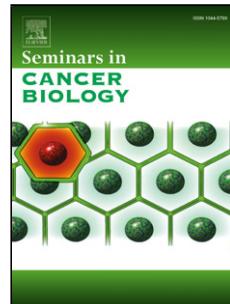


Journal Pre-proof

Structure-based Drug repositioning: Potential and Limits

Melissa F. Adasme, Daniele Parisi, Anastasia Sveshnikova, Michael Schroeder



PII: S1044-579X(20)30010-9

DOI: <https://doi.org/10.1016/j.semancer.2020.01.010>

Reference: YSCBI 1762

To appear in: *Seminars in Cancer Biology*

Received Date: 1 August 2019

Revised Date: 8 January 2020

Accepted Date: 16 January 2020

Please cite this article as: Melissa F. Adasme, Daniele Parisi, Anastasia Sveshnikova, Michael Schroeder, Structure-based Drug repositioning: Potential and Limits, <![CDATA[Seminars in Cancer Biology]]> (2020), doi: <https://doi.org/10.1016/j.semancer.2020.01.010>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.

Structure-based Drug repositioning: Potential and Limits

Melissa F. Adasme^{1,+}, Daniele Parisi^{1,2,+}, Anastasia Sveshnikova¹, and Michael Schroeder^{1,*}

¹Biotechnology Center (BIOTEC), Technische Universität Dresden, 01307 Dresden, Germany

²ESAT-STADIUS, KU Leuven B-3001 Heverlee, Belgium

*to whom correspondence and material requests should be addressed

⁺these authors contributed equally to this work

ABSTRACT

Drug repositioning, the assignment of new therapeutic purposes to known drugs, is an established strategy with many repurposed drugs on the market and many more at experimental stage. We review three use cases, a herpes drug with benefits in cancer, a cancer drug with potential in autoimmune disease, and a selective and an unspecific drug binding the same target (GPCR). We explore these use cases from a structural point of view focusing on a deep understanding of the underlying drug-target interactions. We review tools and data needed for such a drug-centric structural repositioning approach. Finally, we show that the availability of data on targets is an important limiting factor to realise the full potential of structural drug-repositioning.

Drug repositioning

Over the last years the productivity crisis of the pharmaceutical R&D has raised concerns about the future of the field¹. According to Eroom's Law, drug discovery efficiency is declining notwithstanding impressive scientific and technological advances². Taking a new chemical entity to the market has an average estimated cost of US \$2–3 billion³ and a success rate of only 2%⁴. Many hypotheses have been formulated to explain this negative trend, such as the research shift towards more difficult targets¹, the necessity to meet the standard of quality set by previously marketed drugs, and the tendency of overestimating the ability of advances in basic research². One alternative approach to overcome these difficulties is drug repositioning (or drug repurposing), which expands the indications for marketed drugs. This approach offers several advantages, such as a lower risk of failure, a reduced time to market, and less cost (on average only \$300,000⁵). Furthermore, drug repositioning offers opportunities for rare diseases, where lack of pathophysiological characterisations and difficult commercial conditions are the causes for the absence of an FDA approved drug for 95% of conditions^{5,6}, and for personalised medicine where the small size of cohorts makes the discovery process hard and highly expensive⁷. Some successful repositioning cases became very popular stories such as the sildenafil indication change, to erectile dysfunction from hypertension⁸, with worldwide sales of US \$1.88 billion in 2003⁹, and the sadly known thalidomide case, from sedative with severe skeletal birth defects side effects to the erythema nodosum leprosum (ENL, a complication of leprosy)¹⁰ and later to multiple myeloma⁸, of which sales reached US \$224 million in 2003¹¹.

Use cases of structure-based drug repositioning

There are many reasons why a drug could be a cure for two diseases. The diseases might be closely related, the drug's target may play a role in two different diseases, or the drug has two targets each linked to a different disease. Here, we will focus on the latter. Generally, a drug hitting multiple targets is not an exception^{12–14}, as the number of interaction interfaces is considered to be limited^{15,16}. This limited number of interfaces implies re-use of interfaces and the existence of shared binding sites across targets. Haupt et al. showed that the number of targets correlates with the existence of shared binding sites across these targets¹². To this end, structural analysis are crucial to provide an insight to the shared binding sites among targets and to explain the one drug multiple targets concept.

Today, the Protein Data Bank (PDB)¹⁷ comprises structural data for over 1,200 different drug targets¹⁷. Over 60% of these targets are complexed with biologically relevant ligands¹⁸, so that studies of shared binding sites and ultimately structural drug repositioning are possible. Here, we will review three such cases, which differ conceptually. In the first case, the herpes drug BVDU binds a viral thymidine kinase and human heat shock protein implicated in cancer¹⁹. Thus, there is one drug hitting two different targets. In the second case, two chemically dissimilar anti migraine drugs bind the same target 5-HT2C, one specifically, the other one unspecifically²⁰. The third case is the most complex one. The target of the anti-cancer drug ibrutinib shares a binding site with VEGFR2, which is relevant for both cancer and auto-immune diseases²¹. Figure 1 summarises these three cases.

BVDU: From herpes to cancer In the late 1980s, BVDU was first introduced as a treatment against Herpes zoster infection. BVDU is a thymidine analogue and after phosphorylation the virus incorporates BVDU erroneously into its genome, thus bringing viral replication to a halt. The binding mode of BVDU to the viral thymidine kinase is characterised by several specific non-covalent interactions, including its key interaction of a double pi-stacking. Later, based on that key interaction, Heinrich et al.²² showed that BVDU also binds the human heat shock protein Hsp27 (see Figure 1.I), and thereby blocks its anti-apoptotic activity allowing cytotoxic agents to reestablish their efficacy.

