LIBRA-WA: a web application for ligand binding site detection and protein function recognition
Daniele Toti 1, Le Viet Hung 2, Valentina Tortosa 1, Valentina Brandi 1 and Fabio Polticelli 1,3.*

1Department of Sciences, University of Roma Tre, 00146 Rome, Italy
2Department of Science and Technology, Nguyen Tat Thanh University, Ho chi Minh City, Vietnam
3National Institute of Nuclear Physics, Roma Tre Section, 00146 Rome, Italy

Received on XXXX; revised on XXXX; accepted on XXXX

1 SUPPLEMENTARY METHODOLOGY
1.1 Clusterization of ligands according to their mutual similarity
The ligands stored in LIBRA-WA’s binding sites database have been clustered according to their mutual similarity, as computed by an automatic procedure that was meant to compare their SMILES representation. For this purpose, the SMILES information was retrieved from the Protein Data Bank in Europe5 via a corresponding crawling mechanism, and a similarity threshold (set to 0.7 after a fine-tuning process) was placed for generating the clusters. The actual similarity score was computed by using a combination of Levenshtein and Jaccard distances applied to the SMILES strings. The clusters’ semantics is as follows: each ligand has been placed into a cluster so that its similarity score with all of the other members of the cluster is above or equal the given threshold; the ligand with the highest average similarity score with the other members of the cluster has been dubbed the “core” ligand of such a cluster. Each ligand is thus present within one cluster at most. Clusters can be freely browsed by LIBRA-WA users via their corresponding web page, where the ligands making up each cluster are shown.

1.2 Extension and refinements of the alignment output and the scoring mechanism
Results produced by LIBRA-WA have been enriched with respect to LIBRA with additional information and a new scoring mechanism. Specifically, each binding site alignment record displayed within the results of an alignment process now features the information related to the cluster the corresponding ligand falls into (if any), including the core ligand of the cluster, its total size and the number of records within the result page whose ligand falls into that cluster. Alternatively, if the ligand from the record is not included in any cluster, the ligands from other records that are most similar to the considered ligand are shown. Along with the cluster information, a new, refined score mechanism has been implemented to rank more accurately the alignment records. This new score (S\text{new}) combines the contributions given by the aligned binding site’s clique size (C, corresponding to the number of matching residues between the input protein and the target binding site), RMSD value (R) and the relative size of the cluster containing the ligand (K), each normalized according to the following formulas:

\[
C_n = \frac{C - C_{\text{min}}}{C_{\text{max}} - C_{\text{min}}}
\]

\[
R_n = \frac{1 - (R - R_{\text{min}})}{R_{\text{max}} - R_{\text{min}}}
\]

\[
K_n = \frac{K - K_{\text{min}}}{K_{\text{max}} - K_{\text{min}}}
\]

where \text{min} and \text{max} indicate the minimum and maximum value of the respective parameter observed in the considered alignment records.

These normalized values are then weighted and linearly combined according to the following formula:

\[
S_{\text{new}} = w_C \times C_n + w_R \times R_n + w_K \times K_n
\]

where \(w_C = 0.5, w_R = 0.3\) and \(w_K = 0.2\). According to its respective value for \(S_{\text{new}}\), each alignment record shows a “Confidence” indicator for the quality of the alignment itself, ranging from “green” (\(S_{\text{new}} > 0.7\)), to “yellow” (0.4 <= \(S_{\text{new}} < 0.7\)), up to “red” (\(S_{\text{new}} < 0.4\)).

1.3 Identification of ligand binding sites across different protein subunits
As a result of the extensive experimentation carried out on LIBRA (Viet Hung et al., 2015), a number of cases were detected where the program could not identify the correct ligand binding site for the input protein, due to the fact that such a site included residues belonging to different protein subunits or “chains”. As a matter of fact, the original algorithm was able to identify only binding sites whose residues belonged to the same chain. In order to overcome this weakness of the application, a mechanism has been devised for identifying binding sites across different chains, by considering, for each residue of a given chain, its neighbours from different chains within a distance threshold. Since a naive iterative approach for performing this computation would obviously degrade the efficiency of the algorithm, residues have been thus indexed via a kind of r-dimensional indexes called KD-Trees, by using the three-dimensional coordinates of their geometric center as their index key. Then, the identification of their neighbours was only a matter of computing the respective Euclidean distances between residues belonging to different chains by applying an efficient nearest-neighbour algorithm, as implemented by (Levy et al., 2015). This resulted in an extremely fast computation that leaves the original performance of LIBRA’s algorithm virtually unaffected, while increasing its accuracy for such critical cases.

1.4 Additional features and improvements to user experience
Aside from the scientific advancements mentioned in the previous paragraphs, whose purpose was to improve the effectiveness of the alignment process, LIBRA-WA provides the users with an improved user experience and additional features as described below.

1.4.1 User account and personal panel As a web application (sharing the same architectural framework of (Atzeni et al., 2011a,b; Toti et al., 2012)), LIBRA-WA can be freely accessed and used by any web user without

*To whom correspondence should be addressed

http://www.ebi.ac.uk/pdbe-srv/pdbechem/
Fig. 1. Screenshot of a user’s personal panel in LIBRA-WA, where recognition jobs scheduled, running or completed are shown. For the user’s convenience, hovering the mouse over a recognition job displays a summary of the input parameters used when launching it, as shown in the case of the third result from the top. Once available, each result can be browsed in greater detail (“Show”), exported in the LIBRA/LIBRA+ .alg desktop file format (“Export”) or deleted altogether (“Delete”). From this panel, a user can access the page where recognitions can be launched (“Recognize the function of a given protein”), as well as browse the ligand clusters generated from the currently available database (“Display ligand clusters”).

1.4.2 Backward compatibility with LIBRA desktop application and introduction of LIBRA+

The results produced by LIBRA-WA can be easily exported as LIBRA recognition files (.alg) for offline viewing via the desktop application. Additionally, a new version of the desktop application, called LIBRA+, has been released in order to include and display the clusters information and the new scoring mechanism for each of the recognition files: such information is also retroactively available for results produced with the previous version of the desktop application when opened with LIBRA+.
Fig. 2. Screenshot of the panel showing the detailed results for a recognition job. (1) The input protein’s PDB code is reported and the possibility of exporting the results in LIBRA+ .alg file format is provided. (2) The values of the input parameters used for the given execution are recapped. (3) Filters are in place for the user’s convenience when browsing the results obtained; results can be filtered by known protein, ligand and “core” ligand of a cluster. (4) The results are shown in a table which can be re-ordered as needed by clicking on each column’s header. Further information is displayed by hovering the mouse over the known protein and ligand, such as the known protein’s PDB ID and ligand’s full name, respectively, as well as over the “Cluster” information, where all the records falling into the same clusters are shown (as depicted in the screenshot). For each record, via the Jmol HTML 5 plug-in, three-dimensional representations can be visualized, either showing (i) the input protein with its aligned residues highlighted, upon which the corresponding active site from the known protein is superimposed by means of rototranslations of its residues (as exemplified in the screenshot), or (ii) the known protein where the residues from its binding/active site are highlighted instead. Additional information on the aligned protein from each record, including a reference and PubMed ID of the related literature describing the protein’s structure, can be found by clicking on “Info”, which opens a corresponding pop-up window. Coloured circles in the right side of results panel represent the “Confidence” indicator as detailed in the text, with green circles characterizing high-confidence results.
2 SUPPLEMENTARY RESULTS

The effectiveness of LIBRA-WA, and in particular the performance improvement with respect to LIBRA due to the novel scoring mechanism, has been evaluated using the LigaSite set of 373 apo-proteins as a test set. As shown in Table 1 LIBRA-WA finds the biologically relevant ligand/binding site in ~96% of the cases. More importantly, the correct ligand/binding site ranks first in almost all of these cases. In detail, the correct hit is a substrate/substrate analog in 35% of the cases, a specific inhibitor in 16% of the cases, a product/product analog in 13% of the cases, a cofactor/cofactor analog in 13% of the cases, a specific ligand other than the above mentioned categories in 15% of the cases and a transition state analog/reaction intermediate in 3% of the cases. It must be remarked that this performance is in many cases independent from the sequence/structure similarity between the input protein and the holo protein from which the known binding site has been extracted. As already mentioned in the main text, one example of this performance is *E. coli* adenylate kinase (apoprotein PDB code 4AKE) which, upon ADP binding undergoes a large conformational change. As a consequence, LIBRA-WA does not identify the ADP binding site of the corresponding holo protein (PDB code 2ECK) as a correct match. Nonetheless, it correctly identifies the ADP binding site in the input apo protein by virtue of the local, structural similarity with the ADP binding site of the human kinesin-8 motor domain (PDB code 3LRE) which shares a mere 11% sequence identity with the *E. coli* adenylate kinase and practically no significant similarity in the overall protein topology (Figure 3). This observation suggests the possibility to exploit the ligand binding sites database, developed to be used with LIBRA-WA, to analyze if there are common, local structural determinants which govern the binding of specific ligands to different proteins, a possibility which will be explored as one of the future perspectives of this work. Detailed results of the tests performed on the LigaSite set are reported in Table 1. Even removing from the database the holo-proteins present in the LigaSite set, the application still performs fairly well. In fact, LIBRA-WA still identifies a biologically relevant ligand in 88% of the cases, with the correct ligand ranking first in 80% of the cases (Table 4).

For comparative purposes, LIBRA-WA, SiteSeer (Laskowski et al., 2005a,b) and the COACH meta-server (Yang et al., 2013; Roy et al., 2012; Brylinski and Skolnick, 2008; Capra et al., 2009) have been tested on a set of 30 apo-proteins randomly chosen from LigaSite. As it can be seen from Table 2, LIBRA-WA finds a correct solution in 97% of the cases (29 out of 30) as opposed to a success rate of 87% for SiteSeer and 90% for COACH meta-server. Particularly interesting is LIBRA-WA’s performance in comparison to COACH meta-server as the latter incorporates the results of several different sequence-based and structure-based algorithms (TM-Site, S-Site, COFACTOR, FINDSITE and ConCavity), while LIBRA-WA is, by design, exclusively structure-based; incidentally, COACH meta-server was ranked as the best method in the weekly CAMEO ligand Binding Site Prediction Experiments (Haas et al., 2013). This observation has prompted a more extensive comparison of LIBRA-WA’s performance with respect to COACH. For this purpose, COACH’s performance has been evaluated on the entire LigaSite test set. The results, shown in Table 3, indicate that COACH performs slightly better than LIBRA-WA as it finds a biologically relevant ligand in 98% of the cases (versus the 96% success rate of LIBRA-WA). However, the correct result found by COACH ranks first only in 90% of the cases (versus the 94% performance of LIBRA-WA). This is a very good result for LIBRA-WA, given its exclusively structure-based approach and the presence of only small molecules and ions in its database, as opposed to COACH’s

![Apo vs holo](image1)

![Holo vs hit](image2)

![Holo/hit ADP site](image3)

**Fig. 3.** Left panel. Superimposition of the *E. coli* adenylate kinase in the apo (PDB code 4AKE) and the holo ADP-bound (PDB code 2ECK) form. Note the large conformational change occurring upon ADP binding. Central panel. Superimposition of the ADP-bound structures of the *E. coli* adenylate kinase and the human kinesin-8 motor domain (PDB code 3LRE). Note the lack of structural similarity between the two proteins, paralleled by a lack of significant sequence identity (see text for details). Right panel. Detail of the ADP binding site of the two proteins. Note the good superimposition of the ADP molecules and the local structural similarity of a portion of the α helix and the connected loop hosting the ADP binding site in both proteins (circled in red).
database which contains also peptides and nucleic acids. In the light of these extensive tests, it can be stated that LIBRA-WA achieves a performance which is comparable to that of the COACH meta-server.
Table 1: LIBRA-WA's ligand/binding site recognition performance on the LigaSite non-redundant data-set containing 373 apo-proteins for which a corresponding holo structure is available (Desailly et al., 2008). APO and HOLO Protein ID are the PDB codes of the corresponding three-dimensional structures. Ligand ID is the ligand name abbreviation used in the Protein Data Bank.

