

**DATABASE**

**Open Access**

# Creation of a free, Internet-accessible database: the Multiple Target Ligand Database

Chao Chen<sup>1</sup>, Yang He<sup>2</sup>, Jianhui Wu<sup>3</sup> and Jinming Zhou<sup>2\*</sup>

## Abstract

**Background:** Polypharmacology plays an important part in drug discovery, and remains a major challenge in drug development. Identification of the underlying polypharmacology of a drug, as well as development of polypharmacological drugs, have become important issues in the pharmaceutical industry and academia.

**Description:** Herein, through data mining of the Protein Data Bank (PDB), a free, Internet-accessible database called the Multiple Target Ligand Database (MTLD; [www.mtldcadd.com](http://www.mtldcadd.com)) was constructed. The MTLD contains 1,732 multiple-target ligands (MTLs) which bind to 14,996 binding sites extracted from 12,759 PDB structures. Among MTLs, 222 entries are approved drugs and 1,334 entries are drug-like compounds. The MTLD could be an extremely useful tool in the development of polypharmacological drugs. It also sheds light on the side effects of drugs through anticipation of their multiple functions and similarities in the binding sites of multiple targets. The entire database is free for online searching, browsing, and downloading.

**Conclusion:** As a crucial expansion of the PDB, increasing numbers of MTLs will be included in the MTLD. Eventually, it will become an efficient platform to obtain useful information on MTLs and their underlying polypharmacology.

**Keywords:** Polypharmacology, Multiple-target ligands, Database, Drug discovery

## Background

"Polypharmacology" (also termed "drug promiscuity") refers to the action of a single drug on multiple targets through a single pathway or multiple pathways. This phenomenon has been regarded to be the main cause for the severe adverse effects or toxicities of several drugs approved since the 1990s [1-3]. Based on the exponential growth of molecular data and rapid advances in drug development, evidence suggests that polypharmacology is also important for drug efficacy. For instance, clozapine is the "gold standard" anti-psychotic drug exhibiting beneficial effects *via* complicated interactions with multiple target networks [4]. Several highly efficacious drugs such as salicylate, metformin, or imatinib exhibit enhanced therapeutic efficacy through interactions with multiple targets simultaneously.

In general, it is accepted that the activity towards a single target is not sufficient for a complex disease involving multiple pathogenic factors (e.g., cancer, diabetes

mellitus, neurodegenerative syndrome, cardiovascular diseases). Importantly, some of the undesired side effects are due to drugs not hitting their targets, which can confer potential repurposing opportunities for these drugs and provide novel strategies in drug design.

Taken together, polypharmacology plays an important part in drug discovery and remains one of the major challenges in drug development. It opens avenues for rational design of new agents that are more efficient and less toxic than their predecessors [5-10]. Drug discovery using a polypharmacology approach has become a hot topic in the pharmaceutical industry and in academia [5-7,11].

There are hundreds of publicly available databases on drug discovery Protein Data Bank (PDB) [12], DrugBank [13], Kyoto Encyclopedia of Genes and Genomes (KEGG) [14], ZINC [15], Chemical database of European Molecular Biology Laboratory [16], and Therapeutic Target Database [17]. Such databases are key resources that integrate diverse information such as molecular pathways, crystal structures, binding experiments, side effects, and drug targets. Such information is also very useful in prospective drug design using a polypharmacology approach. However, finding information on polypharmacological agents

\* Correspondence: [zhoujinming@imb.pumc.edu.cn](mailto:zhoujinming@imb.pumc.edu.cn)

<sup>2</sup>Institute of Medicinal Biotechnology, Chinese Academy of Medical Science, Beijing, China

Full list of author information is available at the end of the article

is difficult because of the swathes of information contained in such databases. Thus, development of a novel data-mining method archiving polypharmacological information is needed.

The PDB is a repository of detailed three-dimensional (3D) structural information of proteins and other molecules, including information on the binding between ligands and proteins. It is extremely helpful in the elucidation of ligand promiscuity. Recently, based on the information obtained from the PDB, evidence indicates that similarities in binding sites among multiple proteins and the molecular complexity of a ligand could be reasons for the polypharmacology of drugs [18,19]. As a result, several datasets of multiple-target ligands (MTLs) derived from the PDB have been built by comparing the similarities of binding sites (e.g., Kahraman, Extended Kahraman, Huang) [20,21]. However, the overall entries of MTLs in these datasets are  $\leq 100$ . Through analyses of ligand promiscuity based on the PDB, an additional two datasets have been generated, containing 164 and 247 entries, respectively [18,19]. However, these datasets do not include all of the potential MTLs in the PDB, and the information in these datasets is not sufficient. Hence, a database containing all of the potential MTLs in the PDB is needed.

