds per model (48 total) were experimental and 19 new sEH active molecules, with a total hit rate of 40% were found. Six out of 19 compounds displayed IC₅₀ values in the low nanomolar range. In 2018, Tripathi *et al.* applied a 3D-QSAR approach using 26 sEH inhibitors from literature (with a wide range of activity and structural diversity) to design a pharmacophore model containing four features: one hydrogen bond acceptor, two hydrophobic, and one ring aromatic[50]. The selected model was validated using both internal and external test set prediction, Fischer’s randomization, Rm2 metrics, and Güner-Henry scoring method. The chosen model was used to search the National Cancer Institute (NCI) and Maybridge datasets. The hits were screened based on estimated activity and fit value, along with Lipinski’s rule of five. Three hits with very high potency and fit value were chosen for sEH enzyme-based assay and rat aortic ring based vasodilation assay. They were also docked into the sEH. According to the most recent research, Bhagwati *et al.*

presented a new pharmacophore hypothesis based on the molecular modelling studies, including ligand-based pharmacophore modelling, virtual screening, molecular docking, and all-atom molecular dynamics simulations[51]. For the pharmacophore model generation, they chose the top three inhibitors from the available crystal structures based on their IC₅₀ values and interaction with the active site residues, D335 from the catalytic triad, along with Y383 and Y466. The ligand-based approach was used to produce three pharmacophore models using phase module. Then, the best alignment and common feature methods were used to create the pharmacophoric sites. For pharmacophore validation, the test set and decoy set method (with 28 confirmed actives and 1,232 decoys or inactives) were used. Validated pharmacophore was then used to screen the Maybridge database (consisting of around 53,000 small drug-like compounds) to obtain the hit compounds through the phase module. 1,802 hits were further subjected to molecular docking using an extra precision mode of Glide module. The top 100 hits from docking were visually analysed for their molecular interactions with the binding site. After that, nine hits were selected for *in vitro* inhibitor screening assay. Five out of nine compounds showed more than 50% inhibition at 25 μ M concentration. Within this group, three compounds indicated inhibition greater than 60% and were selected for secondary screening at 5 μ M concentration. Two out of three compounds showed inhibition of more than 50% at 5 μ M concentration. To explore the mechanism of inhibition, as well as to analyse the binding modes, a docking procedure was applied. To get even better insight into the validity of docking and biological assay results, the molecular dynamics simulations of the three most active protein-ligand complexes were performed to examine their complexes and to compare them with the reference co-crystallised inhibitor. However, none of the above mentioned methods and trials have considered significant conformational space sampling or alternative binding regions.

In the proposed approach, we focused on searching for regions (pockets identified by cosolvent hot-spots) that due to native or inhibitors induced dynamics of the protein could host inhibitors or their functional groups. Such information was extracted from classical and mixed-solvent MD simulations and we were able to detect potential sites for placing the new functional groups. For instance, different hot-spots have been found at the ends of both branches of the “L”-shaped pocket. After applying the known inhibitors’ structures with identified hot-spots, it can be seen that only a few compounds reach these areas. A particularly interesting idea seems to be the expanding of inhibitors belonging to the C-IV cluster so that their side chains would be directed more towards the cap-domain (within the identified Tcap1 tunnel), where the acetonitrile hot-spot was found. The P1 pocket is surrounded by six amino acids that could be involved in new inhibitor binding – L417, S418, V422, A425, G427, and L428. The mechanism of the cap domain opening could then be blocked.

Another option to consider is to pull out the inhibitors from the binding pockets towards the P2 pocket and position them within the side entry/egress area instead – next to the Tm3 tunnel. Such inhibitors could perhaps fill the gorge between the cap and main domains of hSEH. Within the P2 pocket we identified five amino acids that could be engaged into the inhibitors’ binding – M441, C522, G523, W525, and D529. Three of them provide functional groups that can address strong polar or even covalent interactions with designed inhibitors. The inhibitors classified in the C-IV cluster could probably be used as starting points of synthesis with an aliphatic linker crossing near the W525 residue and polar

functional group targeting P2 pocket. The third option would utilise the small inhibitors clustered into the C-I cluster as a starting point for designing new compounds. The detected P3 pocket provides a space that could host additional functional groups to increase solubility and selectivity of designed inhibitors.

Conclusions

Identifying hot-spots for different (co)solvents showed the importance of protein exploration using different molecular probes. By carrying out the analysis using several molecular probes with different chemical properties, it was possible to identify specific (possible) sites of interactions. Some of them were occupied by more than one type of (co)solvent, whereas some were more specific. This may mean that there are some regions in the protein structure, where the nature of the introduced groups can be various, while other places have a more specific and defined chemical character.

Methods

Molecular dynamic simulations

System setup

The crystal structure of human soluble epoxide hydrolase (PDB ID: 1s8o) was downloaded from the PDB. Using PyMOL[106], the phosphatase domain and the co-crystallised ligand hexaethylene glycol (P6G) were manually removed from the crystal structure.

Classical molecular dynamics simulation

The H++ server[107] was used to protonate the enzyme's structure using standard parameters and pH 6.5. Water molecules were placed using the combination of the Three-Dimensional Reference Interaction Site Model (3D-RISM) implemented in the AMBER package[108] and the Placevent algorithm[109]. The AMBER 14 LEaP[110] was used to immerse a model in a truncated octahedral box of TIP3P water molecules (no counterions were added) and prepare the system for simulation using the ff14SB force field. AMBER 14 software[110] was used to run 50 ns simulation of human soluble epoxide hydrolase. The minimisation procedure consisted of 2000 steps, involving 1000 of the steepest descent steps followed by 1000 steps of conjugate gradient energy minimisation, with decreasing constraints on the protein backbone (500, 125 and 25 kcal x mol⁻¹ x Å⁻²) and a final minimisation with no constraints of conjugate gradient energy minimization. Next, gradual heating was performed from 0 K to 300 K over 20 ps using a Langevin thermostat with a collision frequency of 1.0 ps⁻¹ in periodic boundary conditions with constant volume. The equilibration stage ran using the periodic boundary conditions with constant pressure for 5 ns with 1 fs step using Langevin dynamics with a frequency collision of 1 ps⁻¹ to maintain temperature. Using Langevin dynamics the production stage was run for 50 ns with a 2 fs time step with a collision frequency of 1 ps⁻¹ to maintain constant temperature. Long-range electrostatic interactions were modelled using the Particle Mesh Ewald method with a non-bonded cut-off of 10 Å and the SHAKE

algorithm. The coordinates were saved at an interval of 1 ps. The number of added water molecules is provided in **Supplementary Information, Table S3**.

