Available online at www.sciencedirect.com

ScienceDirect

Mendeleev Commun., 2021, **31**, 291–293

Mendeleev Communications

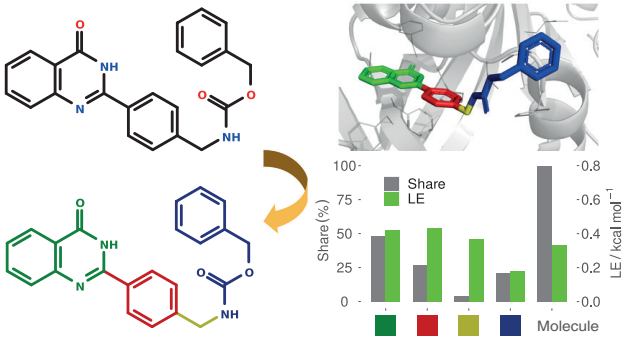
Reverse fragment based drug discovery approach *via* simple estimation of fragment contributions

Dmitry A. Shulga, Nikita N. Ivanov and Vladimir A. Palyulin*

Department of Chemistry, M. V. Lomonosov Moscow State University, 119991 Moscow, Russian Federation. E-mail: vap@qsar.chem.msu.ru

DOI: 10.1016/j.mencom.2021.05.004

Contributions of different fragments of a ligand into the binding/activity to a specified target are of importance to guide hit-to-lead drug discovery, and fragment based drug discovery (FBDD) approach has proven to be quite fruitful. However, the experimental means of FBDD are generally not affordable to many researchers working in the drug discovery field, especially to small medicinal chemistry groups at universities. To partially solve this problem, we propose a Reversed Fragment Based Drug Discovery (R-FBDD) approach in which the contributions of fragments of a molecule are estimated using scoring functions in order to detect whether a fragment is a 'binding anchor' or a ballast, thus guiding further development.



Keywords: fragment based drug discovery, ligand efficiency, molecular modeling, scoring function, drug discovery, hit optimization.

Dedicated to the memory of our teacher Academician Nikolay S. Zefirov

Fragment based drug discovery (FBDD) approach has proven to be useful and intuitively appealing to medicinal chemists.[1,2] The approach relies on the identification of the geometry and energy of binding of small molecules, called fragments, to a target. The experimental study of the fragment binding still remains expensive for widespread applications. On the other hand, a significant part of drug discovery is done outside the Big Pharma: by contract research organisations (CRO), biotech start-ups and university labs.[3] Many university medicinal chemistry labs are lacking the experimental capabilities of employing FBDD despite they are focused on the synthesis of certain organic molecules, containing core fragments of their most expertise, which could serve as scaffolds for potentially active ligands.[4] Thus, a small group tends to be stuck in designing a series of ligands with affordable substituents (from the shelf), followed by testing against available targets, generally leading to low activity and obscure prospects to optimize those hits into leads.

Our idea is to use the ranking capability of the scoring functions (to select the most promising molecules in a set), but to direct it inward to the fragments constituting the whole molecule. Once the value of the scoring function is estimated for each fragment, the usefulness of each fragment in a certain binding mode could be assessed. A special attention is paid to the dedicated fragment a medicinal chemistry group is working with. In the case when the ranking of this scaffold is high, it is likely that further development would bring an active ligand in which the dedicated fragment is crucial. Otherwise, if the ranking of the dedicated scaffold among the other fragments is low, its importance is questionable. The systematic application of fragment contribution analysis constitutes the Reversed Fragment Based Drug Discovery (R-FBDD) approach, since its core ideas are the same as in FBDD, namely, to build molecules from useful fragments. However, the focus stands in the reverse direction, namely, to find useful fragments within a molecule.

Estimation of the fragment contributions is done in four steps. First, a model of ligand–receptor complex is chosen, either experimentally or obtained by docking. Second, the ligand is subdivided by breaking certain single bonds into fragments. In this work, a manual ligand splitting is used, however automatic splitting is also possible.[5] Each fragment is capped with hydrogens to form a correct ground state molecule. Third, the binding energy of each fragment (and that of the whole ligand) to the target is evaluated separately in the positions of fragments taken from the ligand–receptor geometry. Fourth, the unitless fragment energy contributions, ω_j, are estimated according to a simple stakeholder scheme (1). Once ω_j's are defined, an estimation of the whole molecule interaction, ΔE_{mol}, can be factored by equations (2) and (3) into the scaled interaction energy of each fragment, E_j^{Scaled}, in the units of ΔE_{mol}. The value of ΔE_{mol} can be taken from the experiment, if available, or from the modelling data (*e.g.* docking scoring functions).

