W350–W356 *Nucleic Acids Research, 2019, Vol. 47, Web Server issue* *Published online 20 May 2019*
doi: 10.1093/nar/gkz391

Drug ReposER: a web server for predicting similar amino acid arrangements to known drug binding interfaces for potential drug repositioning

Nur Syatila Ab Ghani[1], Effirul Ikhwan Ramlan[2,3] and Mohd Firdaus-Raih[1,4,*]

[1]Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia, [2]School of Computing, Engineering and Intelligent Systems, Ulster University, Northlands Road, Magee Campus, Londonderry BT48 7JL, UK, [3]Malaysia Genome Institute, National Institutes of Biotechnology Malaysia, Jalan Bangi, 43000 Kajang, Selangor, Malaysia and [4]Institute of Systems Biology, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia

Received February 27, 2019; Revised April 24, 2019; Editorial Decision April 26, 2019; Accepted May 02, 2019

ABSTRACT

A common drug repositioning strategy is the re-application of an existing drug to address alternative targets. A crucial aspect to enable such repurposing is that the drug's binding site on the original target is similar to that on the alternative target. Based on the assumption that proteins with similar binding sites may bind to similar drugs, the 3D substructure similarity data can be used to identify similar sites in other proteins that are not known targets. The Drug ReposER (DRug REPOSitioning Exploration Resource) web server is designed to identify potential targets for drug repurposing based on sub-structural similarity to the binding interfaces of known drug binding sites. The application has precomputed amino acid arrangements from protein structures in the Protein Data Bank that are similar to the 3D arrangements of known drug binding sites thus allowing users to explore them as alternative targets. Users can annotate new structures for sites that are similarly arranged to the residues found in known drug binding interfaces. The search results are presented as mappings of matched sidechain superpositions. The results of the searches can be visualized using an integrated NGL viewer. The Drug ReposER server has no access restrictions and is available at http://mfrlab.org/drugreposer/.

INTRODUCTION

The introduction of new approved drugs to the market is both a time consuming and costly endeavour. One approach to circumvent the lengthy development processes and to significantly reduce the costs is to repurpose drugs that are already in use for other indications (1). Despite numerous efforts aimed at the discovery of new drugs, their success rate is rather nominal with the majority of the compounds never reaching clinical trials or failing at the final clinical trials stages (2). The stringent considerations that need to be met for the approval of new therapeutic compounds has resulted in increased interest being directed at the repurposing or repositioning of already approved drugs (3). Both academia and industry have worked together towards proposing new indications for existing drugs with proven track records that would in turn allow for better cost optimization compared to developing and testing new compounds (4).

Thalidomide and sildenafil are two examples of successful drug repurposing initiatives. Thalidomide had been approved by U.S. Food and Drug Administration (FDA) in 1950s and primarily used for morning sickness during the first trimester of pregnancy. It was later withdrawn in 1961 due to teratogenic effects (5) but was recently reintroduced for other indications including multiple myeloma (6) and leprosy (7). Sildenafil, which targets cyclic GMP phosphodiesterase 5 (PDE5), had first been discovered as an antihypertensive due to its vasodilatory effects and was successfully repurposed for erectile dysfunction and pulmonary arterial hypertension (8). However, in the case of sildenafil, the repurposing efforts have focused on different indications that are treatable by targeting PDE-5 (9) and therefore does not involve an alternative target for the drug.

However, drug repositioning can also happen for cases where an approved compound may have an alternative target. The mode of interaction and atomic level details by which a drug exerts its effects are important aspects in understanding its therapeutic mechanism. Knowledge on the ability of a drug molecule to bind to alternate targets and the role of that protein in a particular disease mechanism or pathway is also crucial when investigating the potential off-targets and adverse effects (10) of the candidate. As a

*To whom correspondence should be addressed. Tel: +6 03 89267446; Fax: +6 03 89267972; Email: firdaus@mfrlab.org

result of this understanding, there are numerous entries in the Protein Data Bank (11) of target proteins bound to drug compounds.

The availability of these protein drug complexes allows information about the amino acid residues that are interacting with the drug compounds to be extracted thus making it possible to know the specific 3D arrangements of the amino acids partaking in the interaction. Previously, computer programs such as ASSAM, SPRITE (12) and IMAAAGINE (13) have shown that approaches using comparisons of the amino acid side chain arrangements have certain advantages over approaches that only consider the carbon backbone positions. In the context of protein–drug complexes, it is expected that the majority of the interactions with the drugs will involve the side chain. By modifying such search programs, it is therefore possible to identify other amino acid 3D arrangements that are similar to those known to be involved in the drug interactions.