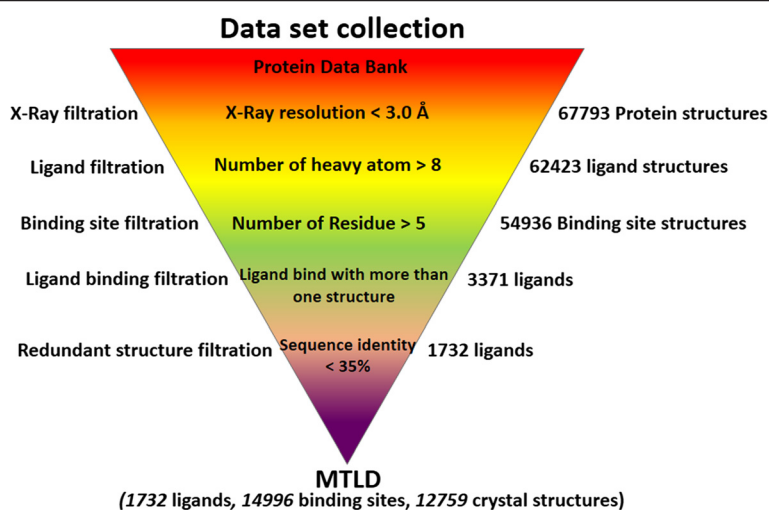
Herein, a database termed the Multiple Target Ligand Database (MTLD, [www.mtldcadd.com](http://www.mtldcadd.com)) based on 3D structural data extracted from the PDB has been constructed. The MTLD contains all of the ligands binding to MTLs and sheds light on the side effects through anticipation of their multiple functions as well as the similarities in the binding site of multiple targets. The entire database is free for online searching, browsing, and downloading. Collectively, the MTLD is extremely useful for the development of polypharmacological drugs, and provides various potential candidates for further optimization.

## Construction and content

Original structural datasets were downloaded from the File Transfer Protocol (FTP) archive (version: December 2012) of the PDB using the script “rsyncPDB.sh”. Datasets were data-mined step-by-step automatically through the command programs written in Perl language (Figure 1).

Atomic structures with a resolution of  $\geq 3.0$  Å could be inferred. Hence, X-ray protein structures with a resolution  $< 3.0$  Å (67,793 entries) from the PDB were selected for the extraction of ligands and their binding sites. To avoid selection of solvent molecules, ligands containing  $> 8$  heavy atoms were extracted from selected PDB files. As a result, 62,423 ligand coordinate files were obtained. Binding sites were defined as all of the protein residues within a radius of 6.0 Å of each atom in binding ligands. Binding sites with  $> 5$  residues were outputted, and 54,936 binding sites were extracted, which could bind with 12,138 ligands. Among these ligands, 3,371 ligands were found to bind to more than one PDB structure, and these ligands were chosen for the next filtration. To remove the redundancy of crystal structural entries bound to the same ligand, the sequence identity between protein pairs was restricted to  $< 35\%$ . Eventually, 1,732 MTLs were extracted from the PDB and archived in the MTLD.

Each ligand entry contains five pieces of information. First, the 3D structures of the ligand extracted directly from the known crystal structures are provided. Second, the two-dimensional structure of the ligand that had been converted into “SMILES” format is given. Third, the structures of the binding site that were outputted according to the coordinates of the ligand are detailed. Fourth, the original crystal structures from the PDB to which the ligand binds is given. Last, information on the sequence of the involved proteins downloaded from



**Figure 1** Dataset collection of the MTLD.

the PDB or Universal Protein Resource (UniProt; [www.uniprot.org](http://www.uniprot.org)) is provided.

Altogether, the MTLTLD comprises 1,732 MTLs,  $\approx 14.3\%$  of total unduplicated extracted ligands (12,138 entries), which bind with 14,996 binding sites from 12,759 crystal structures. Overall, the MTLTLD (Table 1) is the most comprehensive, detailed and complete database of MTLs compared with other existing databases on MTLs.