The knowledge of the shared binding sites is valuable as it can set the base for improved compounds. In their work, Heinrich et al. went beyond BVDU using in-silico docking on a targeted library. First, they docked other thymidine kinase inhibitors than BVDU to both the viral thymidine kinase and Hsp27. Subsequently, they validated compounds with higher affinity than BVDU leading to a new and better Hsp27 inhibitor with a different scaffold than BVDU. Nonetheless, the inhibitor of Heinrich et al. is a kinase inhibitor. To expand beyond this class of inhibitors, Salentin et al. developed a novel approach which allows to screen the entire PDB for any compound-target pair with interactions similar to BVDU and thymidine kinase. This led to the anti-malarial amodiaquine. The repurposing was validated by showing that amodiaquine suppresses chemo resistance in a multiple myeloma cancer cell line and that it inhibits the chaperone function of the cancer target Hsp27.

Ritanserin and ergotamine binding 5-HT2C Recently, a study on the active/inactive state of the serotonin receptor 5-HT2C in complex with ergotamine (ERG) and ritanserin (RIT), respectively, suggested that ergotamine polypharmacology might be extended also to the delta-opioid receptor²⁰. These differential active/inactive states are mirrored by their distinctive binding modes as observed in the 5-HT2C crystal structures, where the only common interaction between both drugs is the salt bridge between the protonated nitrogen of the ligand and the conserved aspartate 134 (Figure 1.II), a canonical interaction for aminergic and many other GPCRs. Compared to the ERG bound structure, RIT binds approximately one helical turn deeper into the

transmembrane bundle, which is outside of the recognized orthosteric site of other solved aminergic GPCR structures. To uncover the molecular basis for ERG's high affinity and polypharmacological profile at aminergic GPCRs, Peng et al. extended the analysis by checking the binding mode of ERG at the 5-HT1B, 2B and 2C-ERG crystal structure complexes. They found the ergoline core to be recognised by nine key residues, eight of which have specific conserved aminoacid properties to enable binding, which are present in all receptors that demonstrate high ERG affinity. Therefore, the analysis revealed a structural basis for promiscuous ligand ergotamine binding across several receptor sub types (e.g. serotonin, dopamine, adrenergic, histamine, muscarinic, opioid).

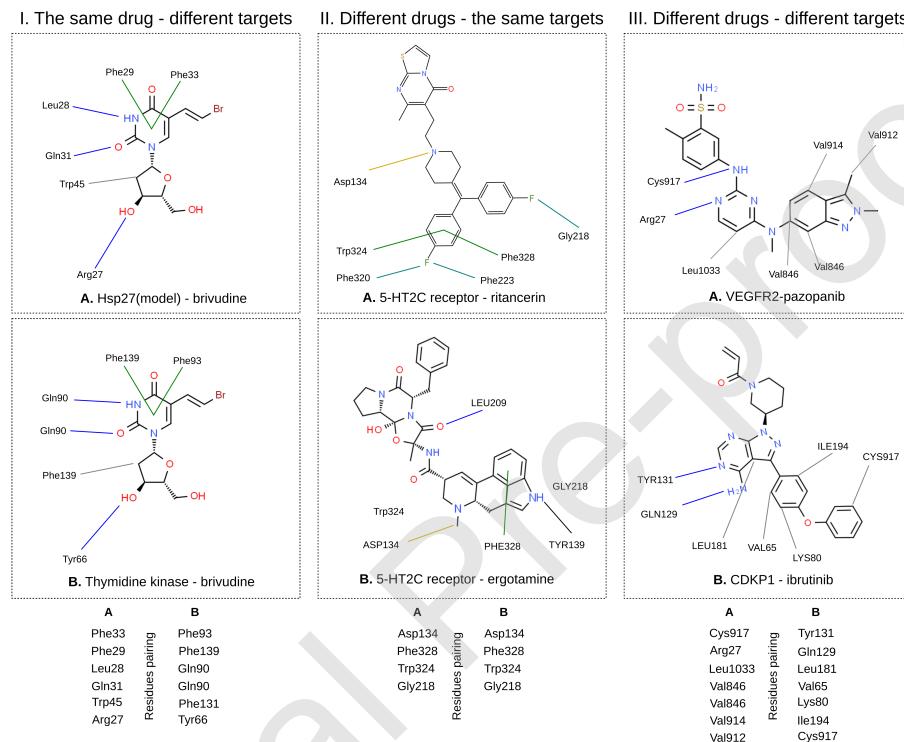


Figure 1. Structural drug-repositioning. I. One drug, two targets. BVDU binds a viral thymidine kinase and the human heat shock protein Hsp27. The two binding sites share at their core two phenylalanine residues. II. Two drugs, one target. Ritancerin, a selective inhibitor, and ergotamine, an unspecific binder, interact with the 5-HT2C receptor through the same four residues. III. One drug, one target. Another drug, another target. The binding sites of Ibrutinib to CDPK1 and of pazopanib and VEGFR2 are similar, so that Ibrutinib also binds VEGFR2. A leucine and a valine play identical roles in both binding sites. Legend to line colors: pi-stacking (green), hydrogen bonds (blue), halogen bond (aquamarine), salt bridges (yellow) and hydrophobic contacts (grey).

Ibrutinib inhibits VEGFR2 Deregulation of B-cells plays an important role in auto-immune disease. Adasme et al. used structural drug repositioning to identify ibrutinib as a VEGFR2 inhibitor and modulator of B-cell activity²¹. Figure 1.III shows two unrelated, but very similar drug-target interactions. Pazopanib is a potent and selective multi-targeted receptor tyrosine kinase inhibitor that blocks tumour growth and inhibits angiogenesis. Figure 1.III.A shows its inhibition of the vascular endothelial growth factor receptor (VEGFR2). Figure 1.III.B depicts the cancer drug ibrutinib binding the calcium dependent protein kinase 1 (CDPK1), an essential enzyme in the opportunistic pathogen *Toxoplasma gondii*. Neither the two drugs,

nor the two targets are structurally similar. Yet, the two binding sites are. Seven individual interactions are shared between the two drug-target pairs, which is highly significant. This finding is remarkable, since ibrutinib was developed as a specific Bruton's tyrosine kinase inhibitor. Its inhibition of VEGFR2 suggests that ibrutinib exerts its anti-cancer effect not only via BTK inhibition, but also via VEGFR2 inhibition.