Proteins containing amino acid arrangements that are similar to that of a known drug binding site may potentially bind to the same drug compound or a similarly structured molecule, and thus such a site can be proposed as a target for drug repositioning. The occurrences of similarly arranged amino acid residues among proteins of different folds and functions are not expected to be as common as between homologs. When such similarities are found to be present in non-homologous proteins, it can be a useful starting point to find unrelated targets for existing drugs and thus allowing for different targets to be explored for an existing compound. In essence, large-scale analysis of such binding site similarities has the potential to yield alternative targets that can pave the way for drug repositioning and also investigations into toxicity or side effects of known drugs.

Currently available resources for drug discovery and drug repurposing

There are various databases currently available for the annotation of drugs, targets and drug–target interactions, such as DrugBank (14) and Therapeutic Target Database (TTD) (15). These databases provide access to candidate compounds with comprehensive information attached and are thus useful for drug design and development. These resources are based on literature analysis of experimental validation and clinical data, as well as information consolidated from other curated databases and regulatory agencies. Several drug repurposing databases have also been developed to facilitate the exploration of approved drugs currently undergoing clinical trials for new indications. For example, RepurposeDB compiles a collection of drugs, drug targets, diseases, and reports of successful drug repositioning activities (16) and RepoDB stores approved and failed drugs from experiments and clinical trials (17). These databases provide benchmarks and valuable insights indicative of the success rate of repurposing target prediction from computational predictions up to the point of commercial availability of the drug.

A number of freely available web servers offer users the ability to explore repurposing opportunities by providing either access to an interface for the virtual screening of small molecules (18), access to a searching function for similar compounds (17–19), or access to a search function for similar binding sites (20,21). Local structural similarity searches for drug binding sites of proteins in the PDB had been implemented by various web servers with varying degrees of success in exploring the potential for repurposing approved drugs (22–24).

The Drug ReposER (DRug REPOSitioning Exploration Resource) web server facilitates the searching of potential alternative targets for known compounds by comparing the similarities of the 3D amino acid arrangements in known targets to other structures available in the PDB. While there are several other databases and web servers that carry the 'drug repurposing' or 'drug repositioning' label, none of them utilize a structural comparison approach of finding similar binding sites as a means to identify and explore potential targets for drug repositioning. For example, other available drug repurposing resources only provide information on the existing status of drugs for repurposing or serve as a repository of repurposed drugs and their targets but do not allow users to actively explore new targets for repositioning or allow investigations into optimizing or modifying existing compounds to interact with a new target with a structural context.

METHODOLOGY

Datasets: Known binding sites for approved drugs

The list of approved drugs used in Drug ReposER were extracted from the 'Drug and Drug Target Mapping' interface in the RCSB PDB database (11); 3203 protein structures bound to 512 selected drug molecules were used to extract patterns containing residues that were defined as binding because there were within 4.0 Å or less distance from the atoms of a molecule that had been classified as a drug. The patterns include those from 66 PDB structures of membrane proteins that are known to bind to various drugs. These binding site patterns and their annotations have been made available via the Drug ReposER web server. The binding site patterns were also compiled into a database of 3D patterns that allow users to identify the presence of similarly arranged amino acids in a query protein structure using a modified version of the SPRITE search engine (12).

Similarly arranged patterns of amino acids predicted to bind to approved drugs

The 3D patterns of known drug binding sites were used as queries to search the PDB for other proteins that contained similarly arranged sites using the ASSAM computer program previously described in Nadzirin *et al*. (12). The program compares each binding site pattern against a non-redundant PDB dataset at a 90% sequence identity cut-off. The searches resulted in more than 100 000 amino acid 3D arrangements that had varying degrees of similarity to amino acid tertiary arrangements that formed the binding sites to known drug compounds. From this total, 689 PDB structures of membrane proteins were found to contain amino acid tertiary arrangements that are similar to known drug binding sites. This dataset is searchable via the Drug ReposER web interface. No further filtering was done on the dataset to reduce its number. As a result, the