Mixed-solvent molecular dynamics simulations

Cosolvent preparation

Five different cosolvents: acetonitrile (ACN), dimethylsulfoxide (DMSO), methanol (MEO), phenol (PHN), and urea (URE) were selected as representatives of different chemical classes to perform the MixMD simulations. The choice of cosolvents was mainly based on literature research (methanol, dimethylsulfoxide, acetonitrile, phenol)[54] and in the case of urea, its choice was dictated by the fact that in the known structure of the pharmacophore directed into hsEH (proposed by Hammock), urea is a preferred component as a “primary pharmacophore”[35]. The chemical structures of all cosolvent molecules were downloaded from the ChemSpider database[111]. Then, a dedicated set of parameters were prepared. Parameters for ACN come from work by Nikitin and Lyubartsev[112] and parameters for URE were modified using the 8Mureabox force field to obtain parameters for a single molecule. For the rest of the molecules, parameters were prepared using Antechamber[113] with Gasteiger charges[114].

Building initial configurations for mixed-solvent molecular dynamic simulations

Packmol software[115] was used to build the initial systems consisting of protein (protonated according to the previously described procedure), water, and particular cosolvent molecules. No counterions were added to the systems. The following factors were taken into account for the selection of the number of cosolvent molecules. The number of cosolvent molecules should be high enough to contribute in exchange with the protein interior and should be tolerated by the protein – the enzyme should be active in the proposed concentration. The percentage concentration of the cosolvent should not exceed 20% (in the case of phenol, the concentration should not exceed 1% to avoid the aggregation of phenol molecules on the protein surface during the simulation) while maintaining the number of water molecules approximate to the number of water molecules in the classical molecular dynamics simulation. Even though we could use higher concentrations, since Archelas *et al.* showed that epoxide hydrolases do not lose the activity in high concentrations of organic solvents[116], we decreased it to 20% limit. As emphasised in the work of Ung *et. al*, a lower cosolvent concentration in MixMD simulations significantly improve the signal-to-noise ratio of genuine binding hot-spots over spurious ones[57]. The number of added water and cosolvent molecules, as well as the calculated final percentage concentrations is provided in **Supplementary Information, Table S3**. Also, the MixMD simulation procedure, carried out using the AMBER 14 package, was identical to the cMD simulation.

Water and cosolvent molecular tracking

AQUA-DUCT 1.0 software was used to track water and cosolvent molecules. Molecules of interest, which have entered the so-called *Object*, defined as 5 Å sphere around the centre of geometry of active site residues, namely D333, Y383, Y466, D496, and H524, were traced within the *Scope* region, defined as the

interior of a convex hull of hsEH Ca atoms. Points at which the molecules of interest enter or leave the Scope (so-called inlets) were clustered to provide statistical data on the transportation of particular molecules through the protein tunnel network.

Hot-spot identification

AQ was used to detect regions occupied by molecules of interests (so-called inner pockets) and identify the densest sites using a local solvent distribution approach. Those hot-spots are considered potential binding sites. For visualisation purposes and clarity, the size of each sphere representing a particular hot-spot has been changed to reflect its occupation level.

LigPlot analysis

100 crystal structures of human soluble epoxide hydrolase complexed with inhibitors (status as of December 2019) were downloaded from the PDB and chain A (P in the case of 1zd2 structure) and were selected for Ligplot[117] analysis. Additionally, the phosphatase domain was manually removed. The first step of the analysis involved reading the 3D coordinates of the protein structures obtained from the PDB database and identification of atoms belonging to inhibitors. The HBPLUS program[118] was used for generating a list of both hydrogen and non-bonded interactions between hsEH and ligands. In the case of identifying the hydrogen bonds, the program computes all possible positions for hydrogen atoms (H) attached to donor atoms (D) that satisfy geometrical criteria with acceptor atoms (A) in the vicinity. The criteria were as follows: H-A distance was $< 2.7 \text{ \AA}$ and the D-A distance was $< 3.35 \text{ \AA}$. The program also lists all possible non-bonded contacts between atoms that are less than a specified distance apart. Here, the cutoff was set to $3.9 \text{ \AA}$. After the Ligplot analysis was carried out, all possible interactions between hsEH and inhibitors were summarised in **Supplementary Information, Table S2**.

Clustering

The hierarchical clustering method[119] was used to group amino acids, as well as inhibitors, based on the interaction similarity expressed in binary form (1 means interaction, 0 means no interaction). A cluster package from SciPy library was used to perform hierarchical clustering. *Jaacard* metric[120] was used to compute the distance between the points and the linkage was performed using the *average* method[121].

List Of Abbreviations

hsEH, human soluble epoxide hydrolase; EETs, epoxyeicosatrienoic acids; cMD, classical molecular dynamics simulations; MixMD, mixed-solvent molecular dynamics simulations; NTD, N-terminal domain; CTD, C-terminal domain; EPHX2, bifunctional epoxide hydrolase 2 gene; DHETs, dihydroxyeicosatrienoic acids; sEH, soluble epoxide hydrolase; 5-LO, 5-lipoxygenase; CMD, cosolvent-based molecular dynamics simulations; AQ, AQUA-DUCT; ACN, acetonitrile; DMSO, dimethylsulfoxide; MEO, methanol; PHN, phenol; URE, urea; 3D-RISM, Three-Dimensional Reference Interaction Site Model; LS, LigandScout

Declarations

Availability of data and materials

Not applicable.