Other study cases of structural drug-repositioning The above cases relied on drug-target interactions present in the PDB. However, with docking, the scope of the structural drug repositioning can be further expanded. In-silico docking aims to predict affinity and pose of a ligand binding a target. The technique is established, widely used, and can lead to promising results if it is applied to targeted libraries.

Kinnings et al. applied docking to repurpose drugs against multi-drug resistant tuberculosis. Among others, they found the anti-Parkinson drug tolcapone due to a shared binding site between its original target COMT and the new target from docking InhA²³. Similarly, nilotinib was validated as potent inhibitor of MAPK14, adding potential to its role as anti-inflammatory drug²⁴ as well as mebendazole-VEGFR2 and celecoxib-CDH11²⁵.

Lim et al. developed a structural systems pharmacology platform (3D-REMAP) that uses ligand binding site comparison and protein-ligand docking to augment sparse chemical genomic data for the machine learning model of genome-scale chemical-protein interaction prediction. As result they predicted levosimendan, a PDE inhibitor for heart failure, as a new inhibitor of the off targets serine/threonine-protein kinase RIOK1 and other kinases. Subsequent experiments and systems biology analyses confirmed this prediction, and suggest that levosimendan is active against multiple cancers, notably lymphoma, through the direct inhibition of RIOK1 and RNA processing pathway²⁶.

Current tools for structure-based drug repositioning

All of the above structural drug repositioning stories build on two algorithmic approaches: docking and binding site comparisons.

Docking Docking aims to predict the orientation of a ligand into a cavity of a target protein including estimation of the binding affinity²⁷. Docking can be applied in computational drug repositioning pipelines in different ways. For example, by screening a single compounds against a library of protein structures, it is possible to identify new drug-target interactions²⁸. Docking can be also coupled with other techniques and placed either at the end of the computational pipeline to evaluate candidates previously selected with other in silico approaches²³, or at the beginning to generate hypotheses which might be subsequently filtered with other data- and knowledge sources²⁹. Several docking algorithms have been developed, each one with its own strong and weak points, which make each tool diverse and suitable for different cases. Some of those docking tools have been listed in Table 1. Limitations to the use of docking in drug repositioning are mostly related with the availability and quality of structural data. Furthermore, results of molecular docking are prone to a high false positive rate³⁰.

Binding site comparison, pharmacophores, and fingerprints Another popular approach is based on protein binding site similarity. Based on the assumption that similar cavities might present a similar pharmacological profile and hence, accommodate the same ligands, several approaches for binding site prediction and comparison have been reported in the recent years, using geometrical criteria as well as chemical descriptors, as shown in Table 1. Limitations for this approach are related to the possibility of the ligand to change conformation upon binding different targets as well as with the noise produced by the flexible chains present in the protein cavity³¹.

Approach	Tool	Description	Access
Docking	Glide ³²	Search of ligand's conformational space. Allows virtual screening, accurate binding mode prediction and universal applicability.	https://omictools.com/glide-tool
	AutoDock ³³	Search of ligand's conformational space to a set of grids describing the target protein. Does not require choosing atom types, it calculates the grids internally. The atomic affinity grids can be visualised.	http://autodock.scripps.edu/
	Flare ³⁴	Calculation of relative binding affinity within a set of ligands using Free Energy Perturbation. Identification of conformational changes and energetics of ligand binding for lead optimization	https://www.cresset-group.com/software/flare/
	Induced Fit ³⁵	Identification of binding modes and conformational changes within the receptor active sites, with reduced van der Waals radii and increased Coulomb-vdW cutoff. Possible removal of highly flexible side chains	https://www.schrodinger.com/induced-fit
	CovDoc-VS ³⁶	Definition of ligand's conformational space to identify covalent forces between the ligand and the receptor binding pocket. Covalent complexes are minimized to score the top.	https://www.schrodinger.com/covdock
	Lead Finder ³⁷	Algorithm and scoring function for virtual screening, designed to dock covalent and non-covalent ligands. It identifies novel leads processing thousands of molecules per hour.	https://www.cresset-group.com/software/lead-finder/
	MolDock ³⁸	Assignment of charges and protonation states, the prediction of cavities as well as the identification of potential binding sites.	https://omictools.com/mvd-tool
	FlexX ³⁹	Exploration of ligands's conformational space and virtual screening. Available enrichment tool for structure-based drug design. Dock huge libraries by using ultra-high speed docking (1s/ligand).	https://www.biosolveit.de/FlexX/
	GOLD ⁴⁰	Virtual screening, lead optimisation, and binding mode definition of molecules. Enables control over speed versus accuracy settings, with a wide range of scoring functions and customisable docking protocols.	https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/
	Sitemap ⁴¹	Binding site identification and evaluation to predict druggability. It can treat entire proteins to locate binding sites whose size, functionality, and extent of solvent exposure meet user specifications.	https://www.schrodinger.com/sitemap
Binding site prediction	FindSite ⁴²	Ligand binding site prediction, ligand screening and molecular function prediction, based on binding site conservation across evolutionary distant proteins. Applicable to high-resolution experimental structures as well as predicted protein models.	http://pwp.gatech.edu/cssb/findsite-comb/
	LigASite ⁴³	Database of biologically relevant binding sites in protein structures. It consists of proteins with one unbound structure and at least one structure of the protein-ligand complex.	http://ligasite.org/
	PDBeMotif ⁴⁴	Binding site characterization of single proteins or classes of proteins and the conserved structural features of their immediate environments either within the same species or across different species.	https://www.ebi.ac.uk/pdbe-site/pdbemotif/
	LigandScout ⁴⁵	Interpretation of ligand topology and identification of the relevant amino acids as pharmacophore models from known and unknown protein-ligand complexes.	https://omictools.com/ligandscout-tool
Interaction similarity	PLIP ⁴⁶	Analysis and 3D representation of non-covalent interactions in protein-ligand complexes. PLIP provides a report profile with all detected non-covalent interactions defining the ligand's binding mode in atom-level detail.	https://projects.biotech.tu-dresden.de/plip-web/plip/index
	SIFT ⁴⁷	Analysis and representation of 3D protein-ligand binding interactions. SIFT generates an interaction fingerprint that converts 3D structural binding information into a one-dimensional binary string.	https://omictools.com/sift-2-tool
	TIFP ⁴⁸	Encoding of protein-ligand coordinates into a fingerprint of 210 molecular interaction patterns. TIFP enables comparison between interaction pattern similarity and ligand or binding site pairwise similarity.	Available on request from Didier Rognan, CNRS-Université de Strasbourg.