### Statistical analyses for the MTLTLD

To better understand constitution of the MTLs in the MTLTLD, statistical analyses of the MTLTLD were undertaken (Figure 2). First, the KEGG database (a database of small molecules, biopolymers, and other chemical substances relevant to biological systems) was used to analyse the relationship between MTLTLD entries and biological processes. In total, 815 MTL entries in the MTLTLD also belonged to the KEGG database ( $\approx 47.1\%$  of overall entries; Figure 2A), which includes various amino acids, saccharides, nucleotides, and lipids. Similarly, in contrast to the known drugs listed in the DrugBank, 222 approved drugs were found in the MTLTLD ( $\approx 12.8\%$  of overall entries; Figure 2B). In particular, by using the module “QuaSAR-Descriptor” included in Molecular Operating Environment (Chemical Computing Group, Montreal, Canada) according to Lipinski's rule of five, 1,334 entries were predicted to be drug-like compounds ( $\approx 76.9\%$  of overall entries; Figure 2C). Analyses of the distribution of the molecular weights of MTLs in the MTLTLD indicated that most of the MTLs had molecular weights  $< 500$  Da, and that a very small portion of MTLs had a molecular weight  $> 1000$  Da (Figure 2D). Thus, statistical analyses suggested that the MTLTLD could be highly relevant to biological processes and the action mechanism of drugs.

Among 1,732 MTLs,  $\approx 45.9\%$  ligands (795 entries) were bound to two distinct proteins (Figure 2E), which was lower than the result (65%) reported by Noé Sturm et al. [19]. This result was probably caused by using a different source of datasets when they adopted the sc-PDB (a database derived from the PDB). Notably, 222 ligands were bound to  $> 10$  proteins, including approved drugs

such as isotretinoin, spermidine, and salicylic acid. The promiscuity of a ligand is related to its conformational flexibility [19]. Hence, analyses of the conformational complexity of the extracted PDB structures for each ligand were conducted through structural alignment using Multiscale Modeling Tools for Structural Biology (Scripps Research Institute, San Diego, CA, USA). The root-mean-square deviation (RMSD) of structure pairs was calculated after the alignment. The maximal RMSD value of the structure pairs was taken to be a criterion of conformational change. Computed RMSD values of most MTLs (1,270 entries,  $\approx 73.3\%$ ) were  $< 2.0$  Å (Figure 2F), indicating that most of the MTLs could bind to different proteins by adopting a similar conformation. However, further comprehensive analyses are needed to identify the effects of other parameters such as the: molecular size and flexibility of MTLs; number of potential targets; mode of interaction between MTLs and targets.