Competing interests

Not applicable.

Funding

The authors' work was supported by the National Science Centre, Poland (grant no. DEC-2013/10/E/NZ1/00649).

Authors' contributions

Conceptualization, A.G.; methodology, M.B., and K.M.; software, M.B. and K.M.; validation, M.B., K.M. and A.G.; formal analysis, M.B.; investigation, M.B.; resources, M.B. and K.M.; data curation, M.B. and K.M.; writing—original draft preparation, M.B.; writing—review and editing, K.M. and A.G.; visualization, M.B., and A.G.; supervision, A.G.; project administration, A.G.; funding acquisition, A.G. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

The authors would like to thank Tomasz Magdziarz for his helpful advice in clustering analysis of data.

References

1. Gomez GA, Morisseau C, Hammock BD, Christianson DW (2004) Structure of Human Epoxide Hydrolase Reveals Mechanistic Inferences on Bifunctional Catalysis in Epoxide and Phosphate Ester Hydrolysis. *Biochemistry* 43:4716–4723. <https://doi.org/10.1021/bi036189j>
2. Newman JW, Morisseau C, Hammock BD (2005) Epoxide hydrolases: Their roles and interactions with lipid metabolism. *Prog Lipid Res* 44:1–51. <https://doi.org/10.1016/j.plipres.2004.10.001>
3. Harris TR, Hammock BD (2013) Soluble epoxide hydrolase: Gene structure, expression and deletion. *Gene* 526:61–74. <https://doi.org/10.1016/j.gene.2013.05.008>
4. Norwood S, Liao J, Hammock BD, Yang G-Y (2019) Epoxyeicosatrienoic acids and soluble epoxide hydrolase: Potential therapeutic target for inflammation and its induced carcinogenesis. *Am J Transl Res* 2:447–457
5. Cronin A, Decker M, Arand M (2011) Mammalian soluble epoxide hydrolase is identical to liver hepoxilin hydrolase. *J Lipid Res* 52:712–719. <https://doi.org/10.1194/jlr.M009639>
6. Liu J-Y (2019) Inhibition of Soluble Epoxide Hydrolase for Renal Health. *Front Pharmacol* 9:. <https://doi.org/10.3389/fphar.2018.01551>

7. Seubert JM, Zeldin DC, Nithipatikom K, Gross GJ (2007) Role of epoxyeicosatrienoic acids in protecting the myocardium following ischemia/reperfusion injury. *Prostaglandins Other Lipid Mediat* 82:50–59. <https://doi.org/10.1016/j.prostaglandins.2006.05.017>
8. Wagner K, Vito S, Inceoglu B, Hammock BD (2014) The role of long chain fatty acids and their epoxide metabolites in nociceptive signaling. *Prostaglandins Other Lipid Mediat* 113–115:2–12. <https://doi.org/10.1016/j.prostaglandins.2014.09.001>
9. Spector AA, Fang X, Snyder GD, Weintraub NL (2004) Epoxyeicosatrienoic acids (EETs): metabolism and biochemical function. *Prog Lipid Res* 43:55–90. [https://doi.org/10.1016/S0163-7827\(03\)00049-3](https://doi.org/10.1016/S0163-7827(03)00049-3)
10. Larsen BT, Gutterman DD, Hatoum OA (2006) Emerging role of epoxyeicosatrienoic acids in coronary vascular function. *Eur J Clin Invest* 36:293–300. <https://doi.org/10.1111/j.1365-2362.2006.01634.x>
11. Thomson SJ, Askari A, Bishop-Bailey D (2012) Anti-Inflammatory Effects of Epoxyeicosatrienoic Acids. *Int J Vasc Med* 2012:1–7. <https://doi.org/10.1155/2012/605101>
12. Jones RD, Liao J, Tong X et al (2019) Epoxy-Oxylipins and Soluble Epoxide Hydrolase Metabolic Pathway as Targets for NSAID-Induced Gastroenteropathy and Inflammation-Associated Carcinogenesis. *Front Pharmacol* 10:. <https://doi.org/10.3389/fphar.2019.00731>
13. Tacconelli S, Patrignani P (2014) Inside epoxyeicosatrienoic acids and cardiovascular disease. *Front Pharmacol* 5:. <https://doi.org/10.3389/fphar.2014.00239>
14. Yang L, Mäki-Petäjä K, Cheriyan J et al (2015) The role of epoxyeicosatrienoic acids in the cardiovascular system. *Br J Clin Pharmacol* 80:28–44. <https://doi.org/10.1111/bcp.12603>
15. Imig JD, Simpkins AN, Renic M, Harder DR (2011) Cytochrome P450 eicosanoids and cerebral vascular function. *Expert Rev Mol Med* 13:e7. <https://doi.org/10.1017/S1462399411001773>
16. Liu X, Davis CM, Alkayed NJ (2018) P450 Eicosanoids and Reactive Oxygen Species Interplay in Brain Injury and Neuroprotection. *Antioxid Redox Signal* 28:987–1007. <https://doi.org/10.1089/ars.2017.7056>
17. Zuloaga KL, Zhang W, Roese NE, Alkayed NJ (2015) Soluble epoxide hydrolase gene deletion improves blood flow and reduces infarct size after cerebral ischemia in reproductively senescent female mice. *Front Pharmacol* 5:. <https://doi.org/10.3389/fphar.2014.00290>
18. Wu C, Shyue S, Hung T et al (2017) Genetic deletion or pharmacological inhibition of soluble epoxide hydrolase reduces brain damage and attenuates neuroinflammation after intracerebral hemorrhage. 1–21. <https://doi.org/10.1186/s12974-017-1005-4>
19. Darwesh AM, Keshavarz-Bahaghighat H, Jamieson KL, Seubert JM (2019) Genetic Deletion or Pharmacological Inhibition of Soluble Epoxide Hydrolase Ameliorates Cardiac Ischemia/Reperfusion Injury by Attenuating NLRP3 Inflammasome Activation. *Int J Mol Sci* 20:3502. <https://doi.org/10.3390/ijms20143502>
20. Samokhvalov V, Jamieson KL, Darwesh AM et al (2019) Deficiency of Soluble Epoxide Hydrolase Protects Cardiac Function Impaired by LPS-Induced Acute Inflammation. *Front Pharmacol* 9:. <https://doi.org/10.3389/fphar.2018.01572>