Table 1. Tools for structure-based drug repositioning. List of published commercial and open source tools used to predict new drug-target interactions starting from structural information of protein targets with or without a bound compound. This list is neither complete nor exhaustive but only a mere collection of well-known examples.

A third structure-based approach is the pharmacophore-based screening. To repurpose a ligand to a new target the molecule is screened against a set of binding sites for matching protein–ligand 3D pharmacophoric features, previously obtained with a pharmacophore modelling algorithm⁴⁹.

Protein-ligand interaction profile similarity is another structure-based approach which uses interactions patterns comparison in form of numerical fingerprints to exploit binding mode similarities of drugs and identifies novel repositioning candidates against targets. Binding patterns are functionally characterised and matched based on geometric criteria. Filtering steps are then applied to eliminate redundant or overlapping interactions⁴⁶. Various methods have been developed for the characterisation and comparison of protein-ligand interaction patterns, each one with different fields of application, advantages and special features⁵⁰. Some well known tools for ligand-protein interaction detection and comparison are listed in Table 1.

Available data for structure-based repositioning

The starting point of any structure-based drug repositioning pipeline is the collection of structural data about drugs and targets. Hence, the development of target-based drug discovery and repositioning approaches have been deeply influenced by an almost exponential growth in the number of experimentally determined protein structures⁵¹. The information about 3D protein structures, mostly obtained with x-ray crystallography or NMR, are stored in PDB. By 2018, PDB provided over 135,000 structures (see Figure 2), with 3D structural information for more than 40,000 different protein sequences and 450 complexes between an FDA-approved drug and a relevant target⁵².

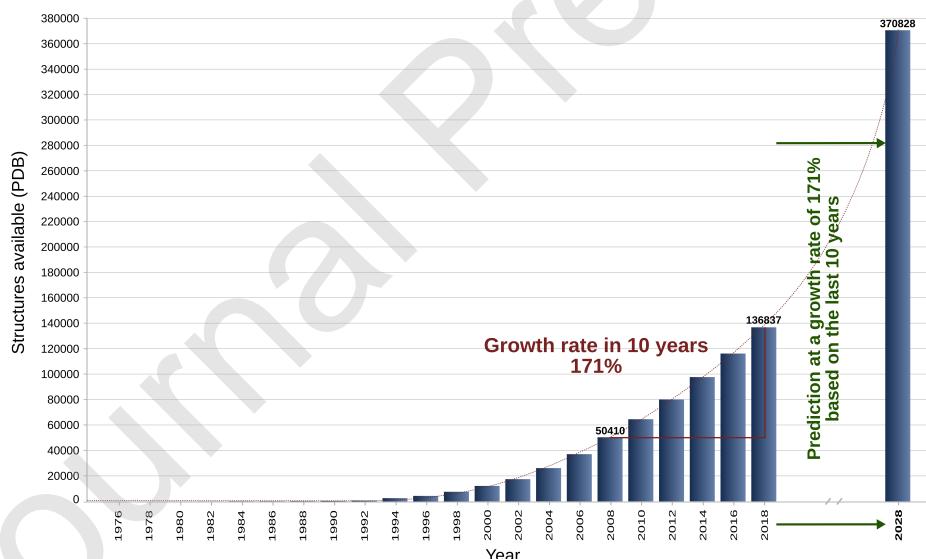


Figure 2. The growing rate of PDB and structures available A PDB growth rate of 171% calculated from the last 10 years (2008-2018) is applied to predict the number of available structures in ten years ahead. By 2028 a total of over 37,000 structures could be available, meaning more than the double of the current amount of structures. This number is only considering the growth rate, but when including the effect of the new upcoming techniques for crystallography, structure modelling and machine learning, among others, the amount of structures might be highly increased.

Over the years, other PDB-related databases have gained importance by focusing on specific structural elements. Het-PDB, for example, collects all the biological relevant small molecules found in a complex with a protein stored in PDB⁵³. On the same

line, in order to facilitate reverse docking and binding site comparison, a collection of pharmacologically relevant protein-ligand complexes has been set up called sc-PDB⁵⁴. Notwithstanding the great amount of structural data available, not all of the therapeutically relevant protein families are equally represented in structural databases. For example, with over 20,000 entries, enzymes are by far the structurally most populated family, while only a handful has been resolved for GPCRs⁵⁵.

Is structure-based drug repositioning ready to contribute to drug discovery?