| APO Protein ID | HOLO Protein ID | Ligand ID | Rank | rmsd (Å) | Ligand details       |
|----------------|-----------------|-----------|------|----------|----------------------|
| 180L           | 186L            | N4B       | 1st  | 0.485    | Specific ligand analog |
| 1A4U           | 1B2L            | NDC       | 1st  | 0.664    | Cofactor analog       |
| 1ADE           | 1HOP            | GCP       | 1st  | 1.023    | Substrate analog      |
| 1AK1           | 1C9E            | MP1       | 1st  | 0.89     | Substrate analog      |
| 1AKZ           | 3FCK            | FCK       | 1st  | 0.593    | Specific inhibitor    |
| 1ARB           | 4NSV            | 2OY       | 1st  | 0.324    | Specific inhibitor    |
| 1ARL           | 3FVL            | BHK       | 1st  | 0.533    | Specific inhibitor    |
| 1BS6           | 3UEX            | STE       | 1st  | 0.791    | Specific ligand       |
| 1BS8           | 1B8U            | NAD       | 1st  | 0.558    | Cofactor              |
| 1BD9           | 2B7A            | OPE       | 1st  | 0.648    | Natural ligand analog |
| 1BK7           | 1UCD            | USP       | 1st  | 0.456    | Reaction product      |
| 1BKZ           | 4UW6            | VV7       | 1st  | 0.512    | Natural ligand        |
| 1BQ6           | 2MAN            | MAN       | 1st  | 0.545    | Reaction product      |
| 1BQ7           | 1DAG            | ACP       | 1st  | 0.857    | Substrate analog      |
| 1C48           | 1CQF            | GAL       | 1st  | 0.69     | Natural ligand        |
| 1C5H           | 2B46            | XYP       | 1st  | 0.434    | Reaction product      |
| 1CE0           | 1CEN            | BGC       | 1st  | 0.672    | Substrate analog      |
| 1CEX           | 1OXM            | TC4       | 1st  | 0.484    | Covalent inhibitor    |
| 1CPF           | 2DC7            | 042       | 1st  | 0.654    | Covalent inhibitor    |
| 1CRW           | 2B4R            | NAD       | 1st  | 0.729    | Cofactor              |
| 1CWY           | 1ESW            | ACR       | 1st  | 0.543    | Specific inhibitor    |
| 1DCO           | 1DCP            | HBI       | 1st  | 0.713    | Cofactor analog       |
| 1DHN           | 2NM3            | MPU       | 1st  | 0.622    | Substrate analog      |
| 1DQ0           | 4P9W            | R3M       | 1st  | 0.726    | Specific ligand analog|
| 1DUP           | 1DUD            | DUD       | 1st  | 0.402    | Substrate analog      |
| 1E4F           | 4A2B            | AGS       | 1st  | 0.817    | Substrate analog      |
| 1E5L           | 1E5Q            | NDP       | 1st  | 0.887    | Cofactor              |
| 1E8Y           | 3L13            | IZW       | 1st  | 0.582    | Specific inhibitor    |
| 1EDQ           | 1X6N            | AO3       | 1st  | 0.444    | Specific inhibitor    |
| 1EPA           | 1EPB            | REA       | 1st  | 0.711    | Natural ligand        |
| 1EWZ           | 2WG1            | PNM       | 1st  | 0.936    | Specific inhibitor    |
| 1EY0           | 1A2T            | THP       | 1st  | 0.937    | Specific inhibitor    |
| 1EZJ           | 1EYR            | CDP       | 1st  | 0.847    | Substrate analog      |
| 1F14           | 1F17            | NAI       | 1st  | 0.409    | Cofactor              |
| 1F1S           | 2BRP            | SIE       | 1st  | 0.509    | Specific inhibitor    |
| 1F2V           | 1I1H            | COJ       | 1st  | 0.665    | Reaction product      |
| 1F41           | 1DVS            | STL       | 1st  | 0.593    | Specific inhibitor    |
| 1FSZ           | 4IMF            | SI3       | 1st  | 0.723    | Substrate            |
| 1FCQ           | 1FCV            | NAG       | 1st  | 0.7      | Substrate fragment    |
| 1FGB           | 1JOY            | A32       | 1st  | 0.883    | Specific inhibitor    |
| 1FO9           | 2APC            | UDM       | 1st  | 0.863    | Substrate analog      |
| 1FSF           | 1HOR            | AGP       | 1st  | 0.673    | Specific inhibitor    |
| 1FTF           | 1FTH            | A3P       | 1st  | 0.64     | Reaction product      |
| 1FTR           | 2FHJ            | MFN       | 1st  | 0.558    | Cofactor              |
| 1FWL           | 1H72            | ANP       | 1st  | 0.66     | Substrate analog      |
| 1G40           | 1RID            | IDS       | 1st  | 1.783    | Natural ligand        |
| PDB Code | Chain | Structure | pKa | Category                |
|----------|-------|------------|-----|-------------------------|
| 1G4E     | 1G4S  | TPS        | 1st | Reaction product        |
| 1G95     | 1G97  | UD1        | 1st | Reaction product        |
| 1GBS     | 1LSP  | BUL        | 1st | Specific inhibitor      |
| 1GCE     | 1GA0  | DVR        | 1st | Specific inhibitor      |
| 1GPS     | 1FXS  | NAP        | 1st | Reaction product        |
| 1G0U     | 1GOY  | 3GP        | 1st | Substrate analog        |
| 1GSH     | 1GSA  | GSH        | 1st | Reaction product        |
| 1GWK     | 1W8U  | BMA        | 1st | Natural ligand          |
| 1GY0     | 1OG1  | TAD        | 1st | Substrate analog        |
| 1HK9     | 3QO3  | ATP        | 1st | Natural ligand          |
| 1HKA     | 3UDV  | J1C        | 1st | Substrate analog        |
| 1HM5     | 4QFH  | G6P        | 1st | Reaction product        |
| 1HNK     | 1HNI  | MLC        | 1st | Substrate               |
| 1H01     | 1H04  | PXP        | 1st | Reaction product        |
| 1H7N     | 1H7L  | ATP        | 1st | Endogenous ligand       |
| 1IAD     | 1QJI  | PKF        | 1st | Specific inhibitor      |
| 1ILV     | 1JKK  | W04        | 1st | Specific inhibitor      |
| 1INL     | 1JQ3  | AAT        | 1st | Substrate analog        |
| 1JS5     | 1MXI  | SAH        | 1st | Cofactor                |
| 1JS8     | 1ISR  | GLA        | 1st | Endogenous ligand       |
| 1HCF     | 1ICG  | ANP        | 1st | Substrate analog        |
| 1JKS     | 1IG1  | ANP        | 1st | Substrate analog        |
| 1JXO     | 1JXM  | MPD        | 1st | Non-specific ligand     |
| 1JYK     | 1JYL  | CDC        | 1st | Reaction product        |
| 1K0M     | 1K0N  | GSH        | 1st | Substrate               |
| 1K3O     | 4H2   | L26        | 1st | Substrate analog        |
| 1K6A     | 1B3O  | XYP        | 1st | Reaction product        |
| 1KAM     | 1KAQ  | DND        | 1st | Substrate               |
| 1KF5     | 1QHC  | PUA        | 1st | Reaction product        |
| 1KHD     | 1KGZ  | PRP        | 1st | Substrate               |
| 1KN9     | 1B12  | 1PN        | 1st | Specific inhibitor      |
| 1KPA     | 1KPE  | ADW        | 1st | Substrate analog        |
| 1KWB     | 1EIR  | BPY        | 1st | Substrate               |
| 1L7D     | 1LXT  | NAD        | 1st | Substrate               |
| 1LBV     | 1LBZ  | FBP        | 1st | Substrate               |
| 1LCI     | 3RJX  | 923        | 1st | Specific Inhibitor      |
| 1LFL4    | 1LF3  | EH5        | 1st | Specific Inhibitor      |
| 1LTU     | 1LTZ  | HHI        | 1st | Cofactor                |
| 1M1Z     | 1MB9  | ATP        | 1st | Substrate               |
| 1MKB     | 4KEH  | 1z3        | 1st | Substrate analog        |
| 1MMI     | 3QSB  | 743        | 1st | Specific inhibitor      |
| 1MR7     | 1Kk4  | ACO        | 1st | Substrate               |
| 1MTZ     | 1MU0  | PHK        | 1st | Specific inhibitor      |
| 1MWK     | 4A61  | ANP        | 1st | Substrate analog        |
| 1MZL     | 1FK4  | STE        | 1st | Natural ligand          |
| 1N05     | 1N07  | FMN        | 1st | Reaction intermediate   |
| 1NDB     | 2H3P  | ACO        | 1st | Substrate               |
| 1NON     | 1XZ8  | 5GP        | 1st | Substrate analog        |
| 1NXM     | 2XL   | TRH        | 1st | Substrate analog        |
| 1O24     | 1KQ4  | FAD        | 1st | Cofactor                |
| 1OEI     | 3EAX  | LZP        | 1st | Specific inhibitor      |
| Code  | Code  | Name     | Code  | Code  | Name     |
|-------|-------|----------|-------|-------|----------|
| 1OFP  | 1OF6  | DTY      | 1OG7  | DTY   | Substrate analog |
| 1OGG  | 1OGO  | GLC      | 1OGM  | 1OGO  | Reaction product |
| 1OJQ  | 1OJZ  | NAD      | 1OLZ  | 3OLZ  | Sugar chain |
| 1O0I  | 2GTE  | VA       | 1O0V  | 3OWY  | Transition state analog |
| 1OXT  | 1OXY  | ANP      | 1O1X  | 1JCJ  | Reaction intermediate |
| 1P5H  | 1T3Z  | CAO      | 1P74  | 3PHI  | Reacation product |
| 1PDB  | 1BOZ  | NDP      | 1P1Q  | 1Q5J  | Substrate analog |
| 1PNG  | 1P1F  | NDG      | 1P24  | 1Q51  | Substrate analog |
| 1P3Q  | 1Q0J  | C2F      | 1P74  | 1Q51  | Substrate analog |
| 1QID  | 1H23  | E12      | 1QTO  | 1JIF  | Specific inhibitor |
| 1QTR  | 1X2B  | STX      | 1R0M  | 1XPY  | Substrate |
| 1R12  | 3ZWP  | AVU      | 1R0M  | 1XPY  | Specific inhibitor |
| 1R29  | 4CP4  | RBT      | 1R55  | 2HV9  | Specific inhibitor |
| 1R55  | 2HV9  | SFG      | 1R7Q  | 3HIXQ | Specific inhibitor |
| 1R7Q  | 2C7J  | IOD      | 1RKM  | 2RKX  | Non-specific ligand |
| 1R7Q  | 2C7J  | IOD      | 1RKM  | 2RKX  | Specific ligand |
| 1RZV  | 1RZU  | ADP      | 1S2L  | 1S2D  | Substrate analog |
| 1S7K  | 1S7L  | COA      | 1S7K  | 1S7L  | Substrate analog |
| 1SGK  | 1DDT  | AP4      | 1S7K  | 1S7L  | Substrate analog |
| 1SGK  | 1DDT  | AP4      | 1SGK  | 1DDT  | Substrate analog |
| 1SGZ  | 3I2S  | MV7      | 1SGZ  | 3I2S  | Specific inhibitor |
| 1S7K  | 1S7L  | COA      | 1S7K  | 1S7L  | Specific inhibitor |
| 1T2I  | 1T2I  | AS1      | 1T2I  | 1T2I  | Cofactor |
| 1T2I  | 1T2I  | AS1      | 1T2I  | 1T2I  | No significant results |
| 1T2I  | 1T2I  | AS1      | 1T2I  | 1T2I  | Specific ligand analog |
| 1T7B  | 1T7B  | OL4      | 1T7B  | 1T7B  | Non-specific ligand |
| 1T7E  | 1T7E  | A3S      | 1T7E  | 1T7E  | Substrate analog |
| 1T7M  | 1T7M  | AS1      | 1T7M  | 1T7M  | Substrate analog |
| 1TVN  | 1TVN  | DCB      | 1TW0  | 1TW0  | Non-specific ligand |
| 1TVN  | 1TVN  | DCB      | 1TVN  | 1TVN  | Non-specific ligand |
| 1TVN  | 1TVN  | DCB      | 1TVN  | 1TVN  | Non-specific ligand |
| 1TVN  | 1TVN  | DCB      | 1TVN  | 1TVN  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| 1TV0  | 1TV0  | 2AN      | 1TV0  | 1TV0  | Non-specific ligand |
| Code | Chain | Name   | Type               | Value | Description               |
|------|--------|--------|--------------------|-------|---------------------------|
| 1ULU | 2WYY   | NAD    | 1st                | 0.444 | Cofactor                  |
| 1UMF | 1UM0   | FMN    | 1st                | 0.465 | Cofactor                  |
| 1UXZ | 1UZ0   | BGC    | 1st                | 0.337 | Specific ligand           |
| 1UYL | 3TOZ   | ATP    | 1st                | 0.934 | Substrate                 |
| 1Vo8 | 2ZE9   | PD7    | 1st                | 0.335 | Reaction product          |
| 1VGZ | 2CX8   | SAH    | 1st                | 0.626 | Cofactor analog           |
| 1V77 | 2CZV   | ACY    | 1st                | 0.357 | Non-specific ligand       |
| 1VFJ | 1V9O   | ADP    | 1st                | 0.643 | Specific ligand           |
| 1VIC | 1H7H   | CDP    | 1st                | 1.172 | Substrate analog          |
| 1VJU | 3DWR   | 0PA    | 1st                | 0.671 | Substrate analog          |
| 1VPS | 1VPS   | SIA    | 1st                | 0.62  | Specific ligand           |
| 1WA9 | 4QVB   | F42    | 1st                | 0.741 | Cofactor                  |
| 1WG3 | 2Z08   | ATP    | 1st                | 0.987 | Specific ligand           |
| 1WNY | 1WNZ   | 2VA    | 1st                | 0.6   | Substrate analog          |
| 1WOS | 1WOP   | FFO    | 1st                | 0.551 | Substrate analog          |
| 1WS9 | 2D29   | FAD    | 1st                | 0.327 | Cofactor                  |
| 1WTJ | 2CWF   | NDP    | 1st                | 0.457 | Cofactor                  |
| 1WXF | 1EE1   | ATP    | 1st                | 0.556 | Substrate analog          |
| 1X56 | 1X55   | NIS    | 1st                | 0.713 | Substrate analog          |
| 1X7O | 1X7P   | SAM    | 1st                | 0.534 | Cofactor                  |
| 1X8F | 1X6U   | DO8    | 1st                | 0.457 | Reaction product          |
| 1XD4 | 1P4N   | UMA    | 1st                | 0.225 | Substrate                 |
| 1XK7 | 1XA4   | COA    | 1st                | 0.569 | Substrate analog          |
| 1XO6 | 1XNY   | I91    | 1st                | 0.548 | Substrate                 |
| 1XQO | 1XQP   | 8HG    | 1st                | 0.575 | Substrate analog          |
| 1XW2 | 2B4F   | XYP    | 1st                | 0.381 | Reaction product          |
| 1Y2Q | 3PD2   | A3S    | 1st                | 0.817 | Substrate analog          |
| 1Y2T | 2F3D   | NGA    | 1st                | 0.896 | Endogenous ligand         |
| 1YBT | 1YBU   | APC    | 1st                | 1.21  | Substrate analog          |
| 1YYY | 1YTM   | ATP    | 1st                | 0.994 | Substrate                 |
| 1Z7G | 4RAB   | 3L3    | 1st                | 0.872 | Specific inhibitor         |
| 1ZAH | 3TU9   | 5MM    | 1st                | 0.319 | Specific inhibitor         |
| 1ZCU | 1ZDF   | UPG    | 1st                | 0.305 | Substrate                 |
| 1ZNW | 1ZNX   | 5GP    | 1st                | 0.279 | Substrate                 |
| 1ZTY | 1ZU0   | CBS    | 1st                | 0.815 | Natural ligand            |
| 1ZUH | 3MUF   | ADP    | 1st                | 1.294 | Reaction product          |
| 2A6Z | 4RAB   | 3L3    | 1st                | 0.872 | Specific inhibitor         |
| 2A8F | 1LRI   | CLR    | 1st                | 0.897 | Specific ligand           |
| 2AD1 | 1BO6   | VO4    | 2nd               | 0.713 | Transition state analog    |
| 2AHF | 2AHG   | UCD    | 1st                | 0.617 | Substrate                 |
| 2AHU | 2AHV   | COA    | 1st                | 0.862 | Cofactor                  |
| 2AMJ | 2B3D   | FAD    | 1st                | 0.449 | Cofactor                  |
| 2B0I | 3F46   | I2C    | 1st                | 0.68  | Cofactor                  |
| 2B6P | 1YMG   | BNG    | 1st                | 0.152 | Non-specific ligand       |
| 2B7B | 3LDF   | SAH    | 1st                | 0.405 | Cofactor                  |
| 2B99 | 2B99   | RDL    | 1st                | 0.53  | Substrate analog inhibitor |
| 2BGT | 1J39   | UPG    | 1st                | 0.96  | Substrate                 |
| 2BJW | 2C96   | ATP    | 1st                | 0.593 | Substrate                 |
| 2BOE | 2BO7   | BGC    | 1st                | 0.267 | Reaction product          |
| 2CS1 | 3B2Q   | ATP    | 1st                | 0.592 | Reaction product          |
| 2CTL | 2ARU   | ATP    | 1st                | 0.555 | Substrate                 |
| 2CAR | 2J4E   |       |                    |       | No significant results    |
| Code  | Protein | Type                  | Score | Description               |
|-------|---------|-----------------------|-------|---------------------------|
| 2CHS  | 1COM    | PRE                   | 1°    | Reaction product          |
| 2CI1  | 2CI1    | KOR                   | 1°    | Specific inhibitor        |
| 2CW9  | 2DY9    | ADP                   | 1°    | Substrate                |
| 2CX5  | 2Z0X    | 5CA                   | 1°    | Substrate analog          |
| 2D5A  | 2D5A    | COA                   | 1°    | Natural ligand            |
| 2DHQ  | 3N76    | CA2                   | 1°    | Specific inhibitor        |
| 2DI6  | 2DFT    | H4B                   | 1°    | Substrate analog          |
| 2DKA  | 2DKC    | 16G                   | 1°    | Substrate                |
| 2DPS  | 2DPT    | PUY                   | 1°    | Substrate analog          |
| 2DQW  | 1EYE    | PMM                   | 1°    | Substrate analog          |
| 2EOC  | 2E3M    | NAP                   | 1°    | Substrate                |
| 2E9K  | 2E0N    | SAH                   | 1°    | Cofactor                 |
| 2E10  | 1WQW    | BT5                   | 1°    | Substrate analog          |
| 2E1V  | 2E1T    | MLC                   | 1°    | Substrate                |
| 2E3S  | 2E3O    | 16C                   | 1°    | Substrate                |
| 2E9R  | 2ED4    | FAD                   | 1°    | Substrate                |
| 2EX0  | 2HQ     | CSP                   | 1°    | Substrate analog          |
| 2F82  | 2F0A    | HMG                   | 1°    | Reaction product          |
| 2FT9  | 2F9W    | PAU                   | 1°    | Substrate                |
| 2FK7  | 3HA7    | B32                   | 1°    | Specific inhibitor        |
| 2FP8  | 2V91    | SS5                   | 1°    | Substrate                |
| 2FSF  | 2FSH    | ANP                   | 1°    | Substrate analog          |
| 2G67  | 2G25    | TZK                   | 1°    | Reaction intermediate analog |
| 2G95  | 3G8E    | IS1                   | 1°    | Specific inhibitor        |
| 2GFV  | 2GFX    | PMN                   | 1°    | Specific inhibitor        |
| 2GG4  | 1RF6    | S3P                   | 1°    | Substrate                |
| 2GQV  | 2P4T    | NAP                   | 4°    | Cofactor                 |
| 2GSF  | 4PSZ    | Q7M                   | 1°    | Specific inhibitor        |
| 2GT2  | 2GT4    | GDD                   | 1°    | Substrate                |
| 2GUB  | 4QDW    | LAI                   | 1°    | Reaction intermediate     |
| 2GWX  | 3GWX    | EPA                   | 1°    | Natural ligand            |
| 2GYY  | 2GZ1    | NAP                   | 1°    | Cofactor                 |
| 2HZ2  | 2A5K    | AZP                   | 1°    | Inhibitor                |
| 2HBJ  | 2HBL    | AMP                   | 1°    | Reaction product          |
| 2HIV  | 2HX     | ATP                   | 1°    | Substrate                |
| 2HK0  | 3VNL    | TAG                   | 1°    | Substrate                |
| 2HVM  | 1LLO    | AMI                   | 1°    | Cofactor                 |
| 2HY7  | 2Q6V    | UDP                   | 1°    | Substrate analog/reaction product |
| 2HZR  | 2HZQ    | STR                   | 1°    | Natural ligand            |
| 2IL4  | 2I4M    | P5D                   | 1°    | Substrate                |
| 2I0B  | 3A2Y    | TS5                   | 1°    | Substrate                |
| 2IRU  | 2IRX    | GTP                   | 1°    | Substrate                |
| 2J71  | 2J72    | GLC                   | 10°   | Natural ligand            |
| 2J8N  | 3DR8    | ACO                   | 1°    | Cofactor                 |
| 2JBR  | 2JBS    | FMN                   | 1°    | Substrate                |
| 2NWD  | 4XAD    | 3ZW                   | 1°    | Substrate analog          |
| 2NCX  | 3CJT    | SAM                   | 1°    | Cofactor                 |
| 209P  | 221S    | CTT                   | 1°    | Substrate                |
| 20AM  | 20A1    | FAD                   | 1°    | Cofactor                 |
| 2OPT  | 3B6A    | ZCT                   | 1°    | Specific inhibitor        |
| 2OU1  | 41GK    | PLM                   | 1°    | Natural ligand            |
| 2OVE  | 20VD    | DAO                   | 1°    | Specific ligand           |
| PDB Code | Chain | Type         | Z Score |
|----------|-------|--------------|---------|
| 2PAW     | 4HJY  | Specific inhibitor | 0.694   |
| 2PFK     | 1PKF  | Reaction product | 1.001   |
| 2PKF     | 4QJG  | Substrate analog | 0.471   |
| 2PPN     | 1BL4  | Specific ligand | 0.916   |
| 2QDR     | 2JGJ  | Reaction product | 0.762   |
| 2QGZ     | 1R3Q  | Reaction product | 0.796   |
| 2QBV     | 2W1A  | Transition state analog | 1.193   |
| 2QDK     | 1RXXU | Reaction product analog | 0.595   |
| 2QEV     | 2QEO  | Natural ligand | 0.412   |
| 2QSU     | 2OTG  | Substrate analog | 0.574   |
| 2QVL     | 2OV7  | Reaction product | 0.518   |
| 2QYS     | 2OW8  | Cofactor | 0.432   |
| 2RX6     | 2R68  | Reaction product | 0.24    |
| 2RG7     | 1AIJ  | Endogenous ligand analog | 2.363   |
| 2RDJ     | 3PHI  | Specific inhibitor | 0.248   |
| 2SGA     | 2QAA  | Reaction intermediate | 0.559   |
| 2SIL     | 2SIM  | Specific inhibitor | 0.179   |
| 2TS1     | 3TS1  | Reaction intermediate | 0.41    |
| 2UYO     | 2UQO  | Cofactor | 0.57    |
| 2V78     | 2VAR  | Substrate | 0.494   |
| 2VFB     | 2VFC  | Cofactor | 0.686   |
| 2VFY     | 2VFK  | Natural ligand | 0.904   |
| 2VQ4     | 4BNF  | Natural ligand | 0.731   |
| 2VUA     | 2UJ9  | Natural ligand | 0.197   |
| 2WN4     | 2WN6  | Cofactor | 0.768   |
| 2WVH     | 3OJ1  | Reaction intermediate | 0.457   |
| 2WZT     | 2WZM  | Cofactor analog | 0.928   |
| 2X5S     | 2XZS  | Reaction product | 0.829   |
| 2YW8     | 1GPM  | Product analog | 1.03    |
| 2YXF     | 1LDS  | Non-specific ligand | 0.9     |
| 2YYA     | 2YW2  | Substrate | 0.833   |
| 2YYY     | 2YYU  | Product analog | 0.481   |
| 2YZG     | 2ZNN  | Substrate analog | 0.944   |
| 2ZBS     | 1SQ9  | Natural ligand analog | 1.022   |
| 2ZCG     | 2A1   | Substrate | 0.767   |
| 2ZCO     | 3ADZ  | Reaction intermediate | 0.8     |
| 2ZGL     | 2ZGM  | Natural ligand analog | 0.398   |
| 2ZHY     | 2HJZ  | Substrate | 0.707   |
| 2ZJ8     | 2ZIA  | Substrate analog | 0.717   |
| 2ZTY     | 3Z93  | Specific inhibitor | 0.645   |
| 3AOY     | 3A0T  | Reaction product | 0.991   |
| 3ASQ     | 1CHW  | Substrate analog | 0.849   |
| 3AAP     | 3AAR  | Specific inhibitor | 0.793   |
| 3ADO     | 3ADP  | Cofactor | 0.632   |
| 3APP     | 1PPM  | Specific inhibitor | 0.66    |
| 3B3G     | 2Y3W  | Cofactor analog | 1.05    |
| 3B1A     | 3BZK  | Cofactor | 0.583   |
| 3B1M     | 1GHP  | Reaction intermediate | 0.481   |
| 3BTV     | 3BTS  | Specific ligand (activator) | 0.827   |
| 3BUE     | 3CAG  | Natural ligand | 0.855   |
| 3BYL     | 3BYN  | Substrate analog | 0.762   |
| 3C2E     | 3C2V  | Specific inhibitor | 0.725   |
| PDB Code | Description                  | Type         | Score | Class       |
|----------|------------------------------|--------------|-------|-------------|
| 3C8N     | 3BAY F42                     | Cofactor     | 0.656 | Specific ligand |
| 3CAF     | 1E0O SGN                     | Specific ligand | 1.313 | Specific ligand |
| 3CB6     | 3CB5 B3P                     | Cofactor     | 0.735 | Specific ligand |
| 3CGZ     | 3CGY RDC                     | Specific inhibitor | 0.849 | Specific inhibitor |
| 3COU     | 2XSQ IMP                     | Reaction product | 0.497 | Reaction product |
| 3CQ1     | No significant results       |              |       |             |
| 3CRM     | 3CRQ DPO                     | Reaction product | 0.502 | Reaction product |
| 3CSR     | 3CSZ NAG                     | Substrate analog | 0.249 | Substrate analog |
| 3CTB     | 1HLH SRL                     | Specific agonist | 0.816 | Specific agonist |
| 3DOO     | 3D4P NAD                     | Cofactor     | 0.603 | Cofactor |
| 3D95     | 3CWK REA                     | Endogenous ligand | 0.683 | Endogenous ligand |
| 3DRD     | 3DOD PLP                     | Cofactor     | 0.954 | Cofactor |
| 3DRE     | 3JU6 ANP                     | Substrate analog | 1.295 | Substrate analog |
| 3DUL     | 3DUW SAH                     | Cofactor     | 1.374 | Cofactor |
| 3E1S     | 2FWW ADP                     | Substrate analog | 1.859 | Substrate analog |
| 3E5B     | 3P0X ICT                     | Substrate     | 0.786 | Substrate |
| 3EEZ     | 3EY POP                      | Substrate     | 1.051 | Substrate |
| 3EK6     | 2BND UDP                     | Reaction product | 1.04  | Reaction product |
| 3ETF     | 3EFV NAD                     | Cofactor     | 0.362 | Cofactor |
| 3EX9     | 3B4P 3B4                     | Substrate analog | 0.528 | Substrate analog |
| 3EXR     | 3EXS 5RP                     | Product analog | 0.574 | Product analog |
| 3F1L     | 3F1K NAP                     | Cofactor     | 0.592 | Cofactor |
| 3F6F     | 3GH6 GSH                     | Substrate     | 0.217 | Substrate |
| 3FTD     | 2ZBR SPG                     | Cofactor analog | 1.578 | Cofactor analog |
| 3FV6     | 3FW5 ANP                     | Natural ligand analogous | 1.079 | Natural ligand analogous |
| 3G1S     | 3G1F 2OM                     | Substrate analog | 0.785 | Substrate analog |
| 3GBT     | 3L3 ATP                      | Substrate     | 0.711 | Substrate |
| 3GD0     | 3GD9 BGC                     | Substrate/product analog | 0.527 | Substrate/product analog |
| 3GLK     | 3GID S1A                     | Specific ligand | 0.483 | Specific ligand |
| 3GP6     | 3GPO APR                     | Specific ligand | 0.58  | Specific ligand |
| 3GQH     | 3GQK ATP                     | Specific ligand | 1.385 | Specific ligand |
| 3GSZ     | 3BR9 DEY                     | Specific inhibitor | 0.547 | Specific inhibitor |
| 3H2G     | 1GKK SEP                     | Specific ligand | 0.806 | Specific ligand |
| 3H38     | 3H39 ATP                     | Substrate     | 0.883 | Substrate |
| 3H49     | 3IN1 ADP                     | Reaction product | 0.624 | Reaction product |
| 3H71     | 2YA6 DAN                     | Specific inhibitor | 0.604 | Specific inhibitor |
| 3HHH     | 2DKV MES                     | Substrate mimic | 0.938 | Substrate mimic |
| 3H15     | 3HIV C2X                     | Transition state analog | 0.613 | Transition state analog |
| 3H44     | 3HY UTP                      | Substrate     | 0.874 | Substrate |
| 3H4X     | 3HP8 SUC                     | Specific ligand | 0.546 | Specific ligand |
| 3DC      | 3V0N 3GW                     | Specific inhibitor | 0.831 | Specific inhibitor |
| 3U2I     | 1QLL TDA                     | Specific inhibitor | 0.688 | Specific inhibitor |
| 3HS      | 3WIC GNP                     | Substrate analog | 1.196 | Substrate analog |
| 3K8      | 4TYO 39X                     | Specific inhibitor | 0.419 | Specific inhibitor |
| 3LY      | No significant results       |              |       |             |
| 3JTY     | 3JTV PSJ                     | Substrate     | 0.523 | Substrate |
| 3U1J     | 3VM ZPR                      | Specific inhibitor | 0.955 | Specific inhibitor |
| 3JYL     | 3YN NDP                      | Cofactor     | 0.641 | Cofactor |
| 3K0M     | 43A 67Z                      | Specific inhibitor | 0.637 | Specific inhibitor |
| 3KAL     | 3KAL HGS                     | Reaction product | 0.633 | Reaction product |
| 3KD1     | 3NR0 PYV                     | Natural ligand mimic | 0.368 | Natural ligand mimic |
Table 2: Comparison of LIBRA-WA, SiteSeek and COACH on their respective ligand/binding site recognition performance, applied on a set of ligand-free proteins randomly chosen from the LigASite. Please note that only the first result provided by each respective system is reported. Wrong results (wrong target protein, ligand not biologically relevant or no results) are highlighted in bold font. Superscript notes for ligands correspond to the following descriptions. 1: Substrate analog; 2: Reaction product; 3: Covalent inhibitor; 4: Cofactor; 5: Specific inhibitor; 6: Substrate; 7: Reaction product analog; 8: Natural ligand analog; 9: Cofactor analog; 10: Reaction intermediate; 11: Natural ligand; 12: Non-specific ligand.