$$\omega_j = \frac{\text{Score (Fragment}_j\text{)}}{\sum_{i=1}^{N} \text{Score (Fragment}_i\text{)}} \quad (1)$$

$$E_j^{Scaled} = \omega_j \, \Delta E_{mol} \quad (2)$$

$$\sum_{i=1}^{N} E_i^{Scaled} = \Delta E_{mol} \sum_{i=1}^{N} \omega_i = \Delta E_{mol} \quad (3)$$

The E_j^{Scaled} can be used also to estimate the effectiveness of each fragment in its position within the whole ligand by means of ligand efficiency (LE) metric (4), often used in FBDD.[6]

$$LE = -\Delta E/NH, \quad (4)$$

where ΔE is ΔE_{mol} (E_j^{Scaled}) for the whole ligand (fragment), in kcal mol^{-1}; NH is the number of heavy (non-hydrogen) atoms in the ligand (fragment).

AutoDock Vina v.1.1.2 was used for scoring ligand and fragment positions as well as to find the plausible ligand pose by molecular docking (example 2, ligand **4**, Figure 1).[7] AutoDockTools v.1.5.6 was used to prepare all ligands and receptors.[8] All processing necessary for the estimation of fragment contributions was made in a specially designed program written in Python (version 2.7) using openbabel (v.2.4.1) library.[9]

An ideal case for the early stage drug discovery is a small ligand which tightly binds to a certain part of the target binding site whereas leaving room for further development by substituting the scaffold to fill the unoccupied pockets. It is better if this ligand contains the dedicated scaffold. Thus, the aim for a small medicinal chemistry group is to early identify the 'binding anchor', a part of a molecule, which specifically and tightly binds to the defined pockets of the binding site and, ideally, contains the dedicated scaffold. The binding anchor provides the surplus of ligand efficiency (LE), which can be spent to expand the ligand to the sizes and activity values pertinent to drugs with simultaneous optimization of selectivity and ADMET properties, without the loss of the binding mode.

We exemplify the use of the technique by a retrospective project, where Tankyrase 2 inhibitors are sought starting from an initial hit **1** (see Figure 1) containing the dedicated scaffold 3,4-dihydroquinazolin-4-one.[10] Three compounds for which both experimental activity and geometry are known form a hypothetical hit-to-lead optimization series of **1** → **2** → **3**. The scaffold contribution was estimated to check the hypothesis whether the scaffold forms a binding anchor. The experimental positions of each ligand–receptor complex (**1**–**3**), showing almost perfect scaffold superposition (see Online Supplementary Materials, Figure S1), were used. According to the obtained results (Table 1), the scaffold contribution is the largest and preserved in the series. It can be assumed that the scaffold is the anchor fragment which keeps the binding mode. Thus, the method is useful for performing a hit-to-lead process with the control of the scaffold position and promptly checking hypotheses *in silico*.

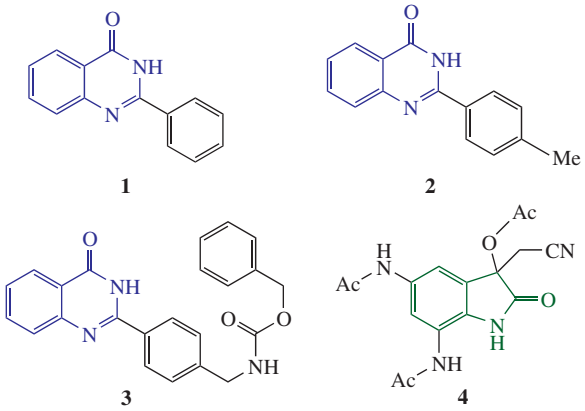


Figure 1 Structures of compounds **1**–**3** with the dedicated scaffold in blue from Tankyrase 2 case, and structure of **4** with the dedicated scaffold in green from MT3/QR2 case.

Another application of R-FBDD relevant to small groups is a simplification of initial hit compounds which tend to be densely substituted with chemical groups arising from the readily available reagents, just in order to get molecule sizes that could reveal significant activity under experimental conditions, where experimental FBDD methods to determine binding mode and energy of small fragments are not affordable. If further development of such hits is hampered by the lack of clear directions, the R-FDBB allows one to take a 'step back' by detecting and eliminating the least contributing groups. LE metric provides insights about such groups. At early stages of drug discovery, LE values should be as high as possible, with a practical threshold LE level being *ca.* 0.3 kcal mol^{-1} atom^{-1}.

As an example, ligand **4** is used, which possesses *in vivo* activity in lowering intraocular pressure as a therapy against glaucoma, presumably acting on MT3/QR2 receptor.[11] The dedicated scaffold in structure **4** is extensively substituted. Further development is hampered by the lack of clear directions.