There are many reasons why a drug could be a cure for two diseases. The diseases might be closely related, the drug's target may play a role in two different diseases, or the drug has two targets each linked to a different disease (see Figure 3). These approaches were introduced by Parisi et al. as disease-, target-, and drug-centric⁵⁶.

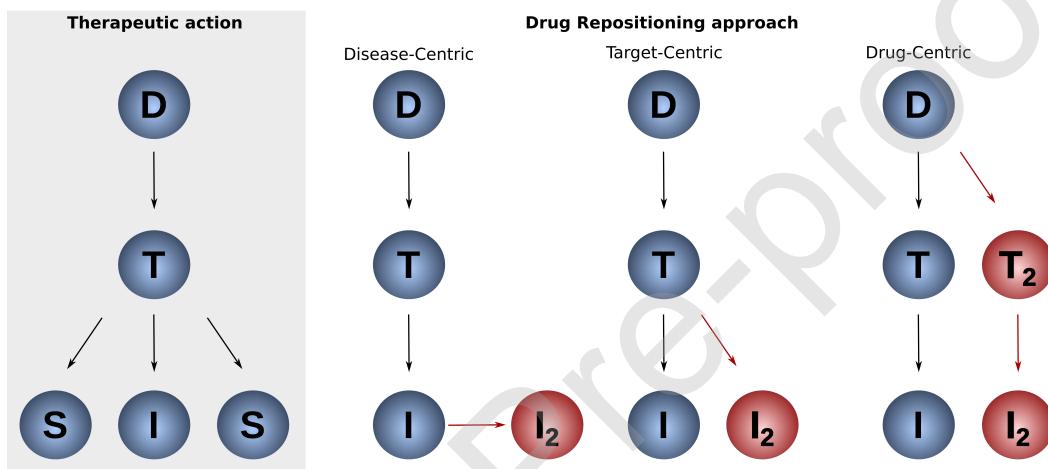


Figure 3. One drug, two indications. A drug (D) has a target (T), which is linked to an indication (I) and possibly side effects (S). There are three ways how the drug can be repurposed. The indications are closely related (disease-centric), the target is linked to two indications (target-centric), and the drug has two targets each linked to a different disease (drug-centric).

Disease-centric repositioning A focus on disease is the most direct approach, since it is driven by the hypothesis that a drug's use can be expanded from the original to a closely related indication. As an example consider nilotinib, a tyrosine kinase inhibitor, approved for the treatment of imatinib-resistant chronic myelogenous leukaemia⁵⁷. A few years later, the Novartis company, proposed the reposition of nilotinib to gastrointestinal stromal tumors. The underlying assumption for disease-centric repositioning is that diseases of the same type have shared guiding principles, e.g. summarised in the hallmarks of cancer⁵⁸. However, despite such commonalities, indications differ and hence repositioning may not succeed. Actually, Novartis' efforts to expand nilotinib to gastrointestinal stromal tumours were abandoned after a phase III trial concluded that it cannot be recommended⁵⁹.

Target-centric repositioning Complementary to a disease-centric approach, target-centric repositioning builds on a novel link between a new indication and an established target. As an example, the protein tyrosine kinase ABL was recently suggested as novel player in Parkinson's disease⁶⁰, and therefore its inhibitors, such as the anti-cancer drug nilotinib, may be effective also against this syndrome⁶¹.

Drug-centric repositioning Besides an established route from a drug via target to an indication, as shown in [Figure 3](#), sometimes a novel target is predicted for the drug which is associated with a new indication. For example, valproic acid is used in bipolar disorder and seizures because its ability to hit the mitochondrial enzymes Succinate-semialdehyde dehydrogenase (ALDH5A1) and 4-aminobutyrate aminotransferase (ABAT). However, due to its off-target interaction with the Histone deacetylase 2 (HDAC2), and the role of this protein in many types of cancers, it has been hypothesised to induce differentiation, growth arrest, and apoptosis in cancer cells. This leads to the repositioning for the treatment of neoplastic conditions such as familial adenomatous polyposis⁶².

In comparison, disease-centric repositioning is the most direct one and may be guided by a good understanding of the molecular mechanisms behind the indications and patents that are broad enough to capture both indications. The target-centric approach is less obvious. The same target can have very different functions, which can be fine-tuned and regulated by the cell. A key is the link of a target to an indication. Often exact causal relationships are not known, but high-throughput techniques such as deep sequencing, micro arrays, and RNAi provide many correlations. The most indirect of the three approaches is the drug-centric one. The use cases discussed previously in this review, are classified as drug-centric approaches. Similar to the target-centric one, this approach requires target-indication links, but it relies additionally on drug-target relations, which have to be derived from the limited structural data available today.

As promising as the use cases described in the previous sections are, we wanted to understand whether they are representative of drug repositioning in general as pursued in commercial and academic settings today.

Can structure-based drug repositioning explain current repositioning cases? To address this question, we considered the extensive analysis described in Parisi et al. for the known repositioning cases to date. They retrieved all repositioning cases, contained in the Repurposed Drug Database (RDD). The database only provides information about the drugs with their original and new indications, therefore, they also integrate the Molecular Drug Targets (MDT) data set⁶³ providing information about the protein targets involved in both indications. The merging of data, led to a compiled report of 196 drug repositioning cases, 263 unique targets and 333 unique indications ([Figure 4.A](#)). In the collected data, more than one third of the cases do not fit a “small molecule drug - protein target” definition. For those cases the drug was usually an antibody and the target usually an RNA or another type of non-protein bio molecule. Therefore, 68 cases were excluded from the data set leaving a total of 128.