| PDB code | Experimental function and LigASite ligand | LIBRA-WA Predicted function (PDB code) and ligand (ligand ID) | SiteSeek | COACH server |
|----------|------------------------------------------|---------------------------------------------------------------|----------|--------------|
| 1AK1     | Ferrochelatase N-methylmesoporphyrin     | Ferrochelatase (1C9E) N-methylmesoporphyrin (MP1) \(^1\) | Ferrochelatase (1C9E) N-methylmesoporphyrin (MP1) \(^1\) | Ferrochelatase (2HRE) N-methylmesoporphyrin IX (PP9) \(^6\) |
| 1BK7     | Ribonuclease MC1 Uridine-5’-monophosphate | Ribonuclease MC1 (1UCD) Uridine-5’-monophosphate (USP) \(^2\) | Ribonuclease NW (1IYB) Guanosine-5’-monophosphate (5GP) \(^2\) | Ribonuclease NW (1IYB) Guanosine-5’-monophosphate (5GP) \(^2\) |
| 1CEX     | Cutinase Butyl-phosphinic acid 2,3-bis-butylcarbamoyloxy-propyl ester group | Cutinase (1OXM) Butyl-phosphinic acid 2,3-bis-butylcarbamoyloxy-propyl ester (TC4) \(^3\) | Cutinase (1XZM) N-undecanoylphosphonate methyl ester group (DPE) \(^3\) | Cutinase (1XZM) N-undecanoylphosphonate methyl ester group (DPE) \(^3\) |
| 1CRW     | d-Glyceraldehyde-3-phosphate dehydrogenase | Glyceraldehyde-3-phosphate dehydrogenase (2B4R) Nicotinamide-adenine-dinucleotide (NAD) \(^4\) | d-Glyceraldehyde-3-phosphate dehydrogenase (1DSS) Sulfate ion (SO4) \(^1\) \(^2\) | Glyceraldehyde-3-phosphate dehydrogenase (4WNC) Nicotinamide-adenine-dinucleotide (NAD) \(^4\) |
| 1EDQ     | Chitinase A Allosamizoline | Chitinase A mutant (1X6N) Allosamidin (AO3) \(^3\) | W167A Chitinase A mutant (1X6N) Allosamidin (AO3) \(^3\) | Chitinase 1 (3W1) N-acetyl-D-glucosamine (NAG) \(^7\) |
| 1EWZ     | Beta-Lactamase OXA-10 (1R)-2-(1-carboxy-2-hydroxy-2-methyl-propyl)-5,5-dimethyl-thiazolidine-4-carboxylic acid | OXA-10 (2WGI) Open form - penicillin g (PNM) \(^5\) | OXA-10 (1FOF) Cobalt (II) ion (Co) \(^1\) \(^11\) | Beta-lactamase OXA-24 8 (3FZC) (2S,3R)-4-(2-aminoethylcarbamoyloxy)-2-[(2-methanoylindolizin-3-yl)amino]-3-methyl-5-sulfino-butanoic acid (AMX) \(^7\) |
| 1GOU     | Ribonuclease Guanosine-3’-monophosphate | Ribonuclease bi (1GOY) Guanosine-3’-monophosphate (3GP) \(^1\) | Ribonuclease bi (1GOY) Sulfate ion (SO4) \(^1\) \(^12\) | Ribonuclease bi (1GOY) Guanosine-3’-monophosphate (3GP) \(^1\) |
| 1MR7     | Streptogramin A Acetyltransferase Acetyl-coenzyme A | Var(D) (1KK4) Acetyl coenzyme A (ACO) \(^6\) | Var(D) (1KK4) Acetyl coenzyme A (ACO) \(^6\) | Hexapetidase-repeat containing-acetyltransferase (3NZ2) Acetyl-coenzyme A (ACO) \(^7\) |
| 1MTZ     | Proline iminopeptidase 3-amino-1-chloro-4-phenylbutanol-2-yl | Tricorn interacting factor F1 (1MU1) (2R,3S)-3-amino-1-chloro-4-phenylbutan-2-ol (PHK) \(^3\) | Tricorn F1-mutant c245q (1XRR) Proline (PRO) \(^7\) | Gamma lactamase (1HL7) 3a,4,7,7a-tetrahydro-benzo [1,3] dioxol-2-one (BD1) \(^1\) \(^2\) |
| 1NDB | Carnitine acetyltransferase | Carnitine acetyltransferase (2H3P) | Acetyl coenzyme A (ACO) | Carnitine acetyltransferase isoform 2 (1SSO) | Carnitine (1SSO) |
|------|---------------------------|----------------------------------|------------------------|---------------------------------|----------------|
| 1R0M | N-acetylglucosamine 2-acetyl-1-glucosamine | N-acetylglucosamine acetylase (1LXY) | N 2'-acetyl-1-glucosamine (NLQ) | Nacylglucosamine acetylase (2FKR) | N 2'-acyetl-1-glucosamine (NLQ) |
|      |                           |                                  |                        | Mandelate racemase/muconate lactonizing enzyme family protein (3TOY) | Calcium ion (Ca) |
| 1RZV | Glycogen synthase 1       | Glycogen synthase 1 (1RZU)       | Adenosine-5'-diphosphate (ADP) | Glycogen synthase 1 (1RZU) | Adenosine-5'-diphosphate (ADP) |
|      | Adenosine-5'-diphosphate (ADP) | Glycogen synthase 1 (1RZU)       | 1,5-anhydro-2-sorbitol (ASO) | Glycogen synthase 3 (3GUH) | 1,5-anhydro-2-sorbitol (ASO) |
| 1SGZ | Beta-secretase inhibitor   | Beta-secretase 1 (3I25)          | [(2S)-3-methyl-1-oxo-1-phenylmethyl]butan-2-ylamine | Beta-secretase 1 (2FDP) | N 1-(28,3S,5R)-3-amino-6-(4-fluorophenylamino)-5-methyl-6-oxo-1-phenylhexan-2-yl)n 2-n,3-n diproplylphosphohalamide (FRP) |
| 1TIB | Lipase                    | Lipase (1GT6)                    | Oleic acid (OLA)        | Lipase (1GT6)                    | Oleic acid (OLA)        |
| 1TVN | Cellulase                 | Endoglucanase cel5a (4A3H)      | 2,4-dinitrophenyl-2-deoxy-2-fluorocetyl-CoA (DCB) | Endoglucanase cel5a (1TVP) | 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (EPE) |
| 1V0S | Phospholipase D           | Phospholipase D (2ZE9)          | 1,2-diyldiheptanate (PDT) | Phospholipase D (1U52)          | 1,2-diacetyl-sn-glycero-3-phosphoethanolamine |
| 1YQ  | Threonyl-tRNA synthetase  | Threonyl-tRNA synthetase (3PD2) | Serine-3'-aminoadenosine (A3S) | Threonyl-tRNA synthetase (2HL0) | Serine-3'-aminoadenosine (A3S) |
|      | Serine-3'-aminoadenosine (A3S) | Shikimate kinase (3MUF)         | Adenosine-5'-diphosphate (ADP) | Shikimate kinase (1ZUI)         | Adenosine-5'-diphosphate (ADP) |
| 2CHS | Chorismate mutase         | Chorismate mutase (1COM)        | Prephenic acid (PRE)    | Chorismate mutase (1COM)        | Prephenic acid (PRE)    |
| 2DPS | Leucyl/phenylalanin-rRNA-protein transferase | Leucyl/phenylalanin-rRNA-protein transferase (2DPT) | Puromycin (PUY) | Leucyl/phenylalanin-rRNA-protein transferase (2DPT) | Puromycin (PUY) |
| 2GT2 | GDP-mannose mannosyl hydrolase | GDP-mannose mannosyl hydrolase (2GT4) | GDP-mannose mannosyl hydrolase (2GT4) | GDP-mannose mannosyl hydrolase (2GT4) | GDP-mannose mannosyl hydrolase (2GT4) |
|      | GDP-mannose mannosyl hydrolase (2GT4) | Guanosine-5'-diphosphate-alpha-d-mannose (GDD) | GDP-mannose mannosyl hydrolase (2GT4) | GDP-mannose mannosyl hydrolase (2GT4) | GDP-mannose mannosyl hydrolase (2GT4) |
| 2OP  | Beta-glucoisidase B       | Beta-glucoisidase B (2ZIS)       | 2-deoxy-2-fluoro-alpha-D-glucopyranosyl (GFP) | Beta-glucoisidase B (2ZIS)       | 2-deoxy-2-fluoro-alpha-D-glucopyranosyl (GFP) |
| 2SI  | Sialidase                 | Sialidase (2SIB)                 | 2-deoxy-2,3-dehydro-N-acetyl-neuraminic acid (ADN) | Trans-sialidase (2AH2) | 3-fluorosalic acid (FSI) |
| 2V78 | Fruktokinase              | 2-keto-3-deoxyglucuronate kinase (2VAR) | Phosphoaminophosphonoc acid-adenylate ester (ANP) | 2-keto-3-deoxyglucuronate kinase (2DCN) | Trihydroxy-5-[(phosphonooxy)methyl]tetrahydrofuran-2-carboxylic acid (CKP) |
|      | Adenosine monophosphate   | Phosphoaminophosphonic acid-adenylate ester (ANP) | 2-keto-3-deoxyglucuronate kinase (2DCN) | Trihydroxy-5-[(phosphonooxy)methyl]tetrahydrofuran-2-carboxylic acid (CKP) | 2-keto-3-deoxyglucuronate kinase (2DCN) |
| 2VQ4 | Glucoamylase A            | Glucoamylase A (3EJB)            | 1-thiobhexopyranoside (HTG) | No results | Glucoamylase A (2V8L) |
|      | Glucose                   | Acyl carrier protein (3EJB)      | 1-thiobhexopyranoside (HTG) | No results | Glucoamylase A (2V8L) |