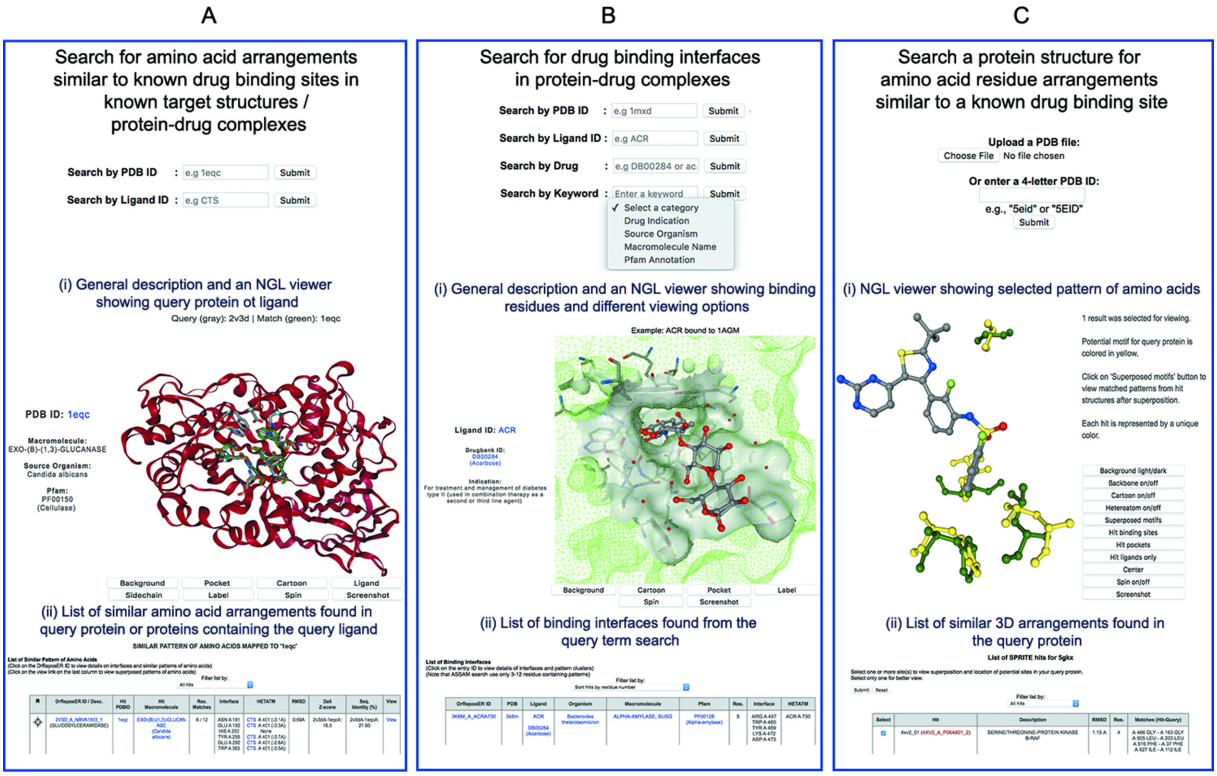

Figure 1. The Drug ReposER search interfaces. (**A**) Searching for amino acid arrangements in PDB structures that are similar to known drug binding sites. (**B**) Searching for drug binding interfaces in protein–drug complexes. (**C**) Searching for residue arrangements in query structure similar to a known drug binding site.

dataset will also contain examples of matches to a known drug binding site for proteins that are not obvious drug targets. We feel that this provides the user with options to look for homologs for which structures are not available in other organisms such as humans. It is therefore the user's prerogative to provide the context to explore the drug repositioning strategy. Users can however carry out their own filtering such as by setting parameters for the number of matches or the superposed 3D similarity of the sites as measured in root mean square deviation (rmsd) values.

Web development and infrastructure

The corresponding database used in the web server was generated using MySQL (version 5.1.44) consisting of two interconnected tables, while all web interfaces and results page were developed using PHP (version 5.3.1) and Python (version 2.7). The Drug ReposER server is hosted on a Linux server (CentOS) physically located at the Malaysia Genome Institute, Malaysia. The server is equipped with dual Intel Xeon processors and 64 GB of memory. The Drug ReposER web application was successfully tested on various platforms including Windows and Mac OS.

The Drug ReposER user interface

The Drug ReposER web server consists of three search options: (i) searching for amino acid 3D arrangements that are similarly postured to known drug binding sites: (ii) searching for 3D structure information of binding sites for approved drugs as retrieved from protein-drug complexes in the PDB; and; (iii) searching a user provided structure for the presence of sites that are similarly arranged to those of known drug binding sites (Supplementary Figure S1). The integrated SPRITE search engine had been reported by Nadzirin *et al*. (12) but was modified with a specific search database of known drug binding sites for use in the Drug ReposER application.