The retrospective analysis of the repositioning cases described by Parisi et al. gave an interesting picture of the current state of drug repositioning ([Figure 4.B](#)). 60% of repositioned drugs analysed (76 cases out of 128) have been redirected to the same disease family. This tendency resulted to be extremely frequent within two categories of therapeutic indications: neoplasms and immune system disorders, which show also the highest number of repositioned cases contained in our database. 30% of the analysed drugs (46 out of 128) have been repurposed to a different condition but to the same protein target. Only the 10% (6 cases) has been repositioned to a different disease and a different target. For none of these six cases PDB holds the necessary structural data to compare binding sites.

In summary, disease-centric repositioning is more prominent than target-centric, which is more prominent than drug-centric. This ordering reflects need for external, novel data, and hence risk. Once disease-centric repositioning is saturated, it becomes vital to obtain the high-quality data needed for target-centric and drug-centric repositioning to take off.

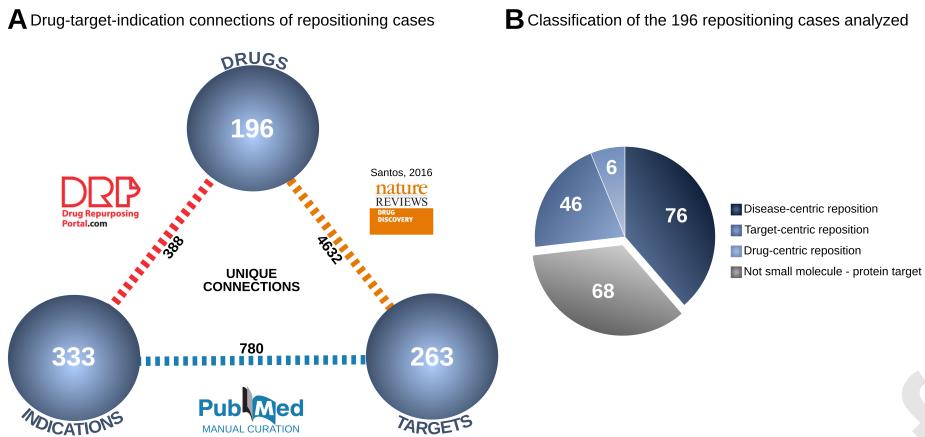


Figure 4. Collection and classification of known repositioning cases⁵⁶. A. Drug, Target, Indication relationships obtained from the Repositioned Drug Database (RDD), the Molecular Target Database (MTD), and PubMed. B. Breakdown of 196 drugs into disease-centric (76), target-centric (46), and drug-centric (6). 68 drugs were discarded (gray) since they were involving non-small molecules or non-protein targets.

Overcoming limiting factors for drug-centric repositioning

In order release the full potential of the drug-centric structure-based drug repositioning, there is an important need for more data along two axes:

Target-indication links A firm link of target to indication is vital, otherwise the repositioning network cannot be constructed. Platforms such as Open Targets⁶⁴ and Beagle⁶⁵ are important to infer new connections and even though they contain multiple links between targets and indications based on literature and experimental evidence, many times those links are only correlations and associations and not causal with well-understood mode of action. Thus, there is a need for data and databases, which curate detailed mode-of-action and causal models disease implication of a target.

Structural drug-target links Only six cases of the RDD database cannot be explained by disease- and target-centric repositioning. And for these six cases, there was no suitable structural data. There are three aspects to overcome these limitations: growth in PDB, modelling techniques and docking. The latter two will benefit from recent breakthroughs in artificial intelligence.

Figure 2 shows the PDB growth rates of the past years. If one extrapolates from the last ten years, then PDB will have doubled in size within the next ten years. This has practical implications, e.g. the drug repositioning campaign of Adasme et al. started from 20 auto-immune targets obtained in an RNAi screen²¹. 15 years, none of these targets did have structure and hence no structural repositioning would have been possible. As of today, the authors used structures for half these 20 targets. With the extrapolation of Figure 2, all of the twenty targets should be structurally available in ten years, a foreseeable time line.

To complement the growth of PDB, homology modelling can generate models of 3D protein structures which are difficult to obtain with crystallographic techniques⁶⁶. Different databases of high-quality 3D protein models such as SWISS-MODEL Repository (SMR)⁶⁷ are already available to support structure-based drug repositioning pipelines. In fact, the repositioning of BVDU from herpes to cancer¹⁹ built on a model of the heat shock protein Hsp27¹⁹. Similarly,²³ used a rat model of the human target COMT²³. Besides, growth of PDB and availability of models, artificial intelligence will improve binding predictions⁶⁸. For example, machine learning was applied to integrate sequence and evolutionary information to predict structures of proteins

uncovered by crystallographic techniques⁶⁹. Artificial intelligence has not only been used to generate new 3D structural data but also to predict the bioactivity of small molecules and new drug-target interactions. In particular AtomNet, a deep convolutional neural network approach, has been developed to model bioactivity and chemical interactions, outperforming previous docking approaches by a large margin⁷⁰.

Overall, these promising developments suggest that drug-centric repositioning will expand over the next decade, while disease-centric, and possibly target-centric repositioning may closer to saturation.

Conclusion

Polypharmacology, a drug binding multiple targets, through remotely related binding sites, is a powerful paradigm for drug repositioning. We have reviewed use cases in cancer, which give an in depth understanding of the drugs mode-of-action and handles to improve drugs after the initial finding. While promising, an in depth analysis of a comprehensive drug repositioning database revealed that only very few cases documented today can build on such a structural, drug-centric repositioning. Instead, the majority of drug repositioning cases today is disease- and target-centric. To realise the full potential of drug-centric repositioning with structural data, there is a need for high-quality, causal links from target to disease and there is a need for more structural data, which may be obtained experimentally or from improved computational *in silico* methods.

Acknowledgements

Thanks to Florian Kaiser and V. Joachim Haupt for valuable feedback.

Author contributions statement

MA, DP conceived the study and wrote the manuscript. MA, AS carried out the structural data collection and figures design. MS wrote the manuscript and supervised the project.