### Internet interface of the MTLTLD

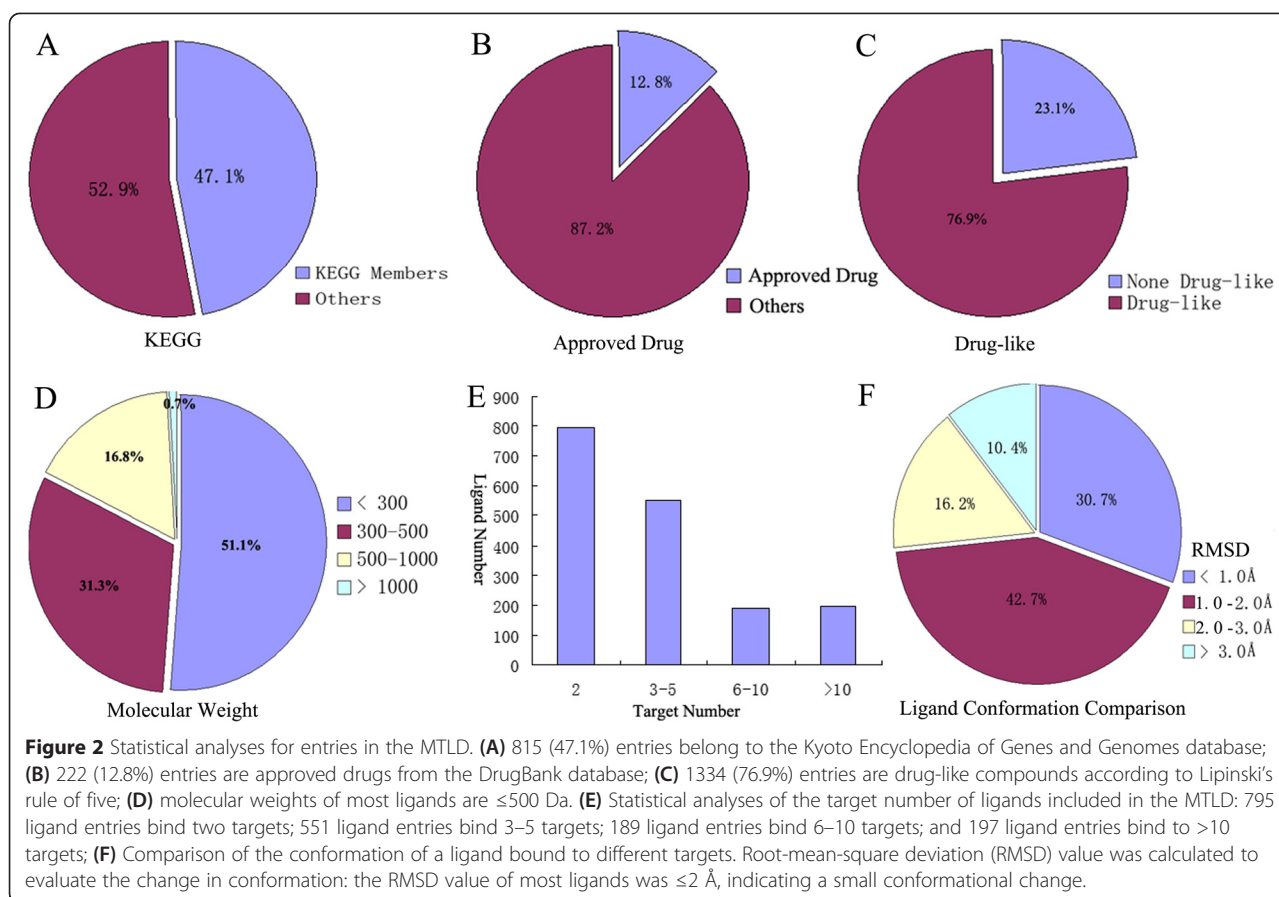
The Internet server of the MTLTLD ([mtldcadd.com](http://mtldcadd.com)) is a free, accessible database of MTLs obtained from data mining of the PDB. The Internet server was built using MySQL, Java, Javascript, and HTML languages on a machine with four 2.13-GHz processors. Java and Javascript enable the search function. Java Runtime Environment needs to be installed on the client/customer side. The 3D structures of ligands and proteins were visualized using an open-source Java viewer: Jmol. The structural similarity of ligands was calculated using the FP2 fingerprint through an open-source chemical toolkit: Open Babel.

The MTLTLD is an easily usable and fully searchable database with many built-in tools. On the MTLTLD homepage and “About” webpage, a brief introduction of the MTLTLD is given. The “Download” webpage provides the download option for all data, including approved drugs, KEGG ligands, and some kinase inhibitors. All can be downloaded conveniently. On the “Statistics” webpage, the results of statistical analyses are provided (as mentioned above) and more statistical results will be revealed on this page in the future.

The link “Search” provides three options. The first is the “Protein” option, which can be searched according to the name, PDB-ID, or UniProt-ID of proteins. For example, a query using the protein name “androgen receptor” was submitted. Five entries were presented in tabular format on the results page showing the: 3D structures of ligands; name, formula, molecular weight of the ligand; ligand-ID of the PDB. One can also proceed to each corresponding webpage of each entry with hyperlinks to other databases such as the PDB, KEGG, DrugBank, and UniProt (Figure 3A). The “Lig” option can be searched by the ligand-ID of the PDB, ligand name, or InChI key. For

**Table 1 Comparison of the MTLTLD with other datasets of multiple-target ligands**

Dataset	MTL entries	PDB structures	Accessibility
Huang	12	143	Yes
Kahraman	9	100	Yes
Extended Kahraman	10	972	Yes
Homogeneous	10	100	Yes
PDB MTL I	164	712	No
PDB MTL II	518	8166	No
<b>MTLTLD</b>	<b>1732</b>	<b>12759</b>	<b>Yes</b>



example, using salicylic acid (ligand-ID: SAL) for searching, 16 non-redundant protein targets binding with it were obtained on the results page (Figure 3B). The “structural” option enables users to draw the queried structures of ligands in the Journal Molecular Editor window. For example, users draw dihydrotestosterone as a query compound with a Tanimoto score cutoff of 0.8 (the Tanimoto score cutoff can be selected from the drop-down menu). Fifteen “hits” were presented in a tabular format on the results page (Figure 3C).