The most plausible binding mode found by docking to the MT3/QR2 receptor model from PDB:2QWX was analyzed in terms of fragment contributions. First, the dedicated scaffold contributes significantly, which means the chances to develop the ligand while retaining the dedicated scaffold are good. Second, two substituents along with the scaffold effectively form a binding anchor, although not excellent in terms of energy and LE (Figure 2, green and yellow). And third, two other substituents contribute to the overall binding with the particularly low LE at least below the LE value for the whole ligand (see Figure 2, in red). Visual analysis confirms that those substituents do not explicitly participate in any directed interaction. An indirect role that they could still play is in placing the otherwise flat ligand relative to the gorge of the binding site by not allowing a tetrahedrally substituted carbon atom at position 3 to penetrate into the binding site (see Online Supplementary Materials, Figure S2). A simple elimination of the two low contributing substituents immediately results in LE increase of the ligand from 0.37 to 0.44 kcal mol^{-1} atom^{-1}.

We believe that R-FBDD approach could be generally used *in silico* to make a 'step back' and to consecutively remove the substituents for increasing ligand efficiency of the whole molecule and utilise different substitution patterns to increase the efficacy of new ligands. This approach definitely does not aim at replacing the experimental testing, instead, it provides

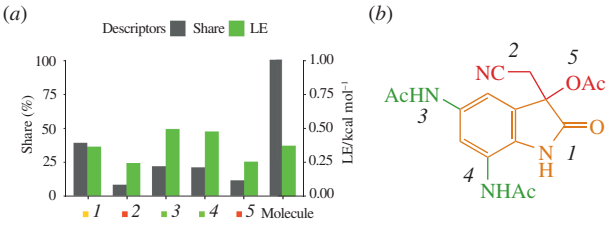


Figure 2 (*a*) Fragment shares (energy) and ligand efficiency (LE) of **4** for docking found position and (*b*) fragments *1*–*5* of structure **4** colored in LE palette: red – low LE, yellow – mean LE and green – high LE values.

Table 1 Estimations of scaffold contribution in the binding energy of compounds **1**–**3** with Tankyrase 2.

Compound	PDB	IC_{50}/nM	ΔE exp. bind./ kcal mol^{-1}	LE (exp.)	ΔE pred. (Vina)	Scaffold energy/ kcal mol^{-1}	Scaffold LE	Scaffold share, % (ΔE)
1	4BU3	2200	–7.8	0.46	–9.5	–5.9	0.54	62
2	4UHG	540	–8.7	0.48	–10.2	–6.1	0.55	60
3	4UI4	45	–10.1	0.35	–9.6	–4.9	0.45	51

lean (*in silico*) tools which could prioritize the experiments and guide further steps of development by providing valuable structure based insights.

The above two cases of using R-FBDD approach naturally form a lean strategy of hit and lead optimization: we either use the most beneficial fragments (presumably with the dedicated chemistry) or get rid of the useless fragments. Both steps potentially result in better ligands.

Online Supplementary Materials

Supplementary data associated with this article can be found in the online version at doi: 10.1016/j.mencom.2021.05.004.

References

1 M. Congreve, C. W. Murray, R. Carr and D. C. Rees, *Annu. Rep. Med. Chem.*, 2007, **42**, 431.
2 D. Joseph-McCarthy, A. J. Campbell, G. Kern and D. Moustakas, *J. Chem. Inf. Model.*, 2014, **54**, 693.
3 R. M. Rydzewski, *Real World Drug Discovery: A Chemist's Guide to Biotech and Pharmaceutical Research*, Elsevier, Oxford, 2010.
4 A. Whitty, *Future Med. Chem.*, 2011, **3**, 797.
5 T. Liu, M. Naderi, C. Alvin, S. Mukhopadhyay and M. Brylinski, *J. Chem. Inf. Model.*, 2017, **57**, 627.
6 A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees and C. H. Reynolds, *Nat. Rev. Drug Discovery*, 2014, **13**, 105.
7 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455.
8 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785.
9 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminf.*, 2011, **3**, 33.
10 A. Nathubhai, T. Haikarainen, P. C. Hayward, S. Muñoz-Descalzo, A. S. Thompson, M. D. Lloyd, L. Lehtiö and M. D. Threadgill, *Eur. J. Med. Chem*., 2016, **118**, 316.
11 E. V. Zaryanova, N. A. Lozinskaya, O. V. Beznos, M. S. Volkova, N. B. Chesnokova and N. S. Zefirov, *Bioorg. Med. Chem. Lett*., 2017, **27**, 3787.

Received: 5th March 2021; Com. 21/6482