Competing interests

In relation to the work described the authors declare no financial/non-financial competing interests.

References

1. Pammolli Fabio, R. M., Magazzini Laura. The productivity crisis in pharmaceutical R&D. *Nature Reviews Drug Discovery* **10**, 428 (2011).
2. Scannell Jack W., B. H. W. B., Blanckley Alex. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery* **11**, 191 (2012).
3. Nosengo, N. Can you teach old drugs new tricks? *Nature* **534**, 314—316 (2016).
4. Yeu, Y., Yoon, Y. & Park, S. Protein localization vector propagation: a method for improving the accuracy of drug repositioning. *Mol. BioSyst.* **11**, 2096–2102 (2015).

5. Pushpakom, S. *et al.* Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery* **18**, 41 EP – (2018).
6. Jegga, A. G. *et al.* Drug repositioning for orphan diseases. *Briefings in Bioinformatics* **12**, 346–356 (2011).
7. Tiriveedhi, V. Impact of precision medicine on drug repositioning and pricing: A too small to thrive crisis. *Journal of Personalized Medicine* **8** (2018).
8. Ashburn, T. T. & Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. *Nature reviews. Drug discovery* **3**, 673—683 (2004).
9. Pfizer. Pfizer 8-k filing. http://www.pfizer.com/download/investors/financial/8k_0122_04.pdf (2004).
10. Teo, S. K. *et al.* Thalidomide in the treatment of leprosy. *Microbes and Infection* **4**, 1193 – 1202 (2002).
11. Pharmaceuticals, C. Thalomid history. <http://www.celgene.com/thalomid/> (2004).
12. Haupt, V. J., Daminelli, S. & Schroeder, M. Drug promiscuity in pdb: Protein binding site similarity is key. *PLOS ONE* **8**, 1–15 (2013).
13. Mestres, J., Gregori-Puigjané, E., Valverde, S. & Solé, R. V. Data completeness—the achilles heel of drug-target networks. *Nature Biotechnology* **26**, 983 (2008).
14. Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology* **4**, 682 (2008). Review Article.
15. Aloy, P. & Russell, R. B. Ten thousand interactions for the molecular biologist. *Nature Biotechnology* **22**, 1317–1321 (2004).
16. Gao, M. & Skolnick, J. Structural space of protein–protein interfaces is degenerate, close to complete, and highly connected. *Proceedings of the National Academy of Sciences* **107**, 22517–22522 (2010).
17. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242 (2000).
18. Haupt, V. J. *et al.* Computational drug repositioning by target hopping: A use case in chagas disease. *Current Pharmaceutical Design* **22**, 3124–3134 (2016).
19. Heinrich, J.-C. Rp101 (brivudine) binds to heat shock protein hsp27 (hspb1) and enhances survival in animals and pancreatic cancer patients. *Journal of Cancer Research and Clinical Oncology* **137**, 1349 (2011).
20. Peng, Y. *et al.* 5-HT_{2C} receptor structures reveal the structural basis of GPCR polypharmacology. *Cell* **172**, 719 – 730.e14 (2018).
21. Adasme, M. F. *et al.* Structure-based drug repositioning reveals ibrutinib as vegfr2 inhibitor. *bioRxiv* (2019).
22. Heinrich, J. C. *et al.* New HSP27 inhibitors efficiently suppress drug resistance development in cancer cells. *Oncotarget* **7**, 68156–68169 (2016).

23. Kinnings, S. L. *et al.* Drug discovery using chemical systems biology: Repositioning the safe medicine comtan to treat multi-drug and extensively drug resistant tuberculosis. *PLOS Computational Biology* **5**, 1–10 (2009).
24. Li, Y. Y., An, J. & Jones, S. J. M. A computational approach to finding novel targets for existing drugs. *PLOS Computational Biology* **7**, 1–13 (2011).
25. Dakshanamurthy, S. *et al.* Predicting new indications for approved drugs using a proteochemometric method. *Journal of Medicinal Chemistry* **55**, 6832–6848 (2012).
26. Lim, H. *et al.* Rational discovery of dual-indication multi-target pde/kinase inhibitor for precision anti-cancer therapy using structural systems pharmacology. *PLOS Computational Biology* **15**, 1–21 (2019).
27. Ekins, S., Mestres, J. & Testa, B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British Journal of Pharmacology* **152**, 9–20 (2007).
28. Zahler, S. *et al.* Inverse in silico screening for identification of kinase inhibitor targets. *Chemistry Biology* **14**, 1207 – 1214 (2007).
29. Chang, R. L., Xie, L., Xie, L., Bourne, P. E. & Palsson, B. Drug off-target effects predicted using structural analysis in the context of a metabolic network model. *PLOS Computational Biology* **6**, 1–18 (2010).
30. Lorber, D. M. & Shoichet, B. K. Flexible ligand docking using conformational ensembles. *Protein Science* **7**, 938–950 (1998).
31. Milletti, F. & Vulpetti, A. Predicting polypharmacology by binding site similarity: From kinases to the protein universe. *Journal of Chemical Information and Modeling* **50**, 1418–1431 (2010).
32. Friesner, R. A. *et al.* Glide: A new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *Journal of Medicinal Chemistry* **47**, 1739–1749 (2004).
33. Forli, S. *et al.* Computational protein-ligand docking and virtual drug screening with the autodock suite. *Nature Protocols* **11**, 905 (2016).
34. Cheeseright, T., Mackey, M., Rose, S. & Vinter, A. Molecular field extrema as descriptors of biological activity: definition and validation. *Journal of Chemical Information and Modeling* **46**, 665–676 (2006).
35. Sherman, W., Beard, H. S. & Farid, R. Use of an induced fit receptor structure in virtual screening. *Chemical Biology & Drug Design* **67**, 83–84 (2006).
36. Zhu, K. *et al.* Docking covalent inhibitors: A parameter free approach to pose prediction and scoring. *Journal of Chemical Information and Modeling* **54**, 1932–1940 (2014).
37. Stroganov, O. V., Novikov, F. N., Stroylov, V. S., Kulkov, V. & Chilov, G. G. Lead finder: An approach to improve accuracy of proteinligand docking, binding energy estimation, and virtual screening. *Journal of Chemical Information and Modeling* **48**, 2371–2385 (2008).