Search for amino acid arrangements similar to known drug binding sites in PDB structures

The Drug ReposER application can be used to determine whether a PDB protein structure contains 3D arrangements of amino acids that are similar to that for known drug binding sites. Users can opt to execute this type of search by providing a PDB ID or PDB ligand ID input that returns a results page with an NGL viewer showing the 3D structure of the query protein or ligand, and a list of similar amino acid arrangements that matches the query term (Figure 1A). The interface searches the query term against a database of 3D amino acid arrangements similar to known drug binding sites, obtained from a non-redundant PDB dataset (Supplementary Figure S1). If the query PDB ID is available in the database, the interface returns all 3D amino acid arrangements (query pattern) found in query protein that are similarly arranged to that of known drug binding sites (hit pattern) and presented together with information regarding any nearby heteroatoms, such as bound annotated drugs, experimental drugs, or metal ions that are at

the vicinity of the query patterns within a distance cut-off of 4.0Å. The presence of nearby heteroatoms can also be searched through the ligand based search by providing a Ligand ID query. Users can also opt to view superpositions of amino acid arrangements in the queries and those from the references using the NGL viewer by clicking the 'View' link available on the last column of the table. Meanwhile, clicking on the DrReposER ID on the first column opens the corresponding results page for the selected binding site.

Search for drug binding interfaces from protein-drug complexes

The interface also allows users to search for known binding sites to approved drugs derived from the PDB, by providing a query term for searching the data set of binding sites for approved drugs. A query of PDB ID, ligand ID, drug name or DrugBank ID returns a list of binding interfaces matching the query term in an NGL viewer (25) showing the 3D structure of the query protein or ligand molecule (Figure 1B). Although there are some similarities of this feature to other services available, they were also integrated in order to allow for users to contain their analysis within the Drug ReposER interface as much as possible without needing to exit into other resources. There are two search options available for this feature; one produces the structural information regarding the protein and the other presents the drug indication as annotated from the DrugBank for the given drug molecule (Figure 1B). Users may also search using keyword queries under different categories that include drug indication, organism, protein name and Pfam annotation (26). Such an interrogation returns a list of binding sites matching the keywords. Clicking a DrReposER ID on the first column returns a results page for the individual data associated with that ID that describes the binding site and a listing for a set of similarly arranged sites in other proteins, along with the matched protein structure's RMSD value from the superposition, the *Z*-score from fold comparison using DaliLite (27) and sequence identity values from sequence comparison using MUSCLE program (28). Users can then click on the 'View' link on the last column to visualize the superposed patterns and compare both the query and the matches.

Search for residue arrangements in query structure similar to known drug binding sites

In cases when the structure of interest is not currently available in the PDB or the dataset used, the user can upload the coordinates of their own protein structure. This is another Drug ReposER specific feature that is unavailable at other similarly intended services. After submitting the query, the SPRITE program searches the query protein against a database of known drug binding sites (12). A typical SPRITE search, although dependent on the current server load and size of the input structure, would typically approximately two minutes to execute. Once the search is completed, users can select to view results for the left-handed and right-handed superposition (12) of matched patterns of amino acids in the query structure (query pattern) against known drug binding sites (hit pattern) (Figure 1C).

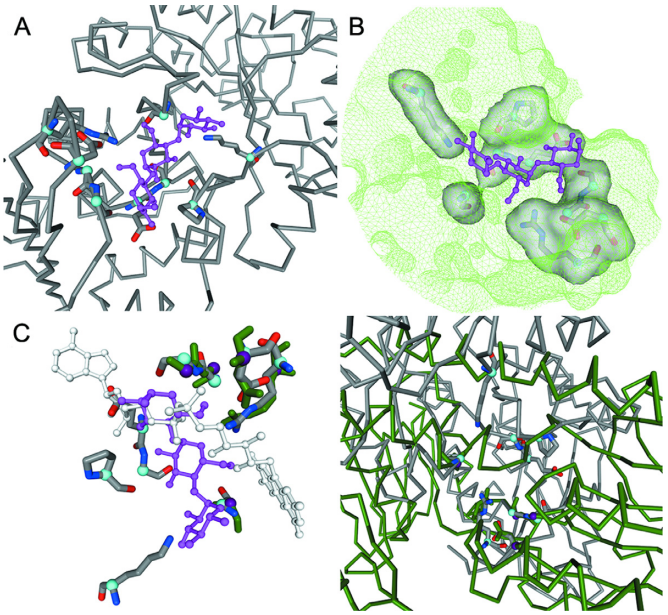

Figure 2. Multiple screenshots of the embedded NGL viewer in the Drug ReposER server. (**A**) Binding site for acarbose from alpha amylase is represented in licorice, while the protein structure can be viewed in backbone representation. (**B**) Binding pocket for acarbose (ACR) from alpha amylase (1mxd) is shown in surface representation and coloured green. (**C**) Superposed patterns of amino acids from the known binding site for acarbose (1mxd, grey) and similarly arranged pattern of amino acids from glutathione reductase (2hqm, green) bound to flavin adenine dinucleotide (FAD); superposed ligands can also be viewed; ACR is coloured in purple and FAD is coloured in white. (**D**) Superposed protein structures of 1mxd (grey) and 2hqm (green) are shown in backbone representation.