21. Ren Q, Ma M, Yang J et al (2018) Soluble epoxide hydrolase plays a key role in the pathogenesis of Parkinson's disease. *Proc Natl Acad Sci* 115:E5815–E5823.
<https://doi.org/10.1073/pnas.1802179115>
22. Zhang G, Kodani S, Hammock BD (2014) Stabilized epoxygenated fatty acids regulate inflammation, pain, angiogenesis and cancer. *Prog Lipid Res* 53:108–123.
<https://doi.org/10.1016/j.plipres.2013.11.003>
23. Zhang G, Panigrahy D, Hwang SH et al (2014) Dual inhibition of cyclooxygenase-2 and soluble epoxide hydrolase synergistically suppresses primary tumor growth and metastasis. *Proc Natl Acad Sci* 111:11127–11132. <https://doi.org/10.1073/pnas.1410432111>
24. Mullin CA, Hammock BD (1982) Chalcone oxides—potent selective inhibitors of cytosolic epoxide hydrolase. *Arch Biochem Biophys* 216:423–439. [https://doi.org/10.1016/0003-9861\(82\)90231-4](https://doi.org/10.1016/0003-9861(82)90231-4)
25. Morisseau C, Du G, Newman JW, Hammock BD (1998) Mechanism of Mammalian Soluble Epoxide Hydrolase Inhibition by Chalcone Oxide Derivatives. *Arch Biochem Biophys* 356:214–228.
<https://doi.org/10.1006/abbi.1998.0756>
26. Morisseau C, Hammock BD (2005) EPOXIDE HYDROLASES: Mechanisms, Inhibitor Designs, and Biological Roles. *Annu Rev Pharmacol Toxicol* 45:311–333.
<https://doi.org/10.1146/annurev.pharmtox.45.120403.095920>
27. Morisseau C, Goodrow MH, Dowdy D et al (1999) Potent urea and carbamate inhibitors of soluble epoxide hydrolases. *Proc Natl Acad Sci* 96:8849–8854. <https://doi.org/10.1073/pnas.96.16.8849>
28. Morisseau C, Hammock BD (2008) Gerry Brooks and epoxide hydrolases: four decades to a pharmaceutical. *Pest Manag Sci* 64:594–609. <https://doi.org/10.1002/ps.1583>
29. Ghosh S, Chiang P-C, Wahlstrom JL et al (2008) Oral Delivery of 1,3-Dicyclohexylurea Nanosuspension Enhances Exposure and Lowers Blood Pressure in Hypertensive Rats. *Basic Clin Pharmacol Toxicol* 102:453–458. <https://doi.org/10.1111/j.1742-7843.2008.00213.x>
30. Imig JD (2006) Cardiovascular Therapeutic Aspects of Soluble Epoxide Hydrolase Inhibitors. *Cardiovasc Drug Rev* 24:169–188. <https://doi.org/10.1111/j.1527-3466.2006.00169.x>
31. Xu D, Li N, He Y et al (2006) Prevention and reversal of cardiac hypertrophy by soluble epoxide hydrolase inhibitors. *Proc Natl Acad Sci* 103:18733–18738.
<https://doi.org/10.1073/pnas.0609158103>
32. Ulu A, Davis BB, Tsai H-J et al (2008) Soluble Epoxide Hydrolase Inhibitors Reduce the Development of Atherosclerosis in Apolipoprotein E-Knockout Mouse Model. *J Cardiovasc Pharmacol* 52:314–323. <https://doi.org/10.1097/FJC.0b013e318185fa3c>
33. Kim I-H, Morisseau C, Watanabe T, Hammock BD (2004) Design, Synthesis, and Biological Activity of 1,3-Disubstituted Ureas as Potent Inhibitors of the Soluble Epoxide Hydrolase of Increased Water Solubility. *J Med Chem* 47:2110–2122. <https://doi.org/10.1021/jm030514j>
34. Kim I-H, Heitzler FR, Morisseau C et al (2005) Optimization of Amide-Based Inhibitors of Soluble Epoxide Hydrolase with Improved Water Solubility. *J Med Chem* 48:3621–3629.
<https://doi.org/10.1021/jm0500929>