- **1NDB**: Carnitine acetyltransferase
- **1R0M**: N-acetylglucosamine 2-acetyl-1-glucosamine
- **1RZV**: Glycogen synthase 1
- **1SGZ**: Beta-secretase inhibitor
- **1TIB**: Lipase
- **1TVN**: Cellulose
- **1V0S**: Phospholipase D
- **1YQ**: Threonyl-tRNA synthetase Serine-3'-aminoadenosine
- **1ZUH**: Shikimate kinase
- **2CHS**: Chorismate mutase
- **2DPS**: Leucyl/phenylalanin-rRNA-protein transferase
- **2GT2**: GDP-mannose mannosyl hydrolase
- **2OP**: Beta-glucoisidase B
- **2SI**: Sialidase
- **2V78**: Fruktokinase
- **2VQ4**: Glucoamylase A
| Enzyme Name | PDB Code | Description |
|-------------|----------|-------------|
| Aldo-ketoreductase | 2WZT | (2R,3R,4R,5R)-5-(6-amino-9H-purin-9-yl)-3-hydroxy-4-(phosphonooxy)-tetrahydrofuran-2-yl]-methyl dihydrogen diphosphate |
| Aldo-ketoreductase | 2WZM | (2R,3R,4R,5R)-5-(6-amino-9H-purin-9-yl)-3-hydroxy-4-(phosphonooxy)-tetrahydrofuran-2-yl]-methyl dihydrogen diphosphate |
| Aldose reductase | IA2Z | NADP nicotinamide-adenine dinucleotide phosphate (NAP) |
| Yvgn protein | 3D3F | NAD(P)H (NDP) |
| Beta-lactamase | 1GHP | Open form - penicillin g (PMN) |
| Beta-lactamase | 1BLC | N-(2-hydroxy-4-oxo-butyl)-N-(3-oxo-transpropenyl)amine (TEM) |
| Beta-lactamase | 4UXZ | (3R,6S)-2-hydroxy-3-[(thiophen-2-ylacetyl)amino]-1,2-oxaborinan-6-ylacetic acid (4D6) |
| Beta-lactamase | 1BLC | N-(2-hydroxy-4-oxo-butyl)-N-(3-oxo-transpropenyl)amine |
| Beta-lactamase | 4XUZ | (3R,6S)-2-hydroxy-3-[(thiophen-2-ylacetyl)amino]-1,2-oxaborinan-6-ylacetic acid |
| RNA polymerase | 3BR9 | (2r)-2-(3-[5-hydroxy-2-(3-methylbutyl)-3-oxo-6-thiophen-2-yl]-2,3-dihydropyridazin-4-yl]-1,1-dioxido-2h-1,2,4-benzothiadiazin-7-yloxy)propanamide (DEY) |
| RNA polymerase | 1YVX | 3-[isopropyl(4-methylbenzoyl)amino]-5-phenylthiophene-2-carboxylic acid (IPC) |
| RNA polymerase | 3CSO | RNA-directed RNA polymerase (3CSO) (11S)-10-acetyl-11-[4-(benzyloxy)-3-chlorophenyl]-3,3-dimethyl-2,3,4,5,10,11-hexahydro-1H-dibenzo[b,e][1,4]diazepin-1-one (XNI) |
| Fucosylgalactoside alpha N-acetylgalactosaminyltransferase | 3V0N | 5-(5-formylthiophen-2-yl)sulfuridine-5’(2-acetamino)-2-deoxy-alpha-D-galactosyl]-diphosphate |
| Galactosyltransferase | 1ZIZ | Beta-D-galactose (GAL) |
| Histo-blood group abo system transferase | 5C1H | Octyl 3-deoxy-2-O-(6-deoxy-alpha-L-galactopyranosyl)-beta-D-xyl-o-hexopyranoside (DA8) |
Table 3: COACH's ligand/binding site recognition performance on the LigaSite non-redundant data-set containing 373 apo-proteins for which a corresponding holo structure is available (Desailly et al., 2008). APO and HOLO Protein ID are the PDB codes of the corresponding three-dimensional structures. Ligand ID is the ligand name abbreviation used in the Protein Data Bank. Ligand type indicates the nature of the ligand: organic (O) or ion (I).