Visualization of structures and superpositions in the NGL viewer

The Drug ReposER web server integrates the NGL viewer (25) for visualization of 3D structures and for visual inspection of the superposed patterns of amino acids. The superposed binding residues can be viewed upon clicking the 'View' link in the 'View' column. Binding residues and drug molecules from the query protein are shown in grey and magenta respectively, while the matched pattern is shown in green. The superposition of query and matched ligands can be viewed upon clicking the 'Ligand on/off' button, while the superposition of their backbone structures can be viewed by clicking the 'Backbone on/off' button. Various pre-set display options are also provided, including changing the background colour, the depiction of ligand binding pockets, and indication of side chain atoms (Figure 2). Users are also given the options to download screenshots of the viewer window.

CASE STUDY: AN EXAMPLE PROCESS FOR DISCOVERING POTENTIAL TARGETS FOR DRUG REPOSITIONING

Repositioning of acetylcysteine for tuberculosis

In this case study, we avoid a direct example that proposes an alternative target for drug repositioning. We instead demonstrate the utility of the application in propos-

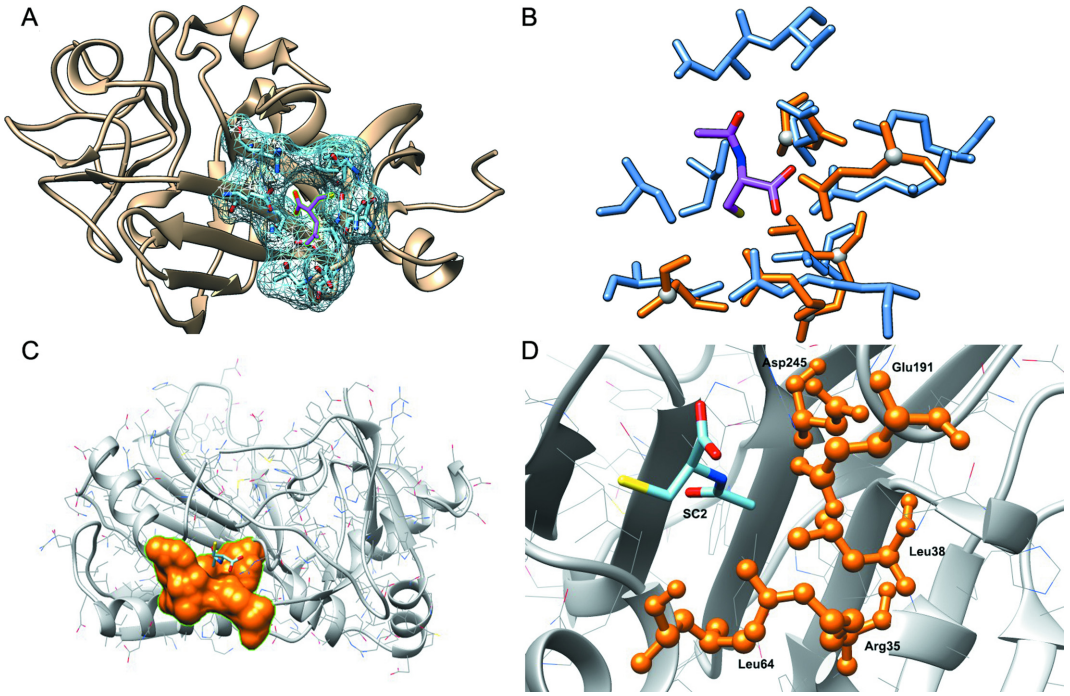

Figure 3. Results of molecular docking prediction performed by SwissDock. (**A**) Binding site for acetylcysteine found in human Ficolin-2 (PDBID: 2j1g). Binding site is coloured blue with surface representation. (**B**) Superposition of binding site for acetylcysteine in human Ficolin-2 (PDBID: 2j1g, light blue colour) to five-residue pattern from Rv3406 protein predicted from the server (PDBID: 4cvy, orange color). (**C**) Model of acetylcysteine against Rv3406 protein from *Mycobacterium tuberculosis*, with ΔG score of −6.34 kcal/mol. (**D**) Predicted binding site for acetylcysteine from Drug ReposER web server is indicated in orange and the conformation of acetylcysteine predicted from the SwissDock is coloured in light blue.