35. Li H-Y, Jin Y, Morriseau C et al (2006) The 5-substituted piperazine as a novel secondary pharmacophore greatly improving the physical properties of urea-based inhibitors of soluble epoxide hydrolase. *Bioorg Med Chem* 14:6586–6592
36. Anandan S-K, Webb HK, Chen D et al (2011) 1-(1-Acetyl-piperidin-4-yl)-3-adamantan-1-yl-urea (AR9281) as a potent, selective, and orally available soluble epoxide hydrolase inhibitor with efficacy in rodent models of hypertension and dysglycemia. *Bioorg Med Chem Lett* 21:983–988. <https://doi.org/10.1016/j.bmcl.2010.12.042>
37. Chen D, Whitcomb R, MacIntyre E et al (2012) Pharmacokinetics and Pharmacodynamics of AR9281, an Inhibitor of Soluble Epoxide Hydrolase, in Single- and Multiple-Dose Studies in Healthy Human Subjects. *J Clin Pharmacol* 52:319–328. <https://doi.org/10.1177/0091270010397049>
38. www.clinicaltrials.gov (2009) The trial was terminated in November 2009 according to www.clinicaltrials.gov and no efficacy results have been reported
39. Morriseau C, Goodrow MH, Newman JW et al (2002) Structural refinement of inhibitors of urea-based soluble epoxide hydrolases. *Biochem Pharmacol* 63:1599–1608. [https://doi.org/10.1016/S0006-2952\(02\)00952-8](https://doi.org/10.1016/S0006-2952(02)00952-8)
40. Lazaar AL, Yang L, Boardley RL et al (2016) Pharmacokinetics, pharmacodynamics and adverse event profile of GSK2256294, a novel soluble epoxide hydrolase inhibitor. *Br J Clin Pharmacol* 81:971–979. <https://doi.org/10.1111/bcp.12855>
41. Amano Y, Yamaguchi T, Tanabe E (2014) Structural insights into binding of inhibitors to soluble epoxide hydrolase gained by fragment screening and X-ray crystallography. *Bioorganic Med Chem* 22:2427–2434. <https://doi.org/10.1016/j.bmc.2014.03.001>
42. Amano Y, Tanabe E, Yamaguchi T (2015) Identification of N-ethylmethylaniline as a novel scaffold for inhibitors of soluble epoxide hydrolase by crystallographic fragment screening. *Bioorganic Med Chem* 23:2310–2317. <https://doi.org/10.1016/j.bmc.2015.03.083>
43. Öster L, Tapani S, Xue Y, Käck H (2015) Successful generation of structural information for fragment-based drug discovery. *Drug Discov Today* 20:1104–1111. <https://doi.org/10.1016/j.drudis.2015.04.005>
44. Xue Y, Olsson T, Johansson CA et al (2016) Fragment Screening of Soluble Epoxide Hydrolase for Lead Generation-Structure-Based Hit Evaluation and Chemistry Exploration. *ChemMedChem* 11:497–508. <https://doi.org/10.1002/cmdc.201500575>
45. Tanaka D, Tsuda Y, Shiyama T et al (2011) A Practical Use of Ligand Efficiency Indices Out of the Fragment-Based Approach: Ligand Efficiency-Guided Lead Identification of Soluble Epoxide Hydrolase Inhibitors. *J Med Chem* 54:851–857. <https://doi.org/10.1021/jm101273e>
46. Xing L, McDonald JJ, Kolodziej SA et al (2011) Discovery of Potent Inhibitors of Soluble Epoxide Hydrolase by Combinatorial Library Design and Structure-Based Virtual Screening. *J Med Chem* 54:1211–1222. <https://doi.org/10.1021/jm101382t>
47. Moser D, Achenbach J, Klingler F-M et al (2012) Evaluation of structure-derived pharmacophore of soluble epoxide hydrolase inhibitors by virtual screening. *Bioorg Med Chem Lett* 22:6762–6765.

<https://doi.org/10.1016/j.bmcl.2012.08.066>

48. Moser D, Wisniewska JM, Hahn S et al (2012) Dual-Target Virtual Screening by Pharmacophore Elucidation and Molecular Shape Filtering. *ACS Med Chem Lett* 3:155–158.
<https://doi.org/10.1021/ml200286e>
49. Waltenberger B, Garscha U, Temml V et al (2016) Discovery of Potent Soluble Epoxide Hydrolase (sEH) Inhibitors by Pharmacophore-Based Virtual Screening. *J Chem Inf Model* 56:747–762.
<https://doi.org/10.1021/acs.jcim.5b00592>
50. Tripathi N, Paliwal S, Sharma S et al (2018) Discovery of Novel Soluble Epoxide Hydrolase Inhibitors as Potent Vasodilators. *Sci Rep* 8:14604. <https://doi.org/10.1038/s41598-018-32449-4>
51. Bhagwati S, Siddiqi MI (2019) Identification of potential soluble epoxide hydrolase (sEH) inhibitors by ligand-based pharmacophore model and biological evaluation. *J Biomol Struct Dyn* 1–11.
<https://doi.org/10.1080/07391102.2019.1691659>
52. Seco J, Luque FJ, Barril X (2009) Binding Site Detection and Druggability Index from First Principles. *J Med Chem* 52:2363–2371. <https://doi.org/10.1021/jm801385d>
53. Lexa KW, Carlson HA (2011) Full Protein Flexibility Is Essential for Proper Hot-Spot Mapping. *J Am Chem Soc* 133:200–202. <https://doi.org/10.1021/ja1079332>
54. Ghanakota P, Carlson HA (2016) Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *J Med Chem* 59:10383–10399.
<https://doi.org/10.1021/acs.jmedchem.6b00399>
55. Ghanakota P, van Vlijmen H, Sherman W, Beuming T (2018) Large-Scale Validation of Mixed-Solvent Simulations to Assess Hotspots at Protein–Protein Interaction Interfaces. *J Chem Inf Model* 58:784–793. <https://doi.org/10.1021/acs.jcim.7b00487>
56. Uehara S, Tanaka S (2017) Cosolvent-Based Molecular Dynamics for Ensemble Docking: Practical Method for Generating Druggable Protein Conformations. *J Chem Inf Model* 57:742–756.
<https://doi.org/10.1021/acs.jcim.6b00791>
57. Ung PMU, Ghanakota P, Graham SE et al (2016) Identifying binding hot spots on protein surfaces by mixed-solvent molecular dynamics: HIV-1 protease as a test case. *Biopolymers* 105:21–34.
<https://doi.org/10.1002/bip.22742>
58. Sabanés Zariquiey F, de Souza JV, Bronowska AK (2019) Cosolvent Analysis Toolkit (CAT): a robust hotspot identification platform for cosolvent simulations of proteins to expand the druggable proteome. *Sci Rep* 9:19118. <https://doi.org/10.1038/s41598-019-55394-2>
59. Lee JY, Krieger JM, Li H, Bahar I (2020) Pharmed: Pharmacophore modeling and hit identification based on druggability simulations. *Protein Sci* 29:76–86. <https://doi.org/10.1002/pro.3732>
60. Ghanakota P, Carlson HA (2016) Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *J Med Chem* 59:10383–10399.
<https://doi.org/10.1021/acs.jmedchem.6b00399>
61. Alvarez-Garcia D, Barril X (2014) Relationship between Protein Flexibility and Binding: Lessons for Structure-Based Drug Design. *J Chem Theory Comput* 10:2608–2614.