| APO Protein ID | HOLO Protein ID | Ligand ID | Rank | Ligand details | Ligand type |
|----------------|-----------------|-----------|------|----------------|-------------|
| 180L           | 3GUM            | PXY       | 1st  | Substrate analog | O           |
| 1A4U           | 2BD0            | NAP       | 1st  | Coafactor analog | O           |
| 1A5Y           | 1G1N            | GDP       | 1st  | Cofactor         | O           |
| 1AK1           | 2HRE            | PP9       | 1st  | Substrate       | O           |
| 1AKZ           | 3EUG            | GOL       | 1st  | Natural ligand  | O           |
| 1ARB           | 3E0N            | ARM       | 1st  | Specific inhibitor | O         |
| 1ARL           | 3WC7            | EF1       | 1st  | Specific inhibitor | O         |
| 1B5E           | 1GX9            | REA       | 1st  | Specific ligand  | O           |
| 1B5P           | 4PLT            | NAI       | 1st  | Cofactor         | O           |
| 1BD9           | 1B7A            | OPE       | 1st  | Natural ligand analog | O         |
| 1BK7           | 1IYB            | 5GP       | 1st  | Reaction product | O           |
| 1BKZ           | 4BL1            | 70B       | 1st  | Natural ligand analog | O         |
| 1BQC           | 5D9M            | GLC       | 1st  | Substrate       | O           |
| 1BYI           | 4WOP            | CTP       | 1st  | substrate       | O           |
| 1C48           | 1CQF            | GLA       | 1st  | Natural ligand  | O           |
| 1C5H           | 1H4G            | EXP       | 1st  | Specific inhibitor | O         |
| 1CEO           | 1CEN            | BGC       | 1st  | Substrate analog | O           |
| 1CEX           | 1XZM            | DPE       | 1st  | Covalent inhibitor | O          |
| 1CPI           | 2XU1            | 424       | 1st  | Specific inhibitor | O         |
| 1CRW           | 4WNC            | NAD       | 1st  | Cofactor         | O           |
| 1CWY           | 1ESW            | ACR       | 1st  | Specific inhibitor | O         |
| 1DCO           | 1DCP            | HBI       | 1st  | Cofactor analog  | O           |
| 1DHIN          | 2NM2            | NEU       | 1st  | Substrate analog | O           |
| 1DQ0           | 3UXA            | XMM       | 1st  | Natural ligand analog | O         |
| 1DUP           | 1SEH            | UMP       | 1st  | Reaction product | O           |
| 1E4F           | 4PL7            | ATP       | 1st  | Substrate       | O           |
| 1E5L           | 3ABI            | NAD       | 1st  | Cofactor         | O           |
| 1EBY           | 3QK0            | QK0       | 1st  | Specific inhibitor | O         |
| 1EDQ           | 3WL1            | NAG       | 1st  | Reaction product | O           |
| 1EPA           | 1EPB            | REA       | 1st  | Natural ligand  | O           |
| 1EWZ           | 3FZC            | MXF       | 1st  | Inhibitor       | O           |
| 1EV0           | 1A2T            | THP       | 1st  | Specific inhibitor | O         |
| 1EZI           | 3K8D            | CTP       | 1st  | Substrate       | O           |
| 1F14           | 4DYD            | NAD       | 1st  | Cofactor         | O           |
| 1F1S           | 10J3            | GCD       | 1st  | Reaction product | O           |
| 1F2V           | 1I1H            | COJ       | 1st  | Reaction product | O           |
| 1F41           | 4K7             | IMN       | 1st  | Specific ligand  | O           |
| 1FSZ           | 4OE7            | GXP       | 1st  | Product analog  | O           |
| 1FCQ           | 1FCV            | NAG       | 1st  | Substrate fragment | O         |
| 1FGB           | 1E1I            | GAA       | 1st  | Specific inhibitor | O         |
| 1F9D           | 2AM3            | UPG       | 1st  | Substrate analog | O           |
| 1F5F           | 1HOR            | AGP       | 1st  | Specific inhibitor | O         |
| 1FTF           | 2WDS            | COA       | 1st  | Substrate       | O           |
| 1FTR           | 2FH1            | MFN       | 1st  | Cofactor         | O           |
| 1FWL           | 1H73            | ANP       | 1st  | Substrate analog | O           |
| Protein ID | Chain | Reference | Type               | Description         |
|------------|-------|-----------|--------------------|---------------------|
| 1G40       |       | 1YSE SVR  | 1st                | Specific inhibitor  |
| 1G4E       |       | 1G4T FTP  | 1st                | Reaction product    |
| 1G9S       |       | 1HM9 UD1  | 1st                | Reaction product    |
| 1GBS       |       | 4CFP AMV  | 1st                | Substrate analog    |
| 1GCE       |       | 3OS8 BSH  | 1st                | Specific inhibitor  |
| 1GF5       |       | 3ICP NAD  | 1st                | Cofactor            |
| 1GOU       |       | 3GP      | 1st                | Substrate analog    |
| 1GSH       |       | 1GSA ADP  | 1st                | Reaction product    |
| 1GWK       |       | 1GWL BMA  | 1st                | Natural ligand      |
| 1GY0       |       | 1OG1 TAD  | 1st                | Substrate analog    |
| 1HK9       |       | 4Y91 NUC  | 1st                | Natural ligand      |
| 1HKA       |       | 1RAO HH2  | 1st                | Product analog      |
| 1HM5       |       | 6GP      | 1st                | Reaction product    |
| 1HNK       |       | 1HND COA  | 1st                | Reaction product    |
| 1HO1       |       | 1M5W DXP  | 3rd                | Substrate           |
| 1I7N       |       | 1ITL ATP  | 1st                | Endogenous ligand   |
| 1IAD       |       | 1QJI Peptide | 1st             | Inhibitor           |
| 1ILV       |       | 2V4O PO4  | 1st                | Reaction product    |
| 1INL       |       | 2OOL S4M  | 1st                | Substrate analog    |
| 1JK9       |       | 4ZDQ CTP  | 2nd                | Substrate           |
| 1K0M       |       | 1EV9 GST  | 1st                | Substrate analog    |
| 1K3O       |       | 3VPQ GSH  | 1st                | Substrate           |
| 1K6A       |       | 5D4Y XYP  | 1st                | Substrate           |
| 1KAM       |       | 1K4M NAD  | 1st                | Substrate           |
| 1KF5       |       | 1JUV C2P  | 1st                | Reaction product    |
| 1KHD       |       | 3R88 PRP  | 1st                | Substrate           |
| 1KN9       |       | 1B12 1PN  | 1st                | Specific inhibitor   |
| 1KPA       |       | 1RZY 5AS  | 1st                | Substrate analog    |
| 1KWB       |       | 1EB BPY  | 1st                | Substrate           |
| 1L7D       |       | 1L7E NA1  | 1st                | Reaction product    |
| 1LBV       |       | 1NV6 PH   | 1st                | Reaction product    |
| 1LC1       |       | 2DS1 SLU  | 1st                | Reaction intermediate analog |
| 1LF4       |       | 1APR Peptide | 1st             | Inhibitor           |
| 1LTU       |       | 1KW0 BH4  | 1st                | Cofactor            |
| 1M1Z       |       | 1JGT APC  | 1st                | Substrate analog    |
| 1MKB       |       | 4KEH 1R3  | 1st                | Substrate analog    |
| 1MMI       |       | 1D1E Peptide | 1st           | Specific inhibitor   |
| 1MR7       |       | 3NZ2 ACO  | 1st                | Substrate           |
| 1MTZ       |       | 1QXQ PCS  | 2nd                | Specific inhibitor   |
| 1MKW       |       | 3KV ANP  | 1st                | Substrate analog    |
| 1MZL       |       | 1UVC STE  | 1st                | Natural ligand      |
| 1N0S       |       | 1N07 FMN  | 1st                | Reaction product    |
| 1NDB       |       | 1SO 152   | 1st                | Substrate           |
| 1NON       |       | 4PS3 USP  | 1st                | Specific ligand     |
| 1NXM       |       | 2XCH TRH  | 1st                | Substrate analog    |
| 1O24       |       | 3HZG FAD  | 1st                | Cofactor            |
| 1O6EM      |       | 2CNH IZB  | 1st                | Specific inhibitor   |
| Code  | Description               | Type     | Code  | Description               | Type     |
|-------|---------------------------|----------|-------|---------------------------|----------|
| 1OFP  | Specific inhibitor         | O        | 1OGH  | Substrate analog          | O        |
| 1OGL  | Substrate analog          | O        | 1OGM  | Reaction product          | O        |
| 1OJQ  | Cofactor                  | O        | 1OLZ  | Specific ligand           | I        |
| 1OOI  | Natural ligand            | O        | 1OPY  | Transition state analog    | O        |
| 1OXT  | Reaction product          | O        | 1PX3  | Reaction product          | O        |
| 1PSH  | Substrate analog          | O        | 1P74  | Substrate analog          | O        |
| 1PDB  | Inhibitor                 | O        | 1PNG  | Substrate analog          | O        |
| 1Q52  | Product analog            | O        | 1Q7M  | Substrate                 | O        |
| 1QCZ  | Substrate analog          | O        | 1QCZ  | Specific ligand           | O        |
| 1QTO  | Specific ligand           | O        | 1QTR  | Substrate                 | O        |
| 1R0M  | Substrate                 | O        | 1R12  | Substrate analog          | O        |
| 1R29  | Natural ligand            | O        | 1R5  | Substrate                 | O        |
| 1RQI  | Substrate analog          | O        | 1R0I  | Non-specific ligand       | I        |
| 1RKH  | Specific ligand           | O        | 1RZV  | Substrate analog          | O        |
| 1S2L  | Substrate                 | O        | 1S2G  | Substrate                 | O        |
| 1S7K  | Substrate                 | O        | 1SGK  | Substrate analog          | O        |
| 1SGZ  | Inhibitor                 | O        | 1SFL  | Substrate                 | O        |
| 1JF1  | Cofactor                  | I        | 1SNT  | Specific inhibitor         | O        |
| 1SGG  | COFACTOR                  | O        | 1SUL  | Specific ligand           | O        |
| 1SWH  | Specific ligand analog    | O        | 1TGN  | Specific inhibitor         | O        |
| 1THV  | Natural ligand            | I        | 1TJG  | Reaction product          | O        |
| 1TJE  | Substrate analog          | O        | 1TIV  | Reaction product          | O        |
| 1TVN  | Substrate analog          | O        | 1TW0  | Non-specific ligand       | O        |
| 1TYF  | Substrate analog          | O        | 1TYV  | Natural ligand            | O        |
| 1TU7  | Substrate                 | O        | 1U0J  | Reaction product          | O        |
| 1U7U  | Substrate                 | O        | 1U9J  | Substrate analog          | O        |
| 1UB1  | Specific ligand           | O        |
| Code | Code | Name | Type     | Status |
|------|------|------|----------|--------|
| 1UF4 | 1UF5 | CDT  | Substrate | O      |
| 1ULU | 1VL8 | NAP  | Cofactor  | O      |
| 1UMF | 1QXO | FMN  | Cofactor  | O      |
| 1UXZ | 1UYX | BGC  | Specific ligand | O |
| 1UYL | 4YKQ | 4EO  | Inhibitor | O      |
| 1VO8 | 1VOR | WOS  | Inhibitor | I      |
| 1V6Z | 2CX8 | SAH  | Cofactor  | O      |
| 1V77 |      |      | No significant results |  |
| 1VFJ | 2XUL | ATP  | Specific ligand | O  |
| 1VIC | 2XWL | CTP  | Substrate | O      |
| 1VIU | 3DWR | 0PA  | Substrate analog | O  |
| 1VPN | 3BWR | SIA  | Specific ligand | O  |
| 1W9A | 2AQ6 | PLP  | Reaction product | O |
| 1WG1 | 3HG     | ATP  | Specific ligand | O  |
| 1WNY | 1W8K | VMS  | Substrate analog | O  |
| 1WOS | 1WOO | THG  | Substrate | O      |
| 1WS9 | 2JIF | FAD  | Cofactor  | O      |
| 1WTJ | 1V9N | NDP  | Cofactor  | O      |
| 1WXF | 2PZ8 | APC  | Substrate analog | O |
| 1X56 | 3A5Y | KAA  | Substrate analog | O |
| 1X7O | 1X7P | SAM  | Cofactor  | O      |
| 1X8F | 3UND | PEP  | Substrate | O      |
| 1XIX | 1P4N | Peptide | Substrate | O  |
| 1XK7 | 1T3Z | CAO  | Substrate analog | O  |
| 1XO6 | 3TV5 | RCMP | Inhibitor | O      |
| 1XOQ | 1XQP | 8HG  | Substrate analog | O  |
| 1XW2 | 1WU5 | XYP  | Reaction product | O  |
| 1Y2Q | 4RR6 | A3S  | Substrate analog | O  |
| 1Y2T | 4ZSN | NDG  | Endogenous ligand analog | O |
| 1YBT | 4WP8 | ZDA  | inhibitor | O      |
| 1YYY | 2GMV | PEP  | Reaction product | O  |
| 1Z7G | 1HMP | 5GP  | Reaction product | O  |
| 1ZAH | 3TU9 | 5MM  | Specific inhibitor | O |
| 1ZCU | 3T7O | UPG  | Substrate | O      |
| 1ZNW | 1ZNX | 5GP  | Substrate | O      |
| 1ZTY | 1ZU0 | CBS  | Natural ligand | O  |
| 1ZUH | 2YJR | SKM  | Substrate | O      |
| 2A6Z | 1N47 | TNR  | Specific ligand | O  |
| 2A8F | 2AB   | ERG  | Specific ligand | O  |
| 2AD1 | 1LS6 | A3P  | Reaction product | O  |
| 2AHF | 2AHG | UCD  | Substrate | O      |
| 2AHU | 2AHV | COA  | Cofactor  | O      |
| 2AMJ | 4FGJ | FAD  | Cofactor  | O      |
| 2B0J | 3F46 | 12C  | Cofactor  | O      |
| 2B6P | 2EUV | GOL  | Non-specific ligand | O  |
| 2B78 | 3VSE | SAH  | Cofactor  | O      |
| 2B98 | 2C9D | PHR  | Inhibitor | O      |
| 2BGT | 1IXY | UDP  | Reaction product | O  |
| 2BFW | 4LZZ | 08T  | Substrate analog | O  |
| 2BOE | 2BOG | SGC  | inhibitor | O      |
| 2CS1 | 3OAA | ANP  | Reaction product analog | O |
| 2C7I | 3A7R | LAQ  | Reaction intermediate | O  |
| Code  | Code  | Code  | Code  | Code  | Code  |
|-------|-------|-------|-------|-------|-------|
| 2CAR  | 3S86  | IMP   | 1st   | Reaction product | O     |
| 2CHS  | 1COM  | PRE   | 1st   | Reaction product | O     |
| 2CI3  | 1LYX  | CYR   | 1st   | Reaction product | O     |
| 2CWK  | 1F3F  | D4T   | 1st   | Substrate analog | O     |
| 2CX5  | 2ZOX  | 5CA   | 1st   | Substrate analog | O     |
| 2DS9  | 4XYL  | COA   | 1st   | Substrate       | O     |
| 2DHQ  | 2BT4  | CA2   | 1st   | Specific inhibitor | O   |
| 2DI6  | 3QNA  | BIO   | 1st   | Substrate       | O     |
| 2DKA  | 2DKD  | PO4   | 1st   | Substrate       | I     |
| 2DPS  | 2Z3O  | PHE   | 1st   | Substrate       | O     |
| 2DQW  | 3TZF  | HH2   | 1st   | Substrate analog | O     |
| 2EOC  | 4AJ3  | NAP   | 1st   | Substrate       | O     |
| 2F9T  | 5B8H  | PAU   | 1st   | Substrate       | O     |
| 2EK7  | 1KPG  | SAH   | 1st   | Cofactor        | O     |
| 2FP8  | 2V91  | S55   | 2nd   | Substrate       | O     |
| 2FSF  | 2FSH  | ANP   | 1st   | Substrate analog | O     |
| 2G67  | 2QTC  | TDK   | 1st   | Reaction intermediate analog | O |
| 2G95  | 2H1D  | NMN   | 1st   | Substrate       | O     |
| 2GFV  | 2FYZ  | DAO   | 1st   | Substrate analog | O     |
| 2GG4  | 3GGA  | S3P   | 1st   | Substrate       | O     |
| 2GQV  | 2P4T  | NAP   | 1st   | Cofactor        | O     |
| 2GSF  | 4HYH  | 1AM   | 1st   | Specific inhibitor | O   |
| 2GT2  | 2GT4  | GDD   | 1st   | Substrate       | O     |
| 2GUB  | 1SSN  | XYL   | 1st   | Inhibitor       | O     |
| 2GWX  | 3V18  | 13M   | 4th   | Natural ligand analog | O |
| 2GYY  | 1DSS  | NAD   | 1st   | Cofactor        | O     |
| 2H2Z  | 2A5I  | AZP   | 1st   | Inhibitor       | O     |
| 2HBJ  | 1ZBH  | MG    | 1st   | Cofactor        | I     |
| 2HV2  | 2HX   | ATP   | 1st   | Substrate       | O     |
| 2HK0  | 2HK1  | FUD   | 1st   | Substrate       | O     |
| 2HVM  | 1FFR  | NAG   | 1st   | Substrate       | O     |
| 2HY7  | 2Q6V  | UDP   | 1st   | Substrate analog/reaction product | O |
| 2HZR  | 2HZQ  | STR   | 1st   | Natural ligand  | O     |
| 2HL1  | 2J3L  | PSA   | 1st   | Reaction intermediate analog | O |
| 2IOB  | 2IO9  | GSH   | 1st   | Substrate       | O     |
| 2IRU  | 3PKY  | UTP   | 1st   | Substrate       | O     |
| 2J71  | 2J72  | GLC   | 1st   | Natural ligand  | O     |
| 2KN   | 2CY2  | ACO   | 1st   | Cofactor        | O     |
| 2BR   | 3PFD  | FDA   | 1st   | Cofactor        | O     |
| 2NWD  | 1BB5  | CTO   | 1st   | Inhibitor       | O     |
| 2NXC  | 3CJT  | SAM   | 1st   | Cofactor        | O     |
| 2NPI  | 1OM   | NOJ   | 1st   | Inhibitor       | O     |
| 2OAM  | 3E1T  | FAD   | 1st   | Cofactor        | O     |
| 2OPT  | 3B6A  | ZCT   | 1st   | Specific inhibitor | O   |
| 2OU1  | 1L6L  | BOG   | 10th  | Natural ligand analog | O |