ing possible mechanisms brought about by repositioning of a known drug that is also supported by clinical trials data. Acetylcysteine (PDB ligand: SC2) is an approved small molecule mucolytic agent that acts by reducing mucus thickness (29), and is also used for the management of acetaminophen poisoning by maintaining the levels of glutathione in the liver (30). The PDB contains several protein structures that are in complex with acetylcysteine. One example, human Ficolin-2 (PDBID: 2j1g) (31), is a protein that recognizes various molecular patterns from pathogen surfaces, such as acetylated carbohydrates and other acetylated compounds and is thus important for the innate immune system (31). Acetylcysteine is known to inhibit the binding of Ficolin-2 to *Streptococcus pneumoniae* (32). L-ficolin had been reported to block the infection of *Mycobacterium tuberculosis* H37Rv in lung cells by binding to the glycolipid region of the bacterial surface (33).

To explore potential new targets for acetylcysteine, we searched for amino acid arrangements that are similar to the acetylcysteine binding site in Ficolin-2 (Figure 3A) using the structure 2j1g as a query. We then selected the Drug ReposER entry 2J1G_F_SC2F1290 for further analysis. One of the similarly arranged amino acid patterns is a five residue pattern in sulfate ester dioxygenase Rv3406 from *M. tuberculosis* H37Rv (PDBID: 4cvy) (34) (Figure 3B). The Rv3406 protein does not share any detectable sequence or structural similarity to human Ficolin-2, but the report from Luo and colleagues clearly indicates the role of Ficolin-2 in defence against *M. tuberculosis* H37Rv infection (33). Molecular docking using the SwissDock web server (35) resulted in several conformations for Rv3406-acetylcysteine binding similarly as presented from the arrangement similarity retrieved by the Drug ReposER server (Figure 3C and D), thus suggesting the Rv3406 protein can serve as an alternative target for acetylcysteine.

Amaral and colleagues had proven the anti-mycobacterial activity of N-acetylcysteine in *M. tuberculosis* through *in-vitro* and *in-vivo* experiments that revealed a reduction of oxidation stress and mycobacterial loads, respectively (36). Due to its antioxidant effects, the ClinicalTrials.gov server (37) also reported an ongoing phase II randomized clinical trial to test the potential role of acetylcysteine as an adjuvant therapy for tuberculosis (ClinicalTrials.gov identifier: NCT03281226), thus suggesting the potential role of acetylcysteine as a tuberculosis drug. Analysis with Drug ReposER can put into context the mechanisms that may be in effect as observed for the clinical trials. As a mucus reducing agent, it is expected that acetylcysteine would be used for patients with tuberculosis. However, acetylcysteine can also target Ficolin-2 and inadvertently affect the innate immune system response. Despite this fact, the clinical trials revealed that acetylcysteine may have other targets that negates that function and is thus able to produce a net anti-mycobacterial effect.

SUMMARY

The Drug ReposER server provides comprehensive information of known drug binding sites that are deployed to identify potentially alternative targets for drug repositioning. The use of the ASSAM and SPRITE computer

programs provide a high-throughput but fast whole PDB search capacity that has not been previously explored by other resources. Furthermore, many previous drug repurposing efforts have concentrated on one compound of interest as a starting point to identify other indications for which it may have therapeutic effects. The outputs provided by the Drug ReposER server can facilitate a larger scale analysis of protein-drug interactions that can guide experimental validation of potentially new protein-drug interactions that is not limited to exploring a single compound of interest as a starting point. The Drug ReposER application is therefore useful for the identification of potentially new targets for existing drugs and/or the rational modification of existing compounds to be fitted to new targets that have local or substructure similarities to that of an existing target site, as well as allowing for investigations into possible toxicity and off-target interactions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Malaysia Genome Institute for the use of computational resources.

FUNDING

Universiti Kebangsaan Malaysia [DIP-2017-013]; Ministry of Science, Technology and Innovation Malaysia [02-01-02-SF1278]. Funding for open access charge: Universiti Kebangsaan Malaysia internal funds.
Conflict of interest statement. None declared.

