Advances in Multi-parameter Optimisation Methods for de Novo Drug Design

Abstract

Introduction

A high quality drug must achieve a balance of physicochemical and ADME properties, safety and potency against its therapeutic target(s). Multi-parameter optimisation (MPO) methods guide the simultaneous optimisation of multiple