Daniele Toti et al
| Code | Code | Type          | Description                        |
|------|------|---------------|------------------------------------|
| 2OVE | 2OVG | DAO           | Specific ligand                     |
| 2PAW | 4ZZZ | FSU           | Specific inhibitor                  |
| 2PFK | 2F88 | FBP           | Reaction product                   |
| 2PKF | 4GI1 | AGI           | Substrate analog                   |
| 2PPN | 1BL4 | API           | Specific ligand                     |
| 2QSR | 2JG1 | ANP           | Cofactor analog                    |
| 2QGV | 1R3Q | 1CP           | Reaction product                   |
| 2QBV | 5CKX | TSA           | Transition state analog             |
| 2QDK | 1RXU | THM           | Substrate analog                   |
| 2QEV | 2QEH | SRO           | Specific ligand                     |
| 2QSU | 3DP9 | BIG           | Substrate analog                   |
| 2QVL | 2QV7 | ADP           | Reaction product                   |
| 2QYS | 3C10 | NAP           | Cofactor                           |
| 2R60 | 3S27 | FRU           | Substrate                          |
| 2RG7 | 2R7A | HEM           | Specific ligand                     |
| 2RJD | 3H6Z | Peptide       | Specific ligand                     |
| 2SGA | 3NCL | CCZ           | Specific inhibitor                  |
| 2Sil | 2AH2 | FSI           | Inhibitor                          |
| 2TS1 | 3TZL | TRP           | Substrate analog                   |
| 2UYO | 1RJD | SAM           | Cofactor                           |
| 2V78 | 3QAI | XNN           | Inhibitor                          |
| 2VFB | 4BS5 | P18           | Specific inhibitor                  |
| 2VFY | 2VFK | AMP           | Specific ligand                     |
| 2VQ4 | 2V8L | GLC           | Natural ligand                     |
| 2VUA | 3RSJ | GAL           | Specific ligand                     |
| 2WN4 | 1GR | NDP           | Substrate                          |
| 2WVT | 2BHI | TPP           | Cofactor                           |
| 2WZT | 3DF3 | NDP           | Cofactor                           |
| 2XXS | 2PA4 | UPG           | Product analog                     |
| 2YWB | 2YWB | AMP           | Product analog                     |
| 2YXF | 4RA3 | TFX           | Non-specific ligand                |
| 2YYA | 3RV3 | ADP           | Reaction product                   |
| 2YYT | 4MUZ | BMP           | Specific inhibitor                  |
| 2YZG | 1EZ1 | ANP           | Substrate analog                   |
| 2ZBS | 5DKS | 5C6           | Specific ligand                     |
| 2ZCG | 4MUZ | BMP           | Specific inhibitor                  |
| 2ZCO | 3W3H | P57           | Reaction intermediate              |
| 2ZGL | 2R0H | CTO           | Natural ligand                     |
| 2ZHY | 3KE5 | ATP           | Substrate                          |
| 2ZJ8 | 4F93 | ATP           | Substrate                          |
| 2ZTY | 3ZZD | G85           | Specific inhibitor                  |
| 3A0Y | 3TZ5 | ADP           | Reaction product                   |
| 3ASQ | 2D52 | COA           | Substrate                          |
| 3AAP | 3CIA | ANP           | Specific inhibitor                  |
| 3ADO | 4OM8 | NAD           | Cofactor                           |
| 3APP | 3K5C | 0B1           | Specific inhibitor                  |
| 3BG3 | 2FYT | SAH           | Cofactor                           |
| 3BA1 | 4NJO | NAD           | Cofactor analog                    |
| 3BLM | 4UXU | 4D6           | inhibitor                          |
| 3BTY | 3NTO | NAD           | Specific ligand                     |
| 3BUE | 2FZ2 | ARG           | Natural ligand                     |
| 3BYL | 1W2T | SUC           | Substrate                          |
| PDB Code | Monomer Abbreviation | Substrate/Co factor Type | Additional Info |
|----------|----------------------|--------------------------|-----------------|
| 3C2E     | IOPQ NTM             | 1st Substrate            | O               |
| 3C8N     | 3B4Y F42             | 1st Cofactor             | O               |
| 3CAF     | 3CU1 SCR             | 2nd Specific ligand      | O               |
| 3CB6     |                      |                          | No significant results |
| 3CGZ     | 4PL9 ADP             | 1st Reaction product     | O               |
| 3COU     | 2A8S GTP             | 2nd Substrate analog     | O               |
| 3CQ1     | 3CQ3 MG              | 1st Non-specific ligand  | I               |
| 3CRM     | 2ZE7 DST             | 2nd Substrate analog     | O               |
| 3CSR     | 3CT5 NAG             | 1st Substrate analog     | O               |
| 3CTB     | 4OK1 198             | 1st Specific agonist     | O               |
| 3DOO     | 1UXJ NAD             | 1st Cofactor             | O               |
| 3D9S     | 1FE3 OLA             | 1st Ligand analog        | O               |
| 3DRD     | 4ATP PLP             | 1st Cofactor             | O               |
| 3DRE     | 4V7N ADP             | 1st Reaction product     | O               |
| 3DUL     | 5FHQ SAM             | 1st Cofactor             | O               |
| 3E1S     | 4C30 ANP             | 1st Substrate analog     | O               |
| 3E5B     | 2HJP PPR             | 1st Substrate analog     | O               |
| 3EZI     | 2AU9 MG              | 1st Specific ligand      | I               |
| 3EK6     | 2J4L UTP             | 1st Specific inhibitor   | O               |
| 3ETF     | 4SP NAD              | 1st Cofactor             | O               |
| 3EX9     | 3JUM AOD             | 2nd Reaction intermediate analog | O |
| 3EXR     | 1X8Z LX1             | 1st Reaction product     | O               |
| 3FL1     | 1VL8 NAP             | 1st Cofactor             | O               |
| 3FSF     | 2IMJ GSH             | 1st Substrate            | O               |
| 3FTD     | 4G9 SAM              | 1st Cofactor             | O               |
| 3FP6     | 4QFG AMP             | 1st Natural ligand analog | O |
| 3GIS     | 3K6 BMP              | 1st Specific inhibitor   | O               |
| 3GBT     | 1BF6 GOL             | 1st Substrate analog     | O               |
| 3GD0     | 3GQ9 BCG             | 1st Substrate/product analog | O |
| 3GLK     | 3O1Z ADP             | 1st Reaction product     | O               |
| 3GPQ     | 3GQ9 APR             | 1st Specific ligand      | O               |
| 3GQH     | 3SUC ATP             | 1st Specific ligand      | O               |
| 3GSZ     | 3SO XNI              | 1st Inhibitor            | O               |
| 3HGZ     | 2XLC DEP             | 1st Specific inhibitor   | O               |
| 3H38     | 1MY CTP              | 1st Substrate            | O               |
| 3H49     | 3QAI XNN             | 1st Specific inhibitor   | O               |
| 3H7I     | 2HA7 ZMR             | 1st Specific inhibitor   | O               |
| 3HBH     | 3HD MXE              | 1st Substrate mimic      | O               |
| 3HIS     | 4MG ADN              | 1st Substrate analog     | O               |
| 3HJ4     | 2QOD ATP             | 1st Substrate analog     | O               |
| 3HJX     | 3HP8 SUC             | 1st Specific ligand      | O               |
| 3HC     | 5C1H DA8             | 1st Inhibitor            | O               |
| 3HJ     | 1KQU BR4             | 1st Substrate analog     | O               |
| 3HS      | 4R98 GNH             | 1st Substrate analog     | O               |
| 3HK8     | 3KD J9Z             | 1st Specific inhibitor   | O               |
| 3LY      | 3OF Peptide          | 8th Substrate mimic      | O               |
| 3TY      | 3TTL LRH             | 1st Reaction product     | O               |
| 3UJ      | 1Q6F PRO             | 1st Substrate analog     | O               |
| 3YJ      | 1H2B NAI             | 1st Cofactor             | O               |
| 30M      | 2POY Peptide         | 1st Specific ligand      | O               |
| 3KAJ     | 3KAL HGS             | 1st Reaction product     | O               |
| 3KDH     | 3NMV PYV             | 1st Natural ligand analog | O |
| APO Protein ID | HOLO Protein ID | Ligand ID | Rank | Ligand details       |
|---------------|----------------|-----------|------|----------------------|
| 1A4U          | 1MG5           | NAI       | 1st  | Cofactor             |
| 1ADE          | 2V40           | GDP       | 1st  | Substrate            |
| 1AK1          | 1C9E           | MP1       | 1st  | Substrate analog     |
| 1AKZ          | 2C53           | DUR       | 5th  | Substrate analog     |
| 1ARL          | 4NSV           | 20Y       | 1st  | Specific inhibitor   |
| 1ART          | 3FVL           | BHK       | 1st  | Specific inhibitor   |
| 1B8E          | 3NQ3           | DKA       | 2nd  | Substrate analog     |
| 1B8P          | 1BMD           | NAD       | 1st  | Cofactor             |
| 1BD9          | 2B7A           | OPE       | 1st  | Natural ligand analog|
| 1BK7          | 1UCA           | U2P       | 1st  | Reaction product     |
| 1BKZ          | 4UW6           | VV7       | 1st  | Natural ligand       |
| 1BQC          | 2MAN           | MAN       | 1st  | Reaction product     |
| 1BY1          | 3QXS           | NAP       | 3rd  | Substrate analog     |
| 1C48          | 1B0S           | GLA       | 1st  | Natural ligand       |
| 1C5H          | 2B46           | XYP       | 1st  | Reaction product     |
| 1CEO          | 3AOF           | BMA       | 1st  | Substrate            |
| 1C6X          | 1X2L           | HEE       | 1st  | Covalent inhibitor   |
| 1CPJ          | 2DC7           | 042       | 1st  | Covalent inhibitor   |
| 1CRW          | 2B4R           | NAD       | 1st  | Cofactor             |
| 1CWY          | 20WC           | ACR       | 1st  | Specific inhibitor   |
| 1DCO          | 2V6T           | H2B       | 3rd  | Substrate-like ligand|
| 1DHN          | 1RRY           | 204       | 1st  | Inhibitor            |
| 1DQ0          | 4P9W           | R3M       | 1st  | Specific ligand analog|
| 1DUJ          | 1SYL           | DUT       | 1st  | Substrate analog     |
| 1E4F          | 4A2B           | AGS       | 1st  | Substrate analog     |