## Discussion

We have demonstrated that the MTL provides readily accessible information for MTLs (e.g., binding sites extracted from PDB structures, drug-like information, structural similarity of ligands) as well as convenient hyperlinks to databases such as UniProt, DrugBank, KEGG, and the PDB. In addition, redundancy is very common in the PDB. For example, in the PDB, dihydrotestosterone has been found to bind with 37 proteins, which belong to only three targets (Table 2). In such circumstances, the MTL exhibits the target information of MTLs clearly by filtration of redundant information. However, filtration can result in the loss of some important information, especially for some kinase families,

which have very similar amino-acid sequences. In such cases, switching off the “35% Sequence Identity Filtration” option in the search webpage gives the full list of proteins to which a ligand binds.

A crystal structure with a bound ligand does not necessarily mean firm binding. The binding affinity or binding energy of the ligand are important parameters to show if the interaction between the ligand and protein is specific, which can help to judge the “true” target of the ligand. Thus, a link to BindingDB was added on the ligand webpage, and one can find the reported data on binding affinity on the linking page. Moreover, most data on the binding affinity of complexes were not available. Hence, evaluation of the binding free energy using the X-Score method according to the complex coordinate [22] was undertaken, and the value of the binding free energy of the complex shown on the webpage. Furthermore, the crystal structures of some target classes (e.g., kinases) can be obtained much more readily than those of other target classes (e.g., G protein-coupled receptors). In such cases, a bias may be added into the MTL because the target will not be shown in the database if a crystal structure for that target has not been solved. Further enhancement is needed to try to include such targets into the MTL.

**A**

Home About **Search** Download Statistic Contact

Protein Lig Structural

androgen receptor

Result

Showing 1 of 5 Targets

Sort: PdbName Page 1 Of 1 GO

Proteinname: ANDROGEN RECEPTOR  
LigId: DHT  
Expression: C<sub>19</sub>H<sub>30</sub>O<sub>2</sub>  
Relevance: 290.45  
Download:

**B**

Protein **Lig** Structural

SAL

Result

Showing 1 of 16 Targets

Sort: Pdb-Name Page 1 Of 4 GO

LigId: SAL  
Expression: C<sub>7</sub>H<sub>6</sub>O<sub>3</sub>  
Molecular weight: 138.12  
Functions: null  
DrugID: KeggID

PDB Name: MULTIPLE ANTIBIOTIC RESISTANCE PROTEIN Binding site: P27245\_1/GS\_SAL\_bindsite

**C**

Protein Lig **Structural**

- Search Products -

Tanimoto: 0.80 Search

Result

Showing 1 of 15 Targets

Sort: Lig-Id Page 1 Of 3 GO

LigId: DHT  
Expression: C<sub>19</sub>H<sub>30</sub>O<sub>2</sub>  
Relevance: 290.45  
Tanimoto: 1.0

**Figure 3** Three examples of searches of MTLs on the MTL D Internet server: **(A)** ligands that bind to the androgen receptor were queried in the “Protein” option; **(B)** ligands with the ligand-ID of “SAL” were queried in the “Lig” option; **(C)** ligands with structural similarity to that of dihydrotestosterone were queried in the “Structural” option.