38. Thomsen, R. & Christensen, M. H. Moldock: a new technique for high-accuracy molecular docking. *Journal of Medicinal Chemistry* **49**, 3315–3321 (2006).
39. A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology* **261**, 470 – 489 (1996).
40. Development and validation of a genetic algorithm for flexible docking11edited by f. e. cohen. *Journal of Molecular Biology* **267**, 727 – 748 (1997).
41. Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chemical Biology & Drug Design* **69**, 146–148 (2007).
42. Skolnick, J. & Brylinski, M. FINDSITE: a combined evolution/structure-based approach to protein function prediction. *Briefings in Bioinformatics* **10**, 378–391 (2009).
43. Dessailly, B. H., Lensink, M. F., Orengo, C. A. & Wodak, S. J. LigASite—a database of biologically relevant binding sites in proteins with known apo -structures . *Nucleic Acids Research* **36**, D667–D673 (2007).
44. Velankar, S. *et al.* PDBe: Protein Data Bank in Europe. *Nucleic Acids Research* **38**, D308–D317 (2009).
45. Wolber, G. & Langer, T. Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling* **45**, 160–169 (2005).
46. Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F. & Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Research* **43**, W443–W447 (2015).
47. Deng, Z., Chuaqui, C. & Singh, J. Structural interaction fingerprint (sift): A novel method for analyzing three-dimensional proteinligand binding interactions. *Journal of Medicinal Chemistry* **47**, 337–344 (2004).
48. Desaphy, J., Raimbaud, E., Ducrot, P. & Rognan, D. Encoding protein–ligand interaction patterns in fingerprints and graphs. *Journal of Chemical Information and Modeling* **53**, 623–637 (2013).
49. Wolber, G. & Langer, T. Ligandscout: 3-d pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling* **45**, 160–169 (2005).
50. Polypharmacology rescored: Protein–ligand interaction profiles for remote binding site similarity assessment. *Progress in Biophysics and Molecular Biology* **116**, 174 – 186 (2014).
51. Representativity of target families in the protein data bank: impact for family-directed structure-based drug discovery. *Drug Discovery Today* **10**, 1629 – 1637 (2005).
52. Knox, C. *et al.* DrugBank 3.0: a comprehensive resource for ‘Omics’ research on drugs. *Nucleic Acids Research* **39**, D1035–D1041 (2010).
53. Yamaguchi, A. *et al.* Het-PDB Navi.: A Database for Protein–Small Molecule Interactions. *The Journal of Biochemistry* **135**, 79–84 (2004).

54. Kellenberger, E. *et al.* sc-pdb: an annotated database of druggable binding sites from the protein data bank. *Journal of Chemical Information and Modeling* **46**, 717–727 (2006).
55. García-Serna, R., Opatowski, L. & Mestres, J. FCP: functional coverage of the proteome by structures. *Bioinformatics* **22**, 1792–1793 (2006).
56. Parisi, D., Adasme, M. F., Sveshnikova, A., Moreau, Y. & Schroeder, M. Drug repositoning or target repositioning: a structural perspective of drug-target-indication relationship for available repurposed drugs. *bioRxiv* (2019).
57. Kantarjian, H. *et al.* Nilotinib in imatinib-resistant cml and philadelphia chromosome–positive all. *New England Journal of Medicine* **354**, 2542–2551 (2006).
58. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
59. Blay, J. Y. *et al.* Nilotinib versus imatinib as first-line therapy for patients with unresectable or metastatic gastrointestinal stromal tumours (ENESTg1): a randomised phase 3 trial. *Lancet Oncol.* **16**, 550–560 (2015).
60. Karuppagounder, S. S. *et al.* The c-Abl inhibitor, nilotinib, protects dopaminergic neurons in a preclinical animal model of Parkinson’s disease. *Sci Rep* **4**, 4874 (2014).
61. Lonskaya, I., Hebron, M. L. & Moussa, C. E.-H. Nilotinib reverses loss of dopamine neurons and improves motor behavior via autophagic degradation of -synuclein in Parkinson’s disease models. *Human Molecular Genetics* **22**, 3315–3328 (2013).
62. Huang, X. & Guo, B. Adenomatous polyposis coli determines sensitivity to histone deacetylase inhibitor–induced apoptosis in colon cancer cells. *Cancer Research* **66**, 9245–9251 (2006).
63. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery* **16**, 19 (2016).
64. Sanseau, P. *et al.* Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Research* **45**, D985–D994 (2016).
65. ElShal, S. *et al.* Beagle: from literature mining to disease-gene discovery. *Nucleic acids research* **44**, e18–e18 (2016).
66. Pieper, U. *et al.* MODBASE: a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research* **34**, D291–D295 (2006).
67. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research* **37**, D387–D392 (2008).
68. Mak, K.-K. & Pichika, M. R. Artificial intelligence in drug development: present status and future prospects. *Drug Discovery Today* **24**, 773 – 780 (2019).
69. AlQuraishi, M. Proteinnet: a standardized data set for machine learning of protein structure. *BMC Bioinformatics* **20**, 311 (2019).
70. Wallach, I., Dzamba, M. & Heifets, A. Atomnet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. *CoRR* **abs/1510.02855** (2015).