Table 4: LIBRA-WA’s ligand/binding site recognition performance excluding from the ligand binding sites database the holo proteins present in the LigaSite non-redundant dataset (Dessailly et al., 2008). APO and HOLO Protein ID are the PDB codes of the corresponding three-dimensional structures. Ligand ID is the ligand name abbreviation used in the Protein Data Bank.
| Code   | Description         | Status |
|--------|---------------------|--------|
| 1E5L   | Cofactor            |        |
| 1E5Y   | Specific inhibitor  |        |
| 1EDQ   | Ligand analog      |        |
| 1EWA   | Specific inhibitor  |        |
| 1EY0   | Specific inhibitor  |        |
| 1EZI   | Reaction product    |        |
| 1F14   | Cofactor            |        |
| 1F1S   | Specific inhibitor  |        |
| 1F2V   | Reaction intermediate|      |
| 1F41   | Specific inhibitor  |        |
| 1F5Z   | Substrate           |        |
| 1FCQ   | Substrate analog    |        |
| 1FGB   | Specific inhibitor  |        |
| 1FDI   | Reaction intermediate|    |
| 1F5Z   | Substrate           |        |
| 1FCQ   | Substrate analog    |        |
| 1F99   | Substrate analog    |        |
| 1FSF   | Specific inhibitor  |        |
| 1FTF   | Reaction product    |        |
| 1FWL   | Substrate analog    |        |
| 1G40   | Competitive inhibitor|     |
| 1G4E   | Reaction product    |        |
| 1G9S   | Inhibitor           |        |
| 1GBS   | Substrate analog    |        |
| 1GCE   | Inhibitor           |        |
| 1GFS   | Reaction product    |        |
| 1GOU   | Substrate analog    |        |
| 1GSH   | Cofactor            |        |
| 1GWS   | Natural ligand      |        |
| 1GY0   | Cofactor            |        |
| 1HK9   | Natural ligand      |        |
| 1HKA   | Inhibitor           |        |
| 1HM5   | Reaction product    |        |
| 1HNK   | Product             |        |
| 1H0I   | Reaction product    |        |
| 1I7N   | Ligand analog       |        |
| 1IAD   | Inhibitor           |        |
| 1ILV   | Specific inhibitor  |        |
| 1INL   | Substrate analog    |        |
| 1J8S   | Cofactor            |        |
| 1JYK   | Specific inhibitor  |        |
| 1K3O   | No significant results|   |
| 1K0M   | Substrate analog    |        |
| 1K3O   | Natural ligand mimic|      |
| 1K6A   | Reaction product    |        |
| 1KAM   | Cofactor            |        |
| 1K5S   | Substrate analog    |        |
| 1KXO   | No significant results|   |