## Utility

Most of the entries in the MTL D are based on drugs. Hence, the MTL D should be useful for developing polypharmacological agents, and may provide potential candidates for further optimization. For example, estrogen receptor- $\alpha$  is a drug target for treating breast cancer, and 17-hydroxysteroid dehydrogenase (17HSD1) is a

putative target for endocrine therapy of hormone-dependent breast cancer [23]. By searching the MTL D, a ligand, estrogen (ligand-ID: EST) binding to 17HSD1 (PDB-ID: 1FDS) and estrogen receptor (PDB-ID: 3Q95) was found. Thus, polypharmacological drugs that act on both targets could be designed based on the structure of estrogen. Conversely, software called Binding Site Match



**Table 2 Multiple targets of selected drugs and natural products in the MTLT**

Ligand (ligand-ID)	Target number	Targets (PDB-ID)
Sutent (B49)	4	Mast/stem cell growth factor (3G0E) Phosphorylase b kinase gamma (2Y7J) Cyclin-dependent kinase 2 (3TI1) Tyrosine-protein kinase ITK/TSK (3MIY)
Imatinib (STI)	5	Homo sapiens v-kit Hardy-Zuckerman (1 T46) Ribosyldihydronicotinamide dehydrogenase (3FW1) Tyrosine-protein kinase SYK (1XBB) Mitogen-activated protein kinase 14 (3HEC) Bcr-Abl protein
Sorafenib (BAX)	4	B-Raf proto-oncogene serine/threonine-protein (1UWH) Vascular endothelial growth factor (4ASD) Cyclin-dependent kinase 8 (3RGF) Mitogen-activated protein kinase 14 (3GCS)
Dihydrotestosterone (DHT)	3	Sex hormone-binding globulin (1D2S) Estradiol 17-beta-dehydrogenase 1 (3KLM) Androgen receptor (1I37)
Estradiol (EST)	5	Short chain 3-hydroxyacyl-CoA dehydrogenase (1E6W) Sex hormone-binding globulin (1LHU) 17-beta-hydroxysteroid-dehydrogenase (1FDS) Estrogen sulfotransferase (1AQU) Estrogen receptor beta (2J7X)
Triiodothyronine (T3)	4	Transthyretin (1SN5) Thyroid receptor alpha (THRA) protein (2H79) Androgen receptor (2PIV) Proliferating cell nuclear antigen (3VKX)

**Table 2 Multiple targets of selected drugs and natural products in the MTLT (Continued)**

Resveratrol (STL)	7	Sulfotransferase family cytosolic 1B (3CKL) Transthyretin (1DVS) Myosin-2 heavy chain (3MNQ) Leukotriene A-4 hydrolase (3FTS) NRH dehydrogenase [quinone] 2 (1SG0) Protein (chalcone synthase) (1CGZ) Methionine adenosyltransferase 2 subunit (2YDX)
Epigallocatechin gallate (EGCG; KDH)	3	Polymerase subunit PA (4AWM) Transthyretin (3NG5) Peptidyl-prolyl <i>cis-trans</i> isomerase NIMA-inte (3OOB)
Theophylline (TEP)	3	Chitinase Apolipoprotein A-I-binding protein Pyridoxal kinase
Pioglitazone (P1B)	2	Peroxisome proliferator-activated receptor ga Amine oxidase [flavin-containing] b

Maker was developed, which could align binding sites according to their similar physiochemical properties and generate common sites. Binding Site Match Maker could be another option for the design of polypharmacological drugs.

The MTLT can shed light on the multiple mechanisms of action of drugs or natural products. For example, imatinib is an efficacious drug for the treatment of chronic myeloid leukaemia (CML). Imatinib prevents Bcr-Abl protein from exerting its actions in the oncogenic pathway in CML [24]. By searching the MTLT, apart from Bcr-Abl, imatinib was found to bind to mitogen-activated protein kinase 14, ribosyldihydronicotinamide dehydrogenase, tyrosine-protein kinase Syk, and c-Kit kinase. The natural product resveratrol (which is present in red wine) exhibits considerable chemical diversity and biological activities [25]. In the MTLT, resveratrol was found to bind with seven targets. More examples of drugs and natural products are listed in Table 2.

Information obtained from the MTLT can be used to address the mechanism of action of the adverse side effects of drugs. For example, the methylxanthine derivative

theophylline (used to treat the symptoms of reversible airflow obstruction) can cause headaches, agitation, and other adverse neuronal side effects. Upon searching the MTLDB, pyridoxal kinase was found to be one of the targets of theophylline, which could be a possible underlying mechanism of the neurotoxic effects of theophylline [26]. Likewise, pioglitazone (used for treating type-2 diabetes mellitus) exhibits neuronal side effects, including central nervous system (CNS) depression. Upon searching the MTLDB, it was found that pioglitazone can bind to monoamine oxidase B, which may be responsible for CNS depression [27].