REFERENCES

1. Nosengo,N. (2016) Can you teach old drugs new tricks? *Nature*, **534**, 314–316.
2. Chartier,M., Morency,L.-P., Zylber,M.I. and Najmanovich,R.J. (2017) Large-scale detection of drug off-targets: hypotheses for drug repurposing and understanding side-effects. *BMC Pharmacol. Toxicol.*, **18**, 18.
3. Ashburn,T.T. and Thor,K.B. (2004) Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discov.*, **3**, 673–683.
4. Oprea,T.I., Bauman,J.E., Bologa,C.G., Buranda,T., Chigaev,A., Edwards,B.S., Jarvik,J.W., Gresham,H.D., Haynes,M.K., Hjelle,B. *et al.* (2011) Drug repurposing from an academic perspective. *Drug Discov. Today Ther. Strateg.*, **8**, 61–69.
5. Ito,T., Ando,H. and Handa,H. (2011) Teratogenic effects of thalidomide: molecular mechanisms. *Cell Mol. Life Sci.*, **68**, 1569–1579.
6. Attal,M., Harousseau,J.-L., Leyvraz,S., Doyen,C., Hulin,C., Benboubker,L., Yakoub Agha,I., Bourhis,J.-H., Garderet,L., Pegourie,B. *et al.* (2006) Maintenance therapy with thalidomide improves survival in patients with multiple myeloma. *Blood*, **108**, 3289–3294.
7. Teo,S.K., Resztak,K.E., Scheffler,M.A., Kook,K.A., Zeldis,J.B., Stirling,D.I. and Thomas,S.D. (2002) Thalidomide in the treatment of leprosy. *Microbes Infect.*, **4**, 1193–1202.
8. Ghofrani,H.A., Osterloh,I.H. and Grimminger,F. (2006) Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. *Nat. Rev. Drug Discov.*, **5**, 689–702.
9. Pantziarka,P., Sukhatme,V., Crispino,S., Bouche,G., Meheus,L. and Sukhatme,V.P. (2018) Repurposing drugs in oncology (ReDO)-selective PDE5 inhibitors as anti-cancer agents. *Ecancermedicalscience*, **12**, 824.
10. Haupt,V.J., Daminelli,S. and Schroeder,M. (2013) Drug promiscuity in PDB: protein binding site similarity is key. *PLoS One*, **8**, e65894.
11. Burley,S.K., Berman,H.M., Bhikadiya,C., Bi,C., Chen,L., Di Costanzo,L., Christie,C., Dalenberg,K., Duarte,J.M., Dutta,S. *et al.* (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.*, **47**, D464–D474.
12. Nadzirin,N., Gardiner,E.J., Willett,P., Artymiuk,P.J. and Firdaus-Raih,M. (2012) SPRITE and ASSAM: web servers for side chain 3D-motif searching in protein structures. *Nucleic Acids Res.*, **40**, W380–W386.
13. Nadzirin,N., Willett,P., Artymiuk,P.J. and Firdaus-Raih,M. (2013) IMAAAGINE: a webserver for searching hypothetical 3D amino acid side chain arrangements in the Protein Data Bank. *Nucleic Acids Res.*, **41**, 432–440.
14. Wishart,D.S., Feunang,Y.D., Guo,A.C., Lo,E.J., Marcu,A., Grant,J.R., Sajed,T., Johnson,D., Li,C., Sayeeda,Z. *et al.* (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.
15. Li,Y.H., Yu,C.Y., Li,X.X., Zhang,P., Tang,J., Yang,Q., Fu,T., Zhang,X., Cui,X., Tu,G. *et al.* (2018) Therapeutic target database update 2018: enriched resource for facilitating bench-to-clinic research of targeted therapeutics. *Nucleic Acids Res.*, **46**, D1121–D1127.
16. Shameer,K., Glicksberg,B.S., Hodos,R., Johnson,K.W., Badgeley,M.A., Readhead,B., Tomlinson,M.S., O'Connor,T., Miotto,R., Kidd,B.A. *et al.* (2018) Systematic analyses of drugs and disease indications in RepurposeDB reveal pharmacological, biological and epidemiological factors influencing drug repositioning. *Brief. Bioinform.*, **19**, 656–678.
17. Brown,A.S. and Patel,C.J. (2017) A standard database for drug repositioning. *Sci. Data*, **4**, 170029.
18. Kringelum,J., Kjaerulff,S.K., Brunak,S., Lund,O., Oprea,T.I. and Taboureau,O. (2016) ChemProt-3.0: a global chemical biology diseases mapping. *Database*, **2016**, 1–7.
19. Nickel,J., Gohlke,B.-O., Erehman,J., Banerjee,P., Rong,W.W., Goede,A., Dunkel,M. and Preissner,R. (2014) SuperPred: update on drug classification and target prediction. *Nucleic Acids Res.*, **42**, W26–W31.
20. Wang,C., Hu,G., Wang,K., Brylinski,M., Xie,L. and Kurgan,L. (2016) PDID: database of molecular-level putative protein-drug interactions in the structural human proteome. *Bioinformatics*, **32**, 579–586.
21. Wang,X., Shen,Y., Wang,S., Li,S., Zhang,W., Liu,X., Lai,L., Pei,J. and Li,H. (2017) PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.*, **45**, W356–W360.
22. Salentin,S., Schreiber,S., Haupt,V.J., Adasme,M.F. and Schroeder,M. (2015) PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res.*, **43**, W443–W447.
23. Konc,J. and Janežič,D. (2014) ProBiS-ligands: a web server for prediction of ligands by examination of protein binding sites. *Nucleic Acids Res.*, **42**, W215–W220.
24. Moriaud,F., Richard,S.B., Adcock,S.A., Chanas-Martin,L., Surgand,J.-S., Ben Jelloul,M. and Delfaud,F. (2011) Identify drug repurposing candidates by mining the protein data bank. *Brief. Bioinform.*, **12**, 336–340.
25. Rose,A.S., Bradley,A.R., Valasatava,Y., Duarte,J.M., Prlic,A. and Rose,P.W. (2018) NGL viewer: web-based molecular graphics for large complexes. *Bioinformatics*, **34**, 3755–3758.
26. El-Gebali,S., Mistry,J., Bateman,A., Eddy,S.R., Luciani,A., Potter,S.C., Qureshi,M., Richardson,L.J., Salazar,G.A., Smart,A. *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.*, **47**, D427–D432.
27. Holm,L. and Park,J. (2000) DaliLite workbench for protein structure comparison. *Bioinformatics*, **16**, 566–567.
28. Edgar,R.C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
29. Sadowska,A.M., Verbraecken,J., Darquennes,K. and De Backer,W.A. (2006) Role of N-acetylcysteine in the management of COPD. *Int. J. Chron. Obstruct. Pulmon. Dis.*, **1**, 425–434.
30. Heard,K.J. (2008) Acetylcysteine for acetaminophen poisoning. *N. Engl. J. Med.*, **359**, 285–292.