<https://doi.org/10.1021/ct500182z>

62. Yu W, Lakkaraju SK, Raman EP et al (2015) Pharmacophore Modeling Using Site-Identification by Ligand Competitive Saturation (SILCS) with Multiple Probe Molecules. *J Chem Inf Model* 55:407–420. <https://doi.org/10.1021/ci500691p>
63. Yang C-Y, Wang S (2010) Computational Analysis of Protein Hotspots. *ACS Med Chem Lett* 1:125–129. <https://doi.org/10.1021/ml100026a>
64. Yang C-Y, Wang S (2011) Hydrophobic Binding Hot Spots of Bcl-xL Protein – Protein Interfaces by Cosolvent Molecular Dynamics Simulation. *ACS Med Chem Lett* 2:280–284. <https://doi.org/10.1021/ml100276b>
65. Bakan A, Nevins N, Lakdawala AS, Bahar I (2012) Druggability Assessment of Allosteric Proteins by Dynamics Simulations in the Presence of Probe Molecules. *J Chem Theory Comput* 8:2435–2447. <https://doi.org/10.1021/ct300117j>
66. Huang D, Rossini E, Steiner S, Caflisch A (2014) Structured Water Molecules in the Binding Site of Bromodomains Can Be Displaced by Cosolvent. *ChemMedChem* 9:573–579. <https://doi.org/10.1002/cmdc.201300156>
67. Tan YS, Spring DR, Abell C, Verma C (2014) The Use of Chlorobenzene as a Probe Molecule in Molecular Dynamics Simulations. *J Chem Inf Model* 54:1821–1827. <https://doi.org/10.1021/ci500215x>
68. Tan YS, Spring DR, Abell C, Verma CS (2015) The Application of Ligand-Mapping Molecular Dynamics Simulations to the Rational Design of Peptidic Modulators of Protein–Protein Interactions. *J Chem Theory Comput* 11:3199–3210. <https://doi.org/10.1021/ct5010577>
69. Prakash P, Sayyed-Ahmad A, Gorfe AA (2015) pMD-Membrane: A Method for Ligand Binding Site Identification in Membrane-Bound Proteins. *PLOS Comput Biol* 11:e1004469. <https://doi.org/10.1371/journal.pcbi.1004469>
70. Arcon JP, Defelipe LA, Modenutti CP et al (2017) Molecular Dynamics in Mixed Solvents Reveals Protein–Ligand Interactions, Improves Docking, and Allows Accurate Binding Free Energy Predictions. *J Chem Inf Model* 57:846–863. <https://doi.org/10.1021/acs.jcim.6b00678>
71. Graham SE, Leja N, Carlson HA (2018) MixMD Probeview: Robust Binding Site Prediction from Cosolvent Simulations. *J Chem Inf Model* 58:1426–1433. <https://doi.org/10.1021/acs.jcim.8b00265>
72. Tan YS, Śledź P, Lang S et al (2012) Using Ligand-Mapping Simulations to Design a Ligand Selectively Targeting a Cryptic Surface Pocket of Polo-Like Kinase 1. *Angew Chem Int E* 51:10078–10081. <https://doi.org/10.1002/anie.20120567>
73. Lexa KW, Goh GB, Carlson HA (2014) Parameter Choice Matters: Validating Probe Parameters for Use in Mixed-Solvent Simulations. *J Chem Inf Model* 54:2190–2199. <https://doi.org/10.1021/ci400741u>
74. Defelipe L, Arcon J, Modenutti C et al (2018) Solvents to Fragments to Drugs: MD Applications in Drug Design. *Molecules* 23:3269. <https://doi.org/10.3390/molecules23123269>
75. English AC, Groom CR, Hubbard RE (2001) Experimental and computational mapping of the binding surface of a crystalline protein. *Protein Eng Des Sel* 14:47–59.

<https://doi.org/10.1093/protein/14.1.47>

76. Lexa KW, Carlson HA (2013) Improving Protocols for Protein Mapping through Proper Comparison to Crystallography Data. *J Chem Inf Model* 53:391–402. <https://doi.org/10.1021/ci300430v>
77. English AC, Done SH, Caves LSD, et al (1999) Locating interaction sites on proteins: The crystal structure of thermolysin soaked in 2–100% isopropanol. *Proteins Struct Funct Genet* 37:628–640. [https://doi.org/10.1002/\(SICI\)1097-0134\(19991201\)37:4<628::AID-PROT13>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0134(19991201)37:4<628::AID-PROT13>3.0.CO;2-G)
78. Ho WC, Luo C, Zhao K et al (2006) High-resolution structure of the p53 core domain: implications for binding small-molecule stabilizing compounds. *Acta Crystallogr Sect D Biol Crystallogr* 62:1484–1493. <https://doi.org/10.1107/S090744490603890X>
79. Basse N, Kaar JL, Settanni G et al (2010) Toward the Rational Design of p53-Stabilizing Drugs: Probing the Surface of the Oncogenic Y220C Mutant. *Chem Biol* 17:46–56. <https://doi.org/10.1016/j.chembiol.2009.12.011>
80. Miranker A, Karplus M (1991) Functionality maps of binding sites: A multiple copy simultaneous search method. *Proteins Struct Funct Genet* 11:29–34. <https://doi.org/10.1002/prot.340110104>
81. Alvarez-Garcia D, Barril X (2014) Molecular Simulations with Solvent Competition Quantify Water Displaceability and Provide Accurate Interaction Maps of Protein Binding Sites. *J Med Chem* 57:8530–8539. <https://doi.org/10.1021/jm5010418>
82. Raman EP, Yu W, Guvench O, MacKerell AD (2011) Reproducing Crystal Binding Modes of Ligand Functional Groups Using Site-Identification by Ligand Competitive Saturation (SILCS) Simulations. *J Chem Inf Model* 51:877–896. <https://doi.org/10.1021/ci100462t>
83. Hall DH, Grove LE, Yueh C et al (2011) Robust Identification of Binding Hot Spots Using Continuum Electrostatics: Application to Hen Egg-White Lysozyme. *J Am Chem Soc* 133:20668–20671. <https://doi.org/10.1021/ja207914y>
84. Ghanakota P, Carlson HA (2016) Moving Beyond Active-Site Detection: MixMD Applied to Allosteric Systems. *J Phys Chem B* 120:8685–8695. <https://doi.org/10.1021/acs.jpcb.6b03515>
85. Magdziarz T, Mitusińska K, Gołdowska S et al (2017) AQUA-DUCT: a ligands tracking tool. *Bioinformatics* 33:2045–2046. <https://doi.org/10.1093/bioinformatics/btx125>
86. Magdziarz T, Mitusińska K, Bzówka M et al (2020) AQUA-DUCT 1.0: structural and functional analysis of macromolecules from an intramolecular voids perspective. *Bioinformatics* 36:2599–2601. <https://doi.org/10.1093/bioinformatics/btz946>
87. Berman HM (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>
88. Mitusińska K, Magdziarz T, Bzówka M et al (2018) Exploring Solanum tuberosum Epoxide Hydrolase Internal Architecture by Water Molecules Tracking. *Biomolecules* 8:143. <https://doi.org/10.3390/biom8040143>
89. Klausen MS, Jespersen MC, Nielsen H et al (2019) NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. *Proteins Struct Funct Bioinforma* 87:520–527. <https://doi.org/10.1002/prot.25674>

90. Hopmann KH, Himo F (2006) Theoretical Study of the Full Reaction Mechanism of Human Soluble Epoxide Hydrolase. *Chem - A Eur J* 12:6898–6909. <https://doi.org/10.1002/chem.200501519>
91. Gora A, Brezovsky J, Damborsky J (2013) Gates of Enzymes. *Chem Rev* 113:5871–5923. <https://doi.org/10.1021/cr300384w>
92. Marques SM, Daniel L, Buryska T et al (2017) Enzyme Tunnels and Gates As Relevant Targets in Drug Design. *Med Res Rev* 37:1095–1139. <https://doi.org/10.1002/med.21430>
93. Huson DH, Scornavacca C (2012) Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. *Syst Biol* 61:1061–1067. <https://doi.org/10.1093/sysbio/sys062>
94. Gomez GA, Morisseau C, Hammock BD, Christianson DW (2006) Human soluble epoxide hydrolase: Structural basis of inhibition by 4- (3-cyclohexylureido) -carboxylic acids. 58–64. <https://doi.org/10.1110/ps.051720206.58>
95. Chen H, Zhang Y, Ye C et al (2014) Insight into the binding modes and inhibition mechanisms of adamantyl-based 1,3-disubstituted urea inhibitors in the active site of the human soluble epoxide hydrolase. *J Biomol Struct Dyn* 32:1231–1247. <https://doi.org/10.1080/07391102.2013.812981>
96. Gomez GA, Morisseau C, Hammock BD, Christianson DW (2004) Structure of Human Epoxide Hydrolase Reveals Mechanistic Inferences on Bifunctional Catalysis in Epoxide and Phosphate Ester Hydrolysis † ‡. *Biochemistry* 43:4716–4723. <https://doi.org/10.1021/bi036189j>
97. Lo HY, Man CC, Fleck RW et al (2010) Substituted pyrazoles as novel sEH antagonist: Investigation of key binding interactions within the catalytic domain. *Bioorg Med Chem Lett* 20:6379–6383. <https://doi.org/10.1016/j.bmcl.2010.09.095>
98. Takai K, Chiyo N, Nakajima T et al (2015) Three-dimensional rational approach to the discovery of potent substituted cyclopropyl urea soluble epoxide hydrolase inhibitors. *Bioorg Med Chem Lett* 25:1705–1708. <https://doi.org/10.1016/j.bmcl.2015.02.076>
99. Pravda L, Berka K, Svobodová Vařeková R et al (2014) Anatomy of enzyme channels. *BMC Bioinformatics* 15:379. <https://doi.org/10.1186/s12859-014-0379-x>
100. Garcin ED, Arvai AS, Rosenfeld RJ et al (2008) Anchored plasticity opens doors for selective inhibitor design in nitric oxide synthase. *Nat Chem Biol* 4:700–707. <https://doi.org/10.1038/nchembio.115>
101. Cheshire DR, Åberg A, Andersson GMK et al (2011) The discovery of novel, potent and highly selective inhibitors of inducible nitric oxide synthase (iNOS). *Bioorg Med Chem Lett* 21:2468–2471. <https://doi.org/10.1016/j.bmcl.2011.02.061>
102. Tuncbag N, Gursoy A, Keskin O (2009) Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics* 25:1513–1520. <https://doi.org/10.1093/bioinformatics/btp240>
103. Zerbe BS, Hall DR, Vajda S et al (2012) Relationship between Hot Spot Residues and Ligand Binding Hot Spots in Protein–Protein Interfaces. *J Chem Inf Model* 52:2236–2244. <https://doi.org/10.1021/ci300175u>
104. Brenke R, Kozakov D, Chuang G-Y et al (2009) Fragment-based identification of druggable ‘hot spots’ of proteins using Fourier domain correlation techniques. *Bioinformatics* 25:621–627.

<https://doi.org/10.1093/bioinformatics/btp036>

105. Kozakov D, Grove LE, Hall DR et al (2015) The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat Protoc* 10:733–755.
<https://doi.org/10.1038/nprot.2015.043>
106. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger; Schrödinger LLC.: New York, NY, USA
107. Gordon JC, Myers JB, Folta T et al (2005) H⁺⁺: a server for estimating pK_as and adding missing hydrogens to macromolecules. *Nucleic Acids Res* 33:W368–W371.
<https://doi.org/10.1093/nar/gki464>
108. Luchko T, Gusarov S, Roe DR et al (2010) Three-Dimensional Molecular Theory of Solvation Coupled with Molecular Dynamics in Amber. *J Chem Theory Comput* 6:607–624.
<https://doi.org/10.1021/ct900460m>
109. Sindhikara DJ, Yoshida N, Hirata F (2012) Placevent: An algorithm for prediction of explicit solvent atom distribution-Application to HIV-1 protease and F-ATP synthase. *J Comput Chem* 33:1536–1543.
<https://doi.org/10.1002/jcc.22984>
110. Case DA, Babin V, Berryman JT et al (2014) AMBER 14
111. Pence HE, Williams A (2010) ChemSpider: An Online Chemical Information Resource. *J Chem Educ* 87:1123–1124. <https://doi.org/10.1021/ed100697w>
112. Nikitin AM, Lyubartsev AP (2007) New six-site acetonitrile model for simulations of liquid acetonitrile and its aqueous mixtures. *J Comput Chem* 28:2020–2026. <https://doi.org/10.1002/jcc.20721>
113. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25:247–260.
<https://doi.org/10.1016/j.jmgm.2005.12.005>
114. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* 36:3219–3228. [https://doi.org/10.1016/0040-4020\(80\)80168-2](https://doi.org/10.1016/0040-4020(80)80168-2)
115. Martínez L, Andrade R, Birgin EG, Martínez JM (2009) PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J Comput Chem* 30:2157–2164.
<https://doi.org/10.1002/jcc.21224>
116. Archelas A, Iacazio G, Kotik M (2016) Epoxide hydrolase and its application in organic synthesis. In: *Green Biocatalysis*
117. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Prot Eng* 8:127–134.
<https://doi.org/10.1093/protein/8.2.127>
118. McDonald IK, Thornton JM (1994) Satisfying Hydrogen Bonding Potential in Proteins. *J Mol Biol* 238:777–793. <https://doi.org/10.1006/jmbi.1994.1334>
119. Köhn H-F, Hubert LJ (2015) Hierarchical Cluster Analysis. In: *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd, Chichester, pp 1–13

120. LEVANDOWSKY M, WINTER D (1971) Distance between Sets. *Nature* 234:34–35.
<https://doi.org/10.1038/234034a0>
121. Müllner D (2011) Modern hierarchical, agglomerative clustering algorithms. *ArXiv*

Figures

Figure 1

(a) Crystal structure of hsEH hydrolase domain. The protein is shown in cartoon representation. The main domain is coloured grey, the cap domain is coloured blue. Internal pockets are shown and coloured beige, and the residues building the active site are presented as red sticks. (b) Localisation of areas of water molecule entry/egress (clusters) to the active site of hsEH. Small balls represent single inlets of water molecule entry/egress, colours correspond to identified tunnels (pink – Tm1 (~78%), green – Tc/m (~20%), cyan – Tg (< 1%), orange – Tm2 (< 1%)). Statistical data of tunnels (clusters) entry/egress utilisation are provided for each tunnel. (c) Comparison of water accessible volume during the MD simulations with solvent accessible volume from the crystal structure.

Figure 2

(a) Localisation of the identified amino acids for particular clusters. Amino acids are shown in two different representations – as spheres in the figure presenting the whole protein and as sticks in figures showing specific cluster locations. The inner pockets of the protein are coloured according to the colours of the amino acids' clusters that occupy them. (b) Classification of amino acid obtained as a result of clustering analysis using the Dendroscope 3 tool[93]. Clusters are presented in the form of a radial

cladogram and stars indicate active site residues. (c) Number of inhibitors identified by LigPlot as interacting with particular residue. The consistent colour scheme is provided for all figures: C-1a (light pink), C-1b (deep pink), C-2 (cyan), C-3 (green), C-4 (yellow), and a set of outliers (grey).

Figure 3

Localisation of inhibitors identified for the main clusters. Inhibitors are shown as sticks. The internal pockets of the protein are coloured according to the colours of the amino acids' clusters that occupy them.

Figure 4

Localisation of hot-spots and pockets for individual (co)solvents. Blue – water (a), orange – acetonitrile (b), green – dimethylsulfoxide (c), yellow – methanol (d), red – phenol (e), purple – urea (f). The upper panels present the cluster areas calculated with kernel density estimation method (shown as isolines) together with pockets (shown as mesh) and hot-spots (shown as spheres reflecting their local density).

The lower panels show the cross-section of the protein with pockets (shown as mesh) and hot-spots (shown as spheres).

Figure 5

Global hot-spots identified during the MD simulations. (a) Localisation of all hot-spots. (b) Superposition of all inhibitors bound in the active site with hot-spots. (c) The close-up on residues surrounding the cavity explored by acetonitrile penetrating the cap domain. (d) The close-up on residues surrounding the water hot-spot located at a gorge between the cap and main domains. (e) The close-up on residues surrounding water/acetonitrile hot-spot located near Tm2 tunnel entrance. Hot-spots colour scheme: blue – water, orange – acetonitrile, green – dimethylsulfoxide, yellow – methanol, red – phenol, purple – urea, colour coding of residues surrounding hot-spots: residues colour reflects the predominant neighbouring solvent, residues name red or blue – unique in the AQUA-DUCT analysis (not identified by Ligplot in the previous analysis). For picture clarity, only the most important residues creating a particular pocket are shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [BzowkainhhsEHcheminformaticSI.docx](#)

- [BzowkainhhsEHcheminformaticSI.docx](#)