Similarities in the binding site have important roles in polypharmacology [18]. Several methodologies have been developed to evaluate similarities in binding sites, such as SiteComp [28] and MultiBind [29]. These methodologies provide useful tools to predict the binding-site similarity of proteins but because of their incompleteness, need to be improved [30]. In such circumstances, the MTLDB could provide various binding sites as the training sets or testing sets for the development of novel methodologies for comparisons of binding sites.

## Conclusion

Here we describe development of a comprehensive, Internet-accessible database called the MTLDB based on datasets extracted from the PDB. To date, the MTLDB comprises 1,732 MTLs that bind to 14,996 binding sites extracted from 12,759 PDB structures. In the MTLDB, the 222 entries are approved drugs and 1,334 entries are drug-like compounds. Thus, the MTLDB could be extremely helpful for developing polypharmacological drugs and could provide potential candidates for further optimization. Moreover, the MTLDB may shed light on the: side effects of drugs; multiple functions of small biological molecules; similarities in binding site of target proteins. As a crucial expansion of the PDB, increasing numbers of MTLs will be included in the MTLDB, which will become an efficient platform to obtain useful information on MTLs.

## Availability and requirements

MTLDB is freely accessible at <http://www.mtdcadd.com>. The data in MTLDB is free for search, download, and further analysis.

## Abbreviations

PDB: Protein Data Bank; MTLs: Multiple Target Ligands; MTLDB: Multiple Target Ligand Database; KEGG: Kyoto Encyclopedia of Genes and Genomes; TTD: Therapeutic target database; DS: Discover studio; RMSD: Root mean square deviation; MMTSB: Multiscale modeling tools for structural biology; ERα: Estrogen receptor α; 17-HSD1: 17-hydroxysteroid dehydrogenase type 1.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

JHW and JMZ conceived the study. CC and JMZ constructed the database and drafted the manuscript. CC and JMZ designed and developed the website. YH, CC, JMZ participated in dataset collecting and processing. All authors read and agreed to the final manuscript.

## Acknowledgements

We thank Mr. Ruiming Li at Institute of Medicinal Biotechnology, Chinese Academy of Medical Science, Beijing, China for the kindly helps at the construction of the website. This work was in part supported by the Nature Science Foundation of China (81311120299, 81271844) and Fundamental Research Funds for the Central Universities of China (3142014100).

## Author details

<sup>1</sup>North China Institute of Science and Technology, Beijing, China. <sup>2</sup>Institute of Medicinal Biotechnology, Chinese Academy of Medical Science, Beijing, China. <sup>3</sup>Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada.

Received: 28 May 2014 Accepted: 30 March 2015

Published online: 15 April 2015

## References

1. Connolly HM, Crary JL, McGoon MD, Hensrud DD, Edwards BS, Edwards WD, et al. Valvular heart disease associated with fenfluramine-phentermine. *N Engl J Med*. 1997;337(9):581–8.
2. Hutcheson JD, Setola V, Roth BL, Merryman WD. Serotonin receptors and heart valve disease—it was meant 2B. *Pharmacol Ther*. 2011;132(2):146–57.
3. Rothman RB, Baumann MH, Savage JE, Rauser L, McBride A, Hufeisen SJ, et al. Evidence for possible involvement of 5-HT(2B) receptors in the cardiac valvulopathy associated with fenfluramine and other serotonergic medications. *Circulation*. 2000;102(23):2836–41.
4. Kane J, Honigfeld G, Singer J, Meltzer H. Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives General Psychiatry*. 1988;45(9):789–96.
5. Costantino L, Barlocco D. Designed multiple ligands: basic research vs clinical outcomes. *Curr Med Chem*. 2012;19(20):3353–87.
6. Medina-Franco JL, Giulianotti MA, Welmaker GS, Houghten RA. Shifting from the single to the multitarget paradigm in drug discovery. *Drug Discov Today*. 2013;18(9–10):495–501.
7. Morphy R, Rankovic Z. Designed multiple ligands. An emerging drug discovery paradigm. *J Medicinal Chem*. 2005;48(21):6523–43.
8. Peters JU. Polypharmacology - foe or friend? *J Med Chem*. 2013;56(22):8955–71.
9. Proschak E. Reconsidering the drug discovery pipeline for designed multitarget drugs. *Drug Discov Today*. 2013;18(23–24):1129–30.
10. Reddy AS, Zhang S. Polypharmacology: drug discovery for the future. *Expert Rev Clin Pharmacol*. 2013;6(1):41–7.
11. Moya-Garcia AA, Ranea JA. Insights into polypharmacology from drug-domain associations. *Bioinformatics*. 2013;29(16):1934–7.
12. Rose PW, Beran B, Bi C, Bluhm WF, Dimitropoulos D, Goodsell DS, et al. The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res*. 2011;39(Database issue):D392–401.
13. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, et al. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res*. 2011;39(Database issue):D1035–41.
14. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res*. 2014;42(1):D199–205.
15. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model*. 2012;52(7):1757–68.
16. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res*. 2012;40(Database issue):D1100–7.
17. Zhu F, Shi Z, Qin C, Tao L, Liu X, Xu F, et al. Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res*. 2012;40(Database issue):D1128–36.
18. Haupt VJ, Daminelli S, Schroeder M. Drug Promiscuity in PDB: Protein Binding Site Similarity Is Key. *PLoS One*. 2013;8(6):e65894.