31. Garlatti,V., Belloy,N., Martin,L., Lacroix,M., Matsushita,M., Endo,Y., Fujita,T., Fontecilla-Camps,J.C., Arlaud,G.J., Thielens,N.M. *et al.* (2007) Structural insights into the innate immune recognition specificities of L- and H-ficolins. *EMBO J.*, **26**, 623–633.
32. Krarup,A., Thiel,S., Hansen,A., Fujita,T. and Jensenius,J.C. (2004) L-ficolin is a pattern recognition molecule specific for acetyl groups. *J. Biol. Chem.*, **279**, 47513–47519.
33. Luo,F., Sun,X., Wang,Y., Wang,Q., Wu,Y., Pan,Q., Fang,C. and Zhang,X.-L. (2013) Ficolin-2 defends against virulent Mycobacteria tuberculosis infection in vivo, and its insufficiency is associated with infection in humans. *PLoS One*, **8**, e73859.
34. Neres,J., Hartkoorn,R.C., Chiarelli,L.R., Gadupudi,R., Pasca,M.R., Mori,G., Venturelli,A., Savina,S., Makarov,V., Kolly,G.S. *et al.* (2015) 2-Carboxyquinoxalines kill mycobacterium tuberculosis through noncovalent inhibition of DprE1. *ACS Chem. Biol.*, **10**, 705–714.
35. Grosdidier,A., Zoete,V. and Michielin,O. (2011) SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.*, **39**, W270–W277.
36. Amaral,E.P., Conceição,E.L., Costa,D.L., Rocha,M.S., Marinho,J.M., Cordeiro-Santos,M., D'Império-Lima,M.R., Barbosa,T., Sher,A. and Andrade,B.B. (2016) N-acetyl-cysteine exhibits potent anti-mycobacterial activity in addition to its known anti-oxidative functions. *BMC Microbiol.*, **16**, 251.
37. Zarin,D.A., Tse,T., Williams,R.J., Califf,R.M. and Ide,N.C. (2011) The ClinicalTrials.gov results database–update and key issues. *N. Engl. J. Med.*, **364**, 852–860.