*Note: Specific inhibitor, Competitive inhibitor, Product, Natural ligand mimic*
| Code | Description                  | Type         |
|------|------------------------------|--------------|
| 1KPA | 3HHN                         | Substrate   |
| 1KWB | 1EIR                         | Substrate   |
| 1L7D | 1NM5                         | Substrate   |
| 1LBV | 1BY                          | Product     |
| 1LCI | 1BA3                         | Inhibitor   |
| 1LF4 | 1LF2                         | Inhibitor   |
| 1LTU | 1DMW                         | Cofactor    |
| 1M1Z | 3SEZ                         | Substrate   |
| 1MKB | 4KEH                         | Substrate analog |
| 1MMI | 3QSB                         | Specific inhibitor |
| 1MR7 | 4HUR                         | Substrate   |
| 1MTZ | 1XRQ                         | Substrate   |
| 1MWK | 4AG1                         | Substrate analog |
| 1MZL | 1UTC                         | Natural ligand |
| 1N05 | 1NB9                         | Substrate   |
| 1NDB | 2H3P                         | Substrate   |
| 1NQN | 4P66                         | Substrate analog |
| 1NXM | 1DZT                         | Substrate analog |
| 1O24 | 4KAT                         | Cofactor    |
| 1OEM | 2QBP                         | Inhibitor   |
| 1OFP | 1OF6                         | Specific inhibitor |
| 1OGH | 3TPS                         | Substrate analog |
| 1OGL | 2YB0                         | Substrate analog |
| 1OGM | 3WWG                         | Substrate analog |
| 1OJQ |                             | No significant results |
| 1OLZ |                             | No significant results |
| 1OOI | 3BS8                         | Substrate analog |
| 1OPY | 3OWY                         | Transition state analog |
| 1OXT | 3C4J                         | Nucleotide analogs |
| 1P1X | 1JCJ                         | Reaction intermediate |
| 1P5H | 1T3Z                         | Substrate analog |
| 1P74 | 3PHJ                         | Reaction product |
| 1PDB | 1BOZ                         | Cofactor    |
| 1PNG | 3PM3                         | Non-specific ligand |
| 1Q52 | 4QII                         | Substrate analog |
| 1Q7M | 2E7F                         | Substrate analog |
| 1QCZ | 2FWP                         | Substrate   |
| 1QID | 1H23                         | Specific inhibitor |
| 1QTO | 4AG1                         | Non-specific ligand |
| 1QTR | 1MU0                         | Inhibitor   |
| 1R0M | 1SJH                         | Product     |
| 1R12 | 3ZWP                         | Substrate analog |
| 1R29 | 4CPJ                         | Specific inhibitor |
| 1R15 | 3BGV                         | Reaction product |
| 1R3Q | 3HXQ                         | Substrate analog |
| 1RJ1 | 2C77                         | Non-specific ligand |
| 1RMK |                             | No significant results |
| 1RZV | 3CX4                         | Substrate   |
| 1S2L | 1S2D                         | Substrate analog |
| 1S7K |                             | No significant results |
| 1SGK | 1DTP                         | Substrate analog |
| 1SGZ | 3Q25                         | Specific inhibitor |
| PDB Code | Chain | Description  |
|----------|-------|-------------|
| 1SJY     |       | Non-specific ligand |
| 1SQI     |       | Inhibitor |
| 1SQG     |       | Cofactor |
| 1SUL     |       | No significant results |
| 1SNT     |       | Specific ligand analog |
| 1TGN     |       | Specific inhibitor |
| 1THV     |       | Non-specific ligand |
| 1TIB     |       | Reaction product |
| 1TJE     |       | Substrate analog |
| 1TVN     |       | Substrate analog |
| 1TV0     |       | Substrate analog |
| 1TJ9     |       | Non-specific ligand |
| 1TJ1     |       | Reaction product |
| 1TJY     |       | Natural ligand analog |
| 1TJZ     |       | No significant results |
| 1U0T     |       | Substrate analog |
| 1UAJ     |       | Substrate analog |
| 1UB4     |       | Specific ligand |
| 1UF4     |       | Substrate |
| 1UJU     |       | Cofactor |
| 1UMF     |       | No significant results |
| 1UXZ     |       | Specific ligand |
| 1ULY     |       | Substrate |
| 1VS0     |       | Reaction product |
| 1V6Z     |       | Substrate analog |
| 1V77     |       | No significant results |
| 1VFJ     |       | Specific ligand |
| 1VJC     |       | Substrate analog |
| 1VJU     |       | No significant results |
| 1VPN     |       | Specific ligand |
| 1W9A     |       | Reaction product analog |
| 1WNY     |       | Substrate analog |
| 1WOS     |       | Reaction product analog |
| 1WS9     |       | Cofactor |
| 1WTJ     |       | No significant results |
| 1WXF     |       | Substrate analog |
| 1X5G     |       | Substrate analog |
| 1X7O     |       | Substrate analog |
| 1X8F     |       | Substrate |
| 1X9X     |       | No significant results |
| 1XK7     |       | Substrate analog |
| 1X06     |       | Substrate |
| 1XQ0     |       | No significant results |
| 1XW2     |       | Reaction product |
| 1Y7Q     |       | Substrate analog |
| 1Y2T     |       | Endogenous ligand |
| 1YBT     |       | Substrate analog |
| 1YYY     |       | Substrate |
| 1Z7G     |       | Specific inhibitor |
| 1ZAH     |       | Substrate |
| 1ZCU     |       | Substrate |
| Code  | Chain | Type                | Substrate | 1st | 2nd | 3rd |
|-------|-------|---------------------|-----------|-----|-----|-----|
| 1ZNW  | F3T   | Substrate analog    |           |     |     |     |
| 1ZTY  |        | No significant results |           |     |     |     |
| 1ZUH  | U3I   | Substrate           |           |     |     |     |
| 2AZ6  |        | No significant results |           |     |     |     |
| 2ASF  | LRI   | Specific ligand     |           |     |     |     |
| 2ADH  | O6D   | Transition state analog |       |     |     |     |
| 2AFH  | U2Z   | Non-specific ligand |           |     |     |     |
| 2AHU  |        | No significant results |           |     |     |     |
| 2AMJ  | P6E   | Cofactor            |           | 1st |     |     |
| 2B09  | F46   | Cofactor            |           |     | 1st |     |
| 2B6P  | B6O   | Specific ligand analog |       |     |     |     |
| 2BT8  | V6E   | Cofactor            |           |     |     | 1st |
| 2B98  |        | No significant results |           |     |     |     |
| 2BG7  | N7F   | Substrate analog    |           |     |     |     |
| 2BFW  | V7I   | Substrate analog    |           |     | 1st |     |
| 2BOE  | O6D   | Reaction product    |           |     |     |     |
| 2G1   | B2Q   | Reaction product    |           |     | 1st |     |
| 2C7I  | R07   | Substrate analog    |           |     |     | 1st |
| 2CAR  |        | No significant results |           |     |     |     |
| 2CHS  | P7P   | Substrate           |           |     |     | 1st |
| 2C13  | C1I   | Specific inhibitor   |           |     | 1st |     |
| 2CWK  | HVD   | Substrate           |           |     |     | 1st |
| 2CX5  |        | No significant results |           |     |     |     |
| 2D59  | Q9N   | Natural ligand      |           |     |     |     |
| 2DHQ  | N76   | Specific inhibitor   |           |     |     | 1st |
| 2D6  | Y13   | Substrate           |           |     | 1st |     |
| 2DKA  | U2W   | Inhibitor           |           |     |     | 93rd |
| 2DPS  | Z3P   | Substrate           |           |     |     | 1st |
| 2DQW  | EYE   | Substrate analog    |           |     | 1st |     |
| 2ECH  | XKD   | Substrate           |           |     |     | 1st |
| 2ECK  | S4D   | Substrate analog    |           |     | 6th |     |
| 2E10  | WOQ   | Substrate analog    |           |     |     | 1st |
| 2E1V  | X7R   | Substrate           |           |     |     | 5th |
| 2E3S  | Z9Z   | Substrate           |           |     | 1st |     |
| 2ECR  | RZ1   | Cofactor            |           |     |     | 1st |
| 2EX0  | S44   | Substrate analog    |           |     |     | 1st |
| 2F82  | WYA   | Reaction product    |           |     | 1st |     |
| 2FPT  | F3I   | Reaction product    |           |     | 1st |     |
| 2FK7  | KPI   | Cofactor            |           |     |     | 1st |
| 2FP8  | V91   | Substrate           |           |     |     | 1st |
| 2FSF  | XZ   | Substrate           |           |     |     | 1st |
| 2G67  | A17   | Reaction product    |           |     | 1st |     |
| 2G95  | M6Q   | Reaction intermediate |         |     |     | 1st |
| 2GFV  | C70   | Inhibitor           |           |     |     | 1st |
| 2GG4  | F67   | Substrate           |           |     | 1st |     |
| 2GQV  | P4T   | Cofactor            |           |     |     | 4th |
| 2GSF  | PSZ   | Specific inhibitor   |           |     |     | 1st |
| 2GT2  | GT4   | Substrate           |           |     |     | 1st |
| 2GBU  | QDW   | Reaction intermediate |       |     |     |     |
| 2GWX  | GWX   | Natural ligand      |           |     |     | 1st |
| 2GYY  | R3N   | Cofactor            |           |     |     | 1st |
| 2H2Z  | A5K   | Inhibitor           |           |     |     | 1st |
| Protein ID | Substrate | Product | Date | Type |
|------------|-----------|---------|------|------|
| 2HBK       | MN        | 1st     | Cofactor |
| 2HIV       | AMP       | 1st     | Reaction product |
| 2HK0       | TAG       | 1st     | Substrate |
| 2HVM       | NAG       | 1st     | Substrate |
| 2HY7       |           |         | No significant results |
| 2HZR       | PLM       | 1st     | Natural ligand analog |
| 2IVL       | PSA       | 1st     | Substrate analog |
| 2IOB       | TS5       | 1st     | Substrate |
| 2IRU       | DTP       | 1st     | Substrate analog |
| 2J7I       |           |         | No significant results |
| 2J8N       | ACO       | 1st     | Cofactor |
| 2JBR       |           |         | No significant results |
| 2NWD       | 3ZW       | 1st     | Substrate analog |
| 2NXC       | SAM       | 1st     | Cofactor |
| 2OP9       | G2F       | 1st     | Product analog |
| 2OAM       | FAD       | 1st     | Cofactor |
| 2OPT       |           |         | No significant results |
| 2OU1       | PLM       | 1st     | Natural ligand |
| 2OVE       | REA       | 1st     | Natural ligand analog |
| 2PAW       | 15R       | 1st     | Specific inhibitor |
| 2PFK       | ADP       | 1st     | Reaction product |
| 2PKF       | AGS       | 1st     | Substrate analog |
| 2PPN       | AP1       | 1st     | Specific ligand |
| 2Q5R       | ANP       | 1st     | Substrate analog |
| 2Q6Z       | ACP       | 1st     | Reaction product |
| 2QBV       | MLT       | 1st     | Product analog |
| 2QDK       | THM       | 1st     | Reaction product analog |
| 2QEV       | LNR       | 1st     | Natural ligand |
| 2QSU       | ADE       | 1st     | Reaction product |
| 2QVL       | AMP       | 1st     | Reaction product analog |
| 2QYS       | NAP       | 1st     | Cofactor |
| 2R60       | P6P       | 1st     | Substrate |
| 2RG7       | BPH       | 55th    | Endogenous ligand analog |
| 2RJD       | P8H       | 1st     | Specific inhibitor |
| 2SGA       | TYR       | 1st     | Reaction intermediate |
| 2SIL       | DAN       | 1st     | Specific inhibitor |
| 2TS1       | YSA       | 1st     | Substrate analog |
| 2UYO       |           |         | No significant results |
| 2Y78       | ADP       | 1st     | Reaction product |
| 2YBV       | COA       | 1st     | Cofactor |
| 2FY4       | A3P       | 13th    | Specific ligand analog |
| 2VQ4       | GLC       | 2nd     | Natural ligand |
| 2VUA       | GAL       | 7th     | Natural ligand |
| 2WN4       | NDP       | 1st     | Cofactor |
| 2VY7       | TDL       | 1st     | Reaction intermediate |
| 2WZT       | FLF       | 9th     | Specific inhibitor |
| 2XS5       | M1P       | 1st     | Substrate |
| 2YWB       | AMP       | 1st     | Product analog |
| 2YXF       | NA        | 10th    | Non-specific ligand |
| 2YYA       | ADP       | 1st     | Substrate analog |
| 2YYT       | XMP       | 1st     | Product analog |
| PDB ID | Chain | Substance | Type |
|--------|-------|-----------|------|
| 2YZG   | 4LJK  | ANP       | Substrate analog |
| 2ZBS   | 159Q  | OHT       | Natural ligand analog |
| 2ZCG   | 3MWU  | UPT       | Covalent inhibitor |
| 2ZCO   | 3ADZ  | PS7       | Reaction intermediate |
| 2ZGL   | 2ZGO  | LAT       | Natural ligand analog |
| 2ZHY   | 3KI3  | ATP       | Substrate |
| 2ZJ8   | 2UIA  | AMP       | Specific ligand |
| 2ZTY   | 3ZZ9  | G83       | Specific inhibitor |
| 3A0Y   | 1Y8O  | ADP       | Reaction product |
| 3A5Q   | 1CHW  | HXC       | Substrate analog |
| 3AAP   | 4BR7  | AU1       | Specific inhibitor |
| 3ADO   | 3F3S  | NAD       | Cofactor |
| 3APP   | 1PPM  | 0F1       | Specific inhibitor |
| 3B3G   | 2Y1W  | SFG       | Cofactor analog |
| 3BAI   | 2DBQ  | NAP       | Cofactor |
| 3BLM   | 1GHP  | PNM       | Reaction intermediate |
| 3BTV   |       |           | No significant results |
| 3BUE   | 2ZFZ  | ARG       | Natural ligand |
| 3BYL   | 1PT2  | SUC       | Substrate |
| 3C2E   | 3C2R  | PHT       | Specific inhibitor |
| 3CN   | 4IC4  | MLY       | Non-specific ligand |
| 3CAF   | 1EOO  | SGN       | Specific ligand |
| 3CB6   |       |           | No significant results |
| 3CGZ   | 2GR7  | CNE       | Specific inhibitor mimic |
| 3COU   | 2A8S  | GTP       | Substrate analog |
| 3CQ1   |       |           | No significant results |
| 3CRM   | 3CRQ  | DPO       | Reaction product |
| 3CSR   |       |           | No significant results |
| 3CTB   | 1ILH  | SRL       | Specific agonist |
| 3DOO   | 1LDN  | NAD       | Cofactor |
| 3D9S   | 1CBQ  | RE9       | Natural ligand analog |
| 3DRD   | 3DOD  | PLP       | Cofactor |
| 3DRE   | 3U6   | ANP       | Substrate analog |
| 3DUL   | 1SUS  | SAH       | Reaction product |
| 3EIS   | 2EWV  | ADP       | Substrate analog |
| 3E5B   | 3P0X  | PYR       | Reaction product analog |
| 3EZ    | 3EFY  | POP       | Substrate |
| 3EK6   | 2BND  | UDP       | Reaction product |
| 3EF   | 4ITA  | NDP       | Cofactor |
| 3EX9   | 3FF0  | GOL       | Substrate mimic |
| 3EXR   | 1KW1  | LG6       | Substrate analog |
| 3FL   | 3GY0  | NAP       | Cofactor |
| 3F6F   | 1PN9  | GTX       | Specific inhibitor |
| 3F7D   | 2ZBR  | SFG       | Cofactor analog |
| 3FVI   | 2QRI  | ADP       | Natural ligand |
| 3G1S   | 3G1F  | 2OM       | Substrate analog |
| 3GBT   | 2W41  | ADP       | Reaction product |
| 3GD0   |       |           | No significant results |
| 3GLK   | 3RX   | S1A       | Specific ligand |
| 3GP   | 3GQQ  | APR       | Specific ligand |
| 3GQH   |       |           | No significant results |
| 3GSZ   | 3BR9  | DEY       | Specific inhibitor |
| Code | Compound | Type | Order | Notes |
|------|----------|------|-------|-------|
| 3H2G | 1GKK     | SEP  | 2nd   | Specific ligand |
| 3H38 | 4WC0     | ATP  | 1st   | Substrate    |
| 3H49 | 2FV7     | ADP  | 1st   | Reaction product |
| 3H71 | 2YA6     | DAN  | 1st   | Specific inhibitor |
| 3HBH | 2DKV     | MES  | 1st   | Substrate mimic |
| 3HS  | 2GA4     | ADE  | 1st   | Reaction product |
| 3HJ4 | 4FH1     | 3AT  | 1st   | Substrate analog |
| 3HNX | 2Y1H1    | MAN  | 1st   | Natural ligand |
| 3BC  | 3V0N     | 3GW  | 1st   | Specific inhibitor |
| 3E1H | 1QLL     | TDA  | 1st   | Specific inhibitor |
| 3IS  | 3WIC     | GNP  | 1st   | Substrate analog |
| 3K8  | 4TY1     | 39X  | 1st   | Specific inhibitor |
| 3LY  | No significant results |       |       | |
| 3IY  | 3ITV     | PSJ  | 1st   | Substrate    |
| 3U1J | 4BCC     | JKT  | 2nd   | Covalent inhibitor |
| 3YL  | 1QOR     | NDP  | 1st   | Cofactor     |
| 3K0M | 455A     | 67Z  | 1st   | Specific inhibitor |
| 3KAJ | 2WYO     | GSH  | 1st   | Product analog |
| 3KDH | 3N10     | PYV  | 1st   | Natural ligand mimic |
| 3J1E | 1Y1R1    | GSP  | 1st   | Substrate analog |
| 3KJT | 4K1M     | TY1  | 1st   | Modified amino acid |
| 3KP7 | 4E1W     | SRY  | 1st   | Specific inhibitor |
| 3K9R | 3XY1     | SAM  | 84th  | Substrate    |
| 3K7X | 3KVU     | ACO  | 1st   | Cofactor     |
| 3LG  | 3LI1     | BGC  | 1st   | Substrate    |
| 3LXZ | 1X18     | GUN  | 3rd   | Specific ligand analog |
| 3M4D | 3M1R     | BCD  | 1st   | Natural ligand |
| 3N6J | 1ECQ     | DXG  | 1st   | Substrate analog |
| 3NK6 | 3GYQ     | SAM  | 1st   | Substrate    |
| 3PTE | 1Y1S     | BSA  | 1st   | Specific inhibitor |
| 3SSW | 1ZT9     | TRP  | 1st   | Specific ligand |
| 3Z1T | 3V1U     | ANP  | 1st   | Specific ligand |
| 4AKE | 3LR3     | ADP  | 1st   | Reaction product |
| 4PGM | 1E59     | V03  | 1st   | Specific inhibitor |
| 4PT1 | 1Y1C     | ABA  | 1st   | Modified amino acid |
Libra-WA

REFERENCES

Atzeni, P., Polticelli, F., Toti, D. (2011) A framework for semi-automatic identification, disambiguation and storage of protein-related abbreviations in scientific literature, Proceedings - International Conference on Data Engineering, 5767646, 59-61

Atzeni, P., Polticelli, F., Toti, D. (2011) An automatic identification and resolution system for protein-related abbreviations in scientific papers, Lecture Notes in Computer Science, 6623, 171-176

Brylinski, M., Skolnick, J. (2008) A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. Proc Natl Acad Sci U S A, 105(1), 129-134

Capra, J.A., Laskowski, R.A., Thornton, J.M., Singh, M., Funkhouser, T.A. (2009) Predicting Protein Ligand Binding Sites by Combining Evolutionary Sequence Conservation and 3D Structure. PLoS Comput Biol, 5(12)

Furnham, N., Holliday, G.L., de Beer, T.A., Jacobsen, J.O., Pearson, W.R., Thornton, J.M. (2014) The Catalytic Site Atlas 2, Nucleic Acids Res., 42, D485-D489

Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., Schwede, T. (2013) The Protein Model Portal—a comprehensive resource for protein structure and model information, Database (Oxford), 2013:bat031.

Levy, et al., (2015) retrieved from the web, http://home.wlu.edu/levys/software/kd/

Dessailly, B.H., et al. (2008) LigASite: a database of biologically relevant binding sites in proteins with known apo-structures, Nucleic Acids Res., 36, 667-673

Laskowski, R.A., et al. (2005a) ProFunc: a server for predicting protein function from 3D structure, Nucleic Acids Res., 3, W89-W93

Laskowski, R.A., et al. (2005b) Protein function prediction using local 3D templates, J Mol Biol, 351, 614-626

Roy, A., Yang, J., Zhang, Y. (2012) COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research, 40, W471-W477

Toti, D., Atzeni, P., Polticelli, F. (2012) Automatic Protein Abbreviations Discovery and Resolution from Full-Text Scientific Papers: The PRAISED Framework. Bio-Algorithm and Med-Systems, 8, 13-

Viet Hung, L., Cattrari, S., Bizai, M., Toti, D., Polticelli, F. (2015) LIBRA: Ligand Binding Site Recognition Application, Bioinformatics, 31(24), 4020-4022

Yang, J., Roy, A., Zhang, Y. (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment, Bioinformatics, 29(20), 2588-95