19. Sturm N, Desaphy J, Quinn RJ, Rognan D, Kellenberger E. Structural insights into the molecular basis of the ligand promiscuity. *J Chem Inf Model*. 2012;52(9):2410–21.
20. Hoffmann B, Zaslavskiy M, Vert JP, Stoven V. A new protein binding pocket similarity measure based on comparison of clouds of atoms in 3D: application to ligand prediction. *BMC Bioinformatics*. 2010;11:99.
21. Sael L, Kihara D. Detecting local ligand-binding site similarity in nonhomologous proteins by surface patch comparison. *Proteins*. 2012;80(4):1177–95.
22. Wang R, Lai L, Wang S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comput Aided Mol Des*. 2002;16(1):11–26.
23. Day JM, Foster PA, Tutill HJ, Parsons MF, Newman SP, Chander SK, et al. 17 $\beta$ -hydroxysteroid dehydrogenase Type 1, and not Type 12, is a target for endocrine therapy of hormone-dependent breast cancer. *Int J Cancer J Int du Cancer*. 2008;122(9):1931–40.
24. Sacha T. Imatinib in Chronic Myeloid Leukemia: an Overview. *Mediterranean J Hematol Infectious Diseases*. 2014;6(1):e2014007.
25. Joseph JD, Lu N, Qian J, Sensintaffar J, Shao G, Brigham D, et al. A clinically relevant androgen receptor mutation confers resistance to second-generation antiandrogens enzalutamide and ARN-509. *Cancer Discovery*. 2013;3(9):1020–9.
26. Gandhi AK, Desai JV, Ghatge MS, di Salvo ML, Di Biase S, Danso-Danquah R, et al. Crystal structures of human pyridoxal kinase in complex with the neurotoxins, ginkgotoxin and theophylline: insights into pyridoxal kinase inhibition. *PLoS One*. 2012;7(7), e40954.
27. Binda C, Aldeco M, Geldenhuys WJ, Tortorici M, Mattevi A, Edmondson DE. Molecular Insights into Human Monoamine Oxidase B Inhibition by the Glitazone Anti-Diabetes Drugs. *ACS Med Chem Lett*. 2011;3(1):39–42.
28. Lin Y, Yoo S, Sanchez R. SiteComp: a server for ligand binding site analysis in protein structures. *Bioinformatics*. 2012;28(8):1172–3.
29. Shulman-Peleg A, Shatsky M, Nussinov R, Wolfson HJ. MultiBind and MAPPI: web servers for multiple alignment of protein 3D-binding sites and their interactions. *Nucleic Acids Res*. 2008;36(Web Server issue):W260–4.
30. Rognan EKCS. How to Measure the Similarity Between Protein Ligand-Binding Sites? *Current Computer Aided-Drug Design*. 2008;4(3):12.

Publish with **ChemistryCentral** and every scientist can read your work free of charge

*“Open access provides opportunities to our colleagues in other parts of the globe, by allowing anyone to view the content free of charge.”*

W. Jeffery Hurst, The Hershey Company.

- available free of charge to the entire scientific community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
<http://www.chemistrycentral.com/manuscript/>



**ChemistryCentral**