Similarity between Flavonoid Biosynthetic Enzymes and Flavonoid Protein Targets Captured by Three-Dimensional Computing Approach

Authors   **Noé Sturm**[1,2], **Ronald J. Quinn**[1], **Esther Kellenberger**[2]

Affiliations   [1] Eskitis Institute for Drug Discovery, Griffith University, Brisbane, Australia
[2] Laboratory of Therapeutic Innovation, Medalis Drug Discovery Center, Université de Strasbourg, Illkirch, France

Correspondence
Dr. Esther Kellenberger
Laboratory of Therapeutic
Innovation
Medalis Drug Discovery Center
UMR 7200 CNRS-University of
Strasbourg
74 Route du Rhin
67400 Illkirch
France
Phone: + 33 3 68 85 42 21
Fax: + 33 3 68 85 43 10
ekellen@unistra.fr

Correspondence
Prof. Dr. Ronald James Quinn
Eskitis Institute for Drug
Discovery
Griffith University
Innovation Park, 46 Don Young
Road
Nathan, Brisbane, QLD 4111
Australia
Phone: + 61 7 37 35 60 00
Fax: + 61 7 37 35 60 01
r.quinn@griffith.edu.au

Abstract
▼

Natural products are made by nature through interaction with biosynthetic enzymes. They also exert their effect as drugs by interaction with proteins. To address the question "Do biosynthetic enzymes and therapeutic targets share common mechanisms for the molecular recognition of natural products?", we compared the active site of five flavonoid biosynthetic enzymes to 8077 ligandable binding sites in the Protein Data Bank using two three-dimensional-based methods (SiteAlign and Shaper). Virtual screenings efficiently retrieved known flavonoid targets, in particular protein kinases. A consistent performance obtained for variable site descriptions (presence/absence of water, variable boundaries, or small structural changes) indicated that the methods are robust and thus well suited for the identification of potential target proteins of natural products. Finally, our results suggested that flavonoid binding is not primarily driven by shape, but rather by the recognition of common anchoring points.

Abbreviations
▼

Introduction
▼

Natural products are chemical compounds synthetized by living organisms. Secondary metabolites are those which are dispensable for survival but give particular species their characteristic features. Secondary metabolites have a broad range of functions, for example, toxins and repellants are used as weapons against prey or predators and attractants are used to attract symbiotic organisms [1]. If they have an extrinsic action on other living organisms, natural products usually disturb an important pathway or trigger a specific biological activity. At the molecular scale, they exert their effect as a drug by interacting with biological macromolecules, especially proteins.

Natural products occupy a diverse chemical space and are involved in a large variety of functions, and therefore represent a rich source of therapeutically useful compounds. Around half of all approved drugs are natural products or their derivatives [2]. Discovery of therapeutic natural products is nevertheless challenging. Extraction, purification, and structure characterization are complex tasks. The determination of potential biological activities is also demanding, requiring many biological assays in a trial and error approach. Computational approaches have recently been proposed to facilitate the identification of targets for a compound of interest. Ligand-based methods, which are based on the assumption that similar compounds bind to the same target, have
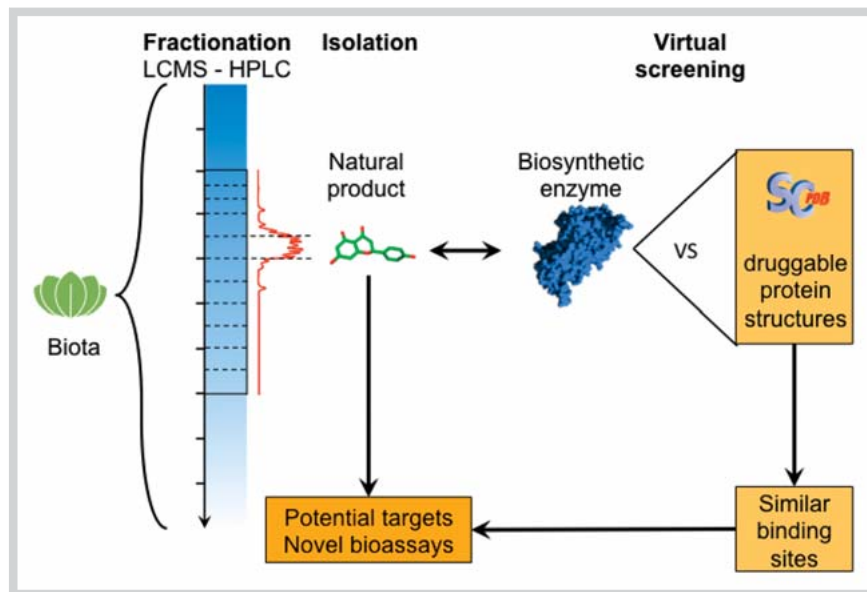
Fig. 1 Ligand-free three-dimensional computing approach to target identification for natural products. (Color figure available online only.)

been successful in drug repositioning and ligand profiling [3]. However, models are predictive only if the biological activity of the explored chemical space is already characterized, thus preventing their application to a novel chemical structure. Structured-based methods in principle circumvent this problem because they interpret the 3D structure of proteins, and do not rely on a training dataset. Docking of a given compound into a series of protein binding sites could efficiently prioritize compounds for experimental testing. A direct comparison of binding sites has also allowed the identification of common ligands of different proteins, assuming that similar binding sites accommodate the same ligand. This second approach is of special interest because it does not depend on a ligand conformational search and gives a robust prediction even if proteins undergo small structural changes [4]. Natural products are made by nature through interaction with biosynthetic enzymes and therefore embed a biological imprint [5,6]. In the present study, we addressed the question "can computing methods find similarity between the active site of biosynthetic enzymes and the binding site of drug targets?". To establish the proof of concept, we focused on flavonoids because different compounds of this class of natural products have been co-crystallized with several biosynthetic enzymes as well as with several protein targets, in particular kinases. The active sites of five different FBEs were used as a query to search the PDB [7] using two different site comparison methods, namely SiteAlign and Shaper (○ **Fig. 1**).

Results and Discussion
▼

In this study, five different proteins were chosen to represent the family of FBEs: CHS, CHI, 2,3QD, DFR, and LAR from the flowering plant *Medicago sativa* (CHS and CHI), the fungus *Aspergillus japonicus* (2,3QD) and the grape vine *Vitis vinifera* (DFR and LAR). These proteins act on nine different substrates in five different pathways of flavonoid metabolism (**Fig. 1S**, Supporting Information) [8], and, therefore, are expected to constitute a representative panel of the possible modes of flavonoid recognition. In support of this hypothesis, the size and composition in amino acids largely differ in the five enzymes (○ **Fig. 2**). In addition, active

sites in the different enzymes are dissimilar, with a single exception (CHS vs. DFR compared using Shaper, **Table 1S**, Supporting Information). The query dataset contains a total of ten different 3D structures, because CHI, 2,3QD, and DFR enzymes were co-crystallized with up to three different flavonoids (○ **Table 1**). Of note, all copies of a given protein site were found to be similar despite slight changes in the site definition and description (**Table 1S**, Supporting Information).

The ten FBE active sites were compared to 8077 protein sites which were selected from the PDB according to their predicted ability to accommodate a small molecular weight ligand with high affinity [9]. The searched set of binding sites, from here on called the screening dataset, represents 2379 proteins (as defined by UniProt identifiers [10]) and 967 enzymatic activities (as described by unique Enzyme Commission numbers [11]). Each protein in the screening dataset was annotated as (1) a FBE if it belonged to the set of query proteins, or (2) a flavonoid target if it was crystallized in complex with a flavonoid (**Table 2S**, Supporting Information) or if a micromolar or better affinity for a flavonoid was reported in the ChEMBL database [12] (IC_{50} or $K_i \leq 10\,\mu M$, **Table 3S**, Supporting Information), or (3) a decoy. Among the 71 flavonoid targets identified, kinases were frequently encountered because the screening dataset is highly enriched in kinases (22% of entries) and in protein kinases (77% of the kinases). Also, flavonoids have been suggested to function as anticancer agents due to the inhibition of protein kinases [13–17]. Several types of steroid receptors, phosphodiesterases, and carbonic anhydrases are also targeted by flavonoids.

Site comparisons were performed using two different methods, namely Shaper and SiteAlign [9,18]. A total of 20 virtual screening experiments were analyzed. Overall performances were assessed by plotting ROC curves [19,20]. The x-axis of ROC curve represents the false positive rate, i.e., selectivity. The y-axis of ROC curve represents the true positive rate, i.e., sensitivity. Here we considered that the number of true positives is the count of FBE and flavonoid targets in the selection and the number of false positives the count of decoys in the selection. Random picking in the screening dataset theoretically produces a diagonal line with an area under the curve (ROCAU) equal to 0.5. Whatever the query site and the comparison method, we observed that ranking
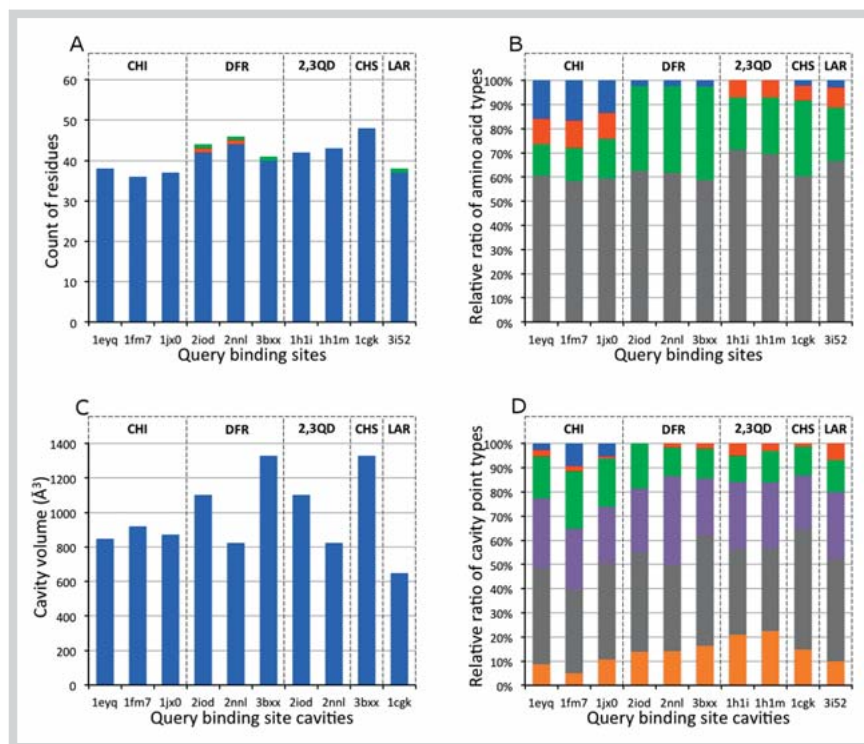
Fig. 2 Description of flavonoid biosynthetic enzyme active sites. **A** Number of amino acids, water molecules, and cofactors in site. Amino acids are colored in blue, water molecules in red, cofactors in green. **B** Composition in amino acids of site. Apolar residues are colored in grey, negatively charged residues in red, positively charged residues in blue, and other polar residues in green. **C** Volume of cavity (Å^3) computed using VolSite. **D** Pharmacophoric description of cavity. Aromatic property is colored in orange, hydrophobic property in grey, hydrogen-bond acceptor in purple, hydrogen-bond donor in green, positive charge in blue, and negative charge in red. (Color figure available online only.)

Table 1 Flavonoid biosynthetic enzymes. Enzyme Commission number indicates the type of reaction catalyzed by the enzyme. UniProt ID is a unique sequence identifier. PDB code is the 3D structure identifier.

| Protein *Species*	Enzyme commission	UniProt ID	Ligand name	PDB code
Chalcone isomerase (CHI) *Medicago sativa*	5.5.1.6	CFI1_MEDSA	Naringenin 5-deoxyflavonol 5-deoxyflavonol	1eyq 1fm7 1jx0
Dihydroflavonol-4-reductase (DFR) *Vitis vinifera*	1.1.1.219	P93 799_VITVI	Myricetin Dihydroquercetin Quercetin	2iod 2 nnl 3bxx
Quercetin 2,3-dioxygenase (2,3QD) *Aspergillus japonicus*	1.13.11.24	QDOI_ASPJA	Quercetin Kaempferol	1h1i 1h1 m
Chalcone Synthase (CHS) *Medicago sativa*	2.3.1.74	CHS2_MEDSA	Naringenin	1cgk
Leucoanthocyanidin reductase 1 (LAR) *Vitis vinifera*	1.17.1.3	Q4W2K4_VITVI	(+)-Catechin	3i52

by similarity is significantly better than random picking (○ **Fig. 3**). The range of ROCAU values was between 0.60 and 0.78 (**Table 4S**, Supporting Information), meaning that predictions were fair to good, respectively.

Comparing methods, we observed that, overall, SiteAlign performed better than Shaper, with ROCAUs in the 0.68–0.78 and 0.60–0.72 ranges, respectively. Since shape superimposition is determinant in predictions made using Shaper while more emphasis is given on pharmacophoric features in SiteAlign, we could postulate that flavonoid binding to flavonoid targets is not primarily driven by shape complementarity, but rather by the recognition of common anchoring points.

For CHI, three 3D structures of the active site were tested as query, yielding almost identical ROC curves and ROCAUs (○ **Fig. 3**; **Table 4S**, Supporting Information). Consistent results were also obtained for the two screenings using DFR queries, and for the three screenings using 2,3QD queries, further demon-

strating that small changes in the size and composition of a query site did not affect the quality of predictions made using SiteAlign and Shaper. Consequently, we concluded that site comparison methods are robust and that there is no quantitative benefit in repeating virtual screening using several similar structures of FBE active site.

To further challenge the methods, we investigated the impact of water molecules on screening results obtained using Shaper (**Table 4S** and **Fig. 2S**, Supporting Information). Noteworthy is that only tightly bound water molecules were included in the sites (more precisely water molecules establishing two or more hydrogen bonds with the protein). FBE sites contained between 0 and 1 water molecules, representing less than 1.3 % of the atoms exposed at the protein site surface. Consequently, water only marginally affected the global description of the query site, with variations in shape and of physicochemical properties being limited to a few spots. These local changes were not sufficient to af-
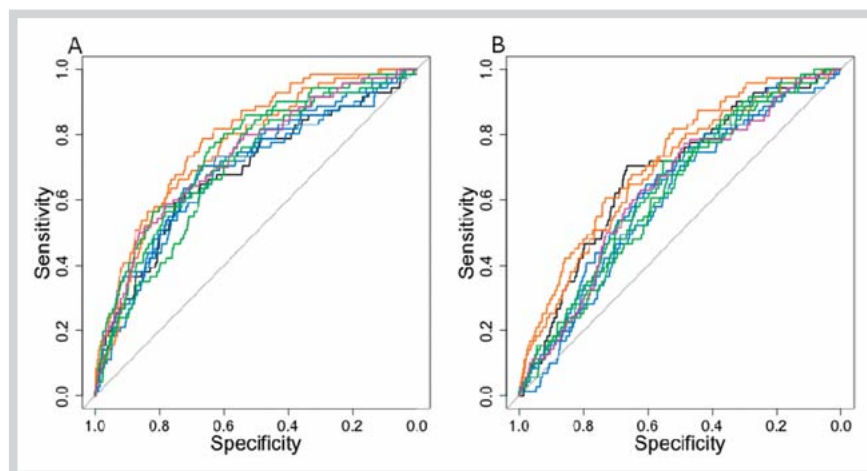
Fig. 3 Receiver operating characteristics curves. **A** SiteAlign. **B** Shaper. Curves are colored according to FBE proteins: CHI in blue, DFR in green, 2,3QD in orange, CHS in black, and LAR in pink. (Color figure available online only.)
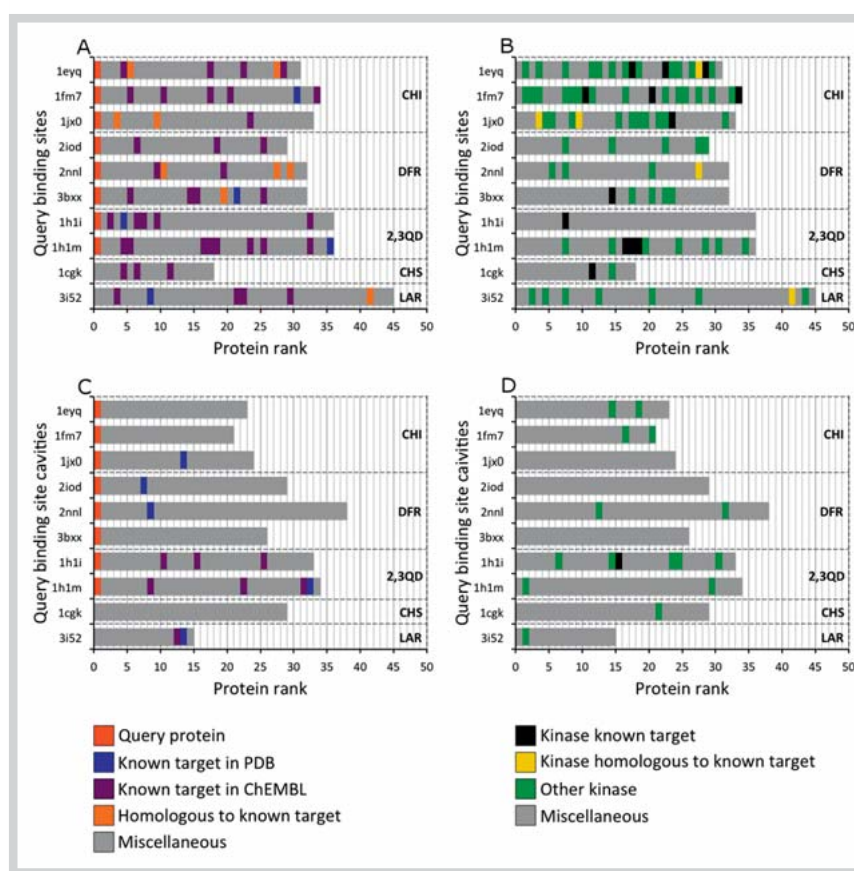
Fig. 4 Composition of hit list. **A** FBE and flavonoid targets in SiteAlign lists. **B** Kinase proteins in SiteAlign hit lists. **C** FBE and flavonoid targets in Shaper lists. **D** Kinase protein in Shaper lists. In **A** and **C**, copies of FBE query are colored in red. Flavonoid targets are colored in blue or purple according to experimental evidence sources (PDB or ChEMBL, respectively). Protein homologs to flavonoid targets are colored in orange. In **B** and **D**, flavonoid targets are colored in black. Kinases homologous to flavonoid targets are colored in yellow. Other kinases are colored in green. (Color figure available online only.)

fect virtual screening results. ROCAU obtained with and without water in the query sites were highly similar.

Given that we aimed at selecting a small number of proteins for experimental testing, methods for virtual screening not only have to be sensitive and selective, i.e., with ROCAUs close to 1, but also have to achieve the early recognition of true targets. Bed-ROC, which increases the weight of true positives in the early fraction of the selection (here the 40 top-ranked entries), indicated that SiteAlign addressed the early recognition of flavonoid targets up to 11 times better than Shaper (**Table 4S**, Supporting Information), as also suggested by the initial slopes of ROC curves (○ **Fig. 3**). The analysis of ROCAU and Bed-ROC revealed that the ability to discriminate FBE and flavonoid targets from decoys also

depends on the query site. Virtual screening experiments using 2,3QD as a query indeed identified the highest number of true positives among top scorers, and exhibited the highest selectivity and sensitivity as well.

In a prospective screening exercise, only top-ranked proteins are submitted for experimental validation. We therefore analyzed hit lists obtained in the retrospective screening exercises. Hit lists were built assuming that similarity is significant if it differs by more than 2.5 standard deviations from the mean value of the distribution of scores. All distributions of scores were unimodal and could be approximated to the normal distribution with a slight skew on the tails (**Fig. 3S-6S**, Supporting Information). All 20 hit lists had relatively small and consistent sizes (between 18

and 45 using SiteAlign, and between 15 and 38 using Shaper, see ○ **Fig. 4**). A few nonselective flavonoid targets were found in several hit lists. Steroid receptors were present in all SiteAlign lists. These proteins have promiscuous binding sites [21]. For example, human peroxisome proliferator-activated receptor γ [22] was found in seven different hit lists (SiteAlign combined with CHI or 2,3QD, Shaper combined with CHI, DFR, or LAR). Carbonic anhydrase 2 [23] was also frequently encountered in hit lists.

Detailed analysis of each hit list showed that the composition was characteristic of each FBE screening. We especially observed FBE-specific flavonoid targets, thereby suggesting that there is not a single flavonoid imprint across the FBE family. Some flavonoid targets were found in only one FBE query. For example, human RAC-α serine/threonine protein kinase [24], human mitogen-activated protein kinase 1 [25], and human phosphatidylinositol 4,5-biphosphate 3-kinase catalytic subunit γ isoform [17] were only present in CHI hit lists. Many kinases, and more specifically serine/threonine protein kinases, were actually present in CHI hit lists, but not in other hit lists (○ **Fig. 4 B, D**). The flavonoid biological imprint embedded in CHI thus constituted a good bait to identify kinases which potentially bind flavonoids. CHI is involved in the formation of the isoflavan scaffold by catalyzing ring closure on chalcone substrates, and thus may retain an imprint of the complete isoflavan scaffold (**Fig. 1S**, Supporting Information). In addition, the active site composition in CHI differs from that in other FBEs. Especially CHI, like the kinases retrieved from the screening dataset, contains more charged residues than other FBEs (○ **Fig. 2**).

Considering that all the proteins homologous to flavonoid targets in the SiteAlign hit lists are putative true positives, the performance of retrospective screenings was probably underestimated. For example, proto-oncogene tyrosine-protein kinase Src from both humans and chickens [24] were present in the CHI hit list (1eyq), while only the human enzyme was marked as a flavonoid target. Androgen receptors from both humans and chimpanzees were identified in the CHI hit list (1eyq), while only the human enzyme was marked as a flavonoid target.

Finally, we asked the question "can similarity score be interpreted into common structural features?". To that end, we displayed the 3D alignment for a selection of similar pairs and observed that secondary structure elements are well superimposed although the protein global 3D structures are different. As shown on ○ **Fig. 5**, the active site of CHI is formed by α1 and α2 helices and a β1 three-stranded sheet and β2 strand. The similar binding site in RAC-α serine/threonine protein kinase is made of α3 and α4 helices that well superimpose to α1 and α2 in CHI. In addition, the β3 three-stranded sheet and α5 helix in the kinase well match β1 and β2 in CHI. Interestingly, secondary structure elements with a conserved position in space do not necessarily match secondary structure elements of the same type, as illustrated by the superimposition of the β2 strand from CHI to the α5 helix in the kinase.

In this retrospective study, we were able to use FBE as bait to retrieve flavonoid targets from a large set of ligandable proteins. Protein similarity based on shape (Shaper) returned hit lists with up to 14.7% of flavonoid targets. We demonstrated that shape-based similarity is not the method of choice, especially with promiscuous natural products in particular flavonoids. In this study, protein similarity based on molecular anchoring points (SiteAlign) returned hit lists containing up to 27% of flavonoid targets. SiteAlign successfully identified alternate domains of a helix and a β-sheet as possible equivalent anchoring points. The diversity of
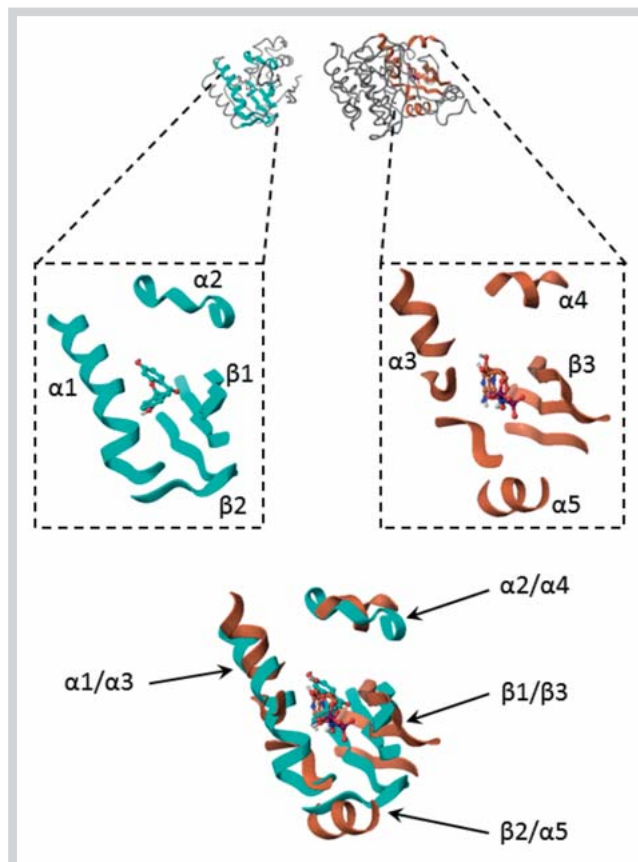

Fig. 5 Three-dimensional alignment of sites in chalcone isomerase and Ras-related C3 botulinum toxin substrate-α serine/threonine protein kinase. The active site of CHI (pdb code: 1fm7) is represented by cyan ribbons and the ATP-binding site of RAC-α serine/threonine protein kinase (pdb code: 4ekk) by orange ribbons. Ligands are rendered with a ball and stick. Sites were aligned using SiteAlign. (Color figure available online only.)

flavonoid targets and other proteins retrieved using different FBE queries suggested that the biological imprint gained during biosynthesis of natural products is unique to each biosynthetic enzyme (here, FBE) rather than there being a single unique flavonoid biological imprint across the FBE family. All FBE queries retrieved known flavonoid targets as well as a set of non-related flavonoid targets. This methodology promises to deliver non-related flavonoid targets as an enriched bioassay screening set.

Material and Methods
▼

Three-dimensional structures of protein binding sites

FBEs and the screening dataset were extracted from the 2012 release of the sc-PDB database [26]. The sc-PDB provides an all-atom description of complexes between a small molecular weight ligand and a ligandable protein, which includes all protein chains, metal ion(s), cofactor(s), and water molecule(s) (establishing at least two hydrogen bonds with the protein chains) in the vicinity of the ligand. For each protein, the binding site was defined as all protein residues delimiting the cavity detected using Volsite [9] and with at least one heavy atom distant from less than 6.5 Å from any ligand heavy atom. Last, we verified that the FBE active site was consistent with the amino acid sequence of the native protein as described in the UniProt database [10].
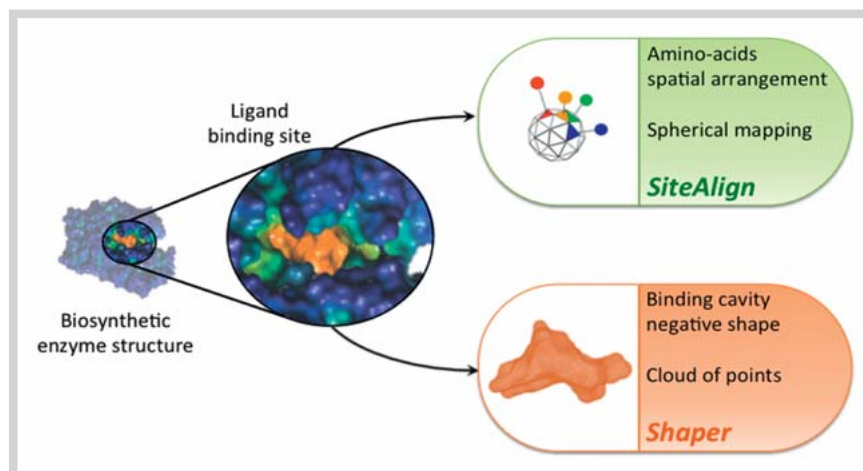
Fig. 6 Principle of protein binding sites comparison in SiteAlign and Shaper. (Color figure available online only.)

Binding site comparison

Site similarity was evaluated using two programs based on different methods, SiteAlign [18] and Shaper [9] (○ **Fig. 6**). Briefly, SiteAlign represents a binding site with an 80-triangle polyhedron centered on the protein cavity. Physicochemical properties of binding site amino acids are projected onto triangles of the polyhedron (cofactors, metal ions, and water molecules are ignored). Null property is assigned to triangles not hit by the projection of an amino acid. Binding sites are aligned by optimizing the superimposition of two polyhedrons for the best match of physicochemical properties. SiteAlign quantifies site similarity using two distances, whether considering all matched triangles ($D1$ score) or only matched triangles with non-null properties in the two polyhedrons ($D2$ score).

In the present study, the $D1$ score was used as a filter; two sites were dissimilar if $D1$ was lower than 0.6. The $D2$ score was used to rank solutions.

Shaper represents the negative image of a binding site, including amino acids, cofactor(s), and water molecule(s); 1.5 Å-spaced grid points filling the cavity are annotated with pharmacophoric properties of the nearest protein atoms. Binding sites are aligned by maximizing the geometric overlap of grids. Shaper quantifies site similarity by computing the proportion in the query site of the grid points with position and properties common to that in the compared site (*RefTversky* score).

Virtual screening

FBE active sites were compared to all the 8077 entries of the sc-PDB using Shaper and SiteAlign. Each screening experiment yielded a ranked list of 8076 binding sites, sorted by decreasing similarity to the query. For a given query, a hit list was obtained by selecting all proteins with at least one copy having a similarity score better than the mean of the distribution plus 2.5 standard deviations.

ROCAUs were computed using the package *pROC* [27] in R. Bed-ROC values were computed using the package *enrichvs* in R. The alpha coefficient for Bed-ROC was set to 200.

Supporting information

Tables showing the similarity between active sites of FBEs, sc-PDB proteins in a complex with a flavonoid, proteins with a micromolar or better affinity for flavonoids, as well as ROCAU and Bed-ROC values are available as Supporting Information. Also, figures displaying the biosynthetic reactions catalyzed by FBEs,

ROC curves for site comparison using Shaper, distribution of SiteAlign distances, as well as SiteAlign score and Shaper similarity score distributions can be found in this section.

Conflict of Interest
▼
The authors declare no conflict of interest.

References
1 *Demain AL, Fang A.* The natural functions of secondary metabolites. Adv Biochem Eng Biotechnol 2000; 69: 1–39
2 *Newman DJ, Cragg GM.* Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod 2012; 75: 311–335
3 *Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, Lavan P, Weber E, Doak AK, Cote S, Shoichet BK, Urban L.* Large-scale prediction and testing of drug activity on side-effect targets. Nature 2012; 486: 361–367
4 *Kellenberger E, Schalon C, Rognan D.* How to measure the similiarty between protein ligand-binding sites. Curr Comput Aided Drug Des 2008; 4: 209–220
5 *McArdle BM, Campitelli MR, Quinn RJ.* A common protein fold topology shared by flavonoid biosynthetic enzymes and therapeutic targets. J Nat Prod 2006; 69: 14–17
6 *Kellenberger E, Hofmann A, Quinn RJ.* Similar interactions of natural products with biosynthetic enzymes and therapeutic targets could explain why nature produces such a large proportion of existing drugs. Nat Prod Rep 2011; 28: 1483–1492
7 *Gutmanas A, Alhroub Y, Battle GM, Berrisford JM, Bochet E, Conroy MJ, Dana JM, Fernandez Montecelo MA, van Ginkel G, Gore SP, Haslam P, Hatherley R, Hendrickx PM, Hirshberg M, Lagerstedt I, Mir S, Mukhopadhyay A, Oldfield TJ, Patwardhan A, Rinaldi L, Sahni G, Sanz-Garcia E, Sen S, Slowley RA, Velankar S, Wainwright ME, Kleywegt GJ.* PDBe: Protein Data Bank in Europe. Nucleic Acids Res 2014; 42: D285–D291
8 *Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A, Shearer AG, Travers M, Weerasinghe D, Zhang P, Karp PD.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 2012; 40: D742–D753
9 *Desaphy J, Azdimousa K, Kellenberger E, Rognan D.* Comparison and druggability prediction of protein-ligand binding sites from pharmacophore-annotated cavity shapes. J Chem Inf Model 2012; 52: 2287–2299

10 *Consortium TU.* Activities at the Universal Protein Resource (UniProt). Nucleic Acids Res 2014; 42: D191–D198

11 *Bairoch A, Apweiler R.* The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 2000; 28: 45–48

12 *Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, Kruger FA, Light Y, Mak L, McGlinchey S, Nowotka M, Papadatos G, Santos R, Overington JP.* The ChEMBL bioactivity database: an update. Nucleic Acids Res 2014; 42: D1083–D1090

13 *Sak K.* Site-specific anticancer effects of dietary flavonoid quercetin. Nutr Cancer 2014; 66: 177–193

14 *Peer WA, Murphy AS.* The science of flavonoids. In: Grotewold E, editor. Flavonoids as signal molecules: targets of flavonoid action. New York: Springer; 2006: 239–268

15 *Lu X, Jung J, Cho HJ, Lim DY, Lee HS, Chun HS, Kwon DY, Park JH.* Fisetin inhibits the activities of cyclin-dependent kinases leading to cell cycle arrest in HT-29 human colon cancer cells. J Nutr 2005; 135: 2884–2890

16 *Havsteen BH.* The biochemistry and medical significance of the flavonoids. Pharmacol Ther 2002; 96: 67–202

17 *Walker EH, Pacold ME, Perisic O, Stephens L, Hawkins PT, Wymann MP, Williams RL.* Structural determinants of phosphoinositide 3-kinase inhibition by wortmannin, LY294002, quercetin, myricetin, and staurosporine. Mol Cell 2000; 6: 909–919

18 *Schalon C, Surgand JS, Kellenberger E, Rognan D.* A simple and fuzzy method to align and compare druggable ligand-binding sites. Proteins 2008; 71: 1755–1778

19 *Swets JA, Dawes RM, Monahan J.* Better decisions through science. Sci Am 2000; 283: 82–87

20 *Hawkins PC, Warren GL, Skillman AG, Nicholls A.* How to do an evaluation: pitfalls and traps. J Comput Aided Mol Des 2008; 22: 179–190

21 *Sturm N, Desaphy J, Quinn RJ, Rognan D, Kellenberger E.* Structural insights into the molecular basis of the ligand promiscuity. J Chem Inf Model 2012; 52: 2410–2421

22 *Puhl AC, Bernardes A, Silveira RL, Yuan J, Campos JL, Saidemberg DM, Palma MS, Cvoro A, Ayers SD, Webb P, Reinach PS, Skaf MS, Polikarpov I.* Mode of peroxisome proliferator-activated receptor gamma activation by luteolin. Mol Pharmacol 2012; 81: 788–799

23 *Ekinci D, Karagoz L, Ekinci D, Senturk M, Supuran CT.* Carbonic anhydrase inhibitors: *in vitro* inhibition of alpha isoforms (hCA I, hCA II, bCA III, hCA IV) by flavonoids. J Enzyme Inhib Med Chem 2013; 28: 283–288

24 *El Amrani M, Lai D, Debbab A, Aly AH, Siems K, Seidel C, Schnekenburger M, Gaigneaux A, Diederich M, Feger D, Lin W, Proksch P.* Protein kinase and HDAC inhibitors from the endophytic fungus *Epicoccum nigrum.* J Nat Prod 2014; 77: 49–56

25 *Tasdemir D, Mallon R, Greenstein M, Feldberg LR, Kim SC, Collins K, Wojciechowicz D, Mangalindan GC, Concepcion GP, Harper MK, Ireland CM.* Aldisine alkaloids from the Philippine sponge *Stylissa massa* are potent inhibitors of mitogen-activated protein kinase kinase-1 (MEK-1). J Med Chem 2002; 45: 529–532

26 *Desaphy J, Bret G, Rognan D, Kellenberger E.* sc-PDB: a 3D-database of ligandable binding sites – 10 years on. Nucleic Acids Res 2015; 43: D399–D404

27 *Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011; 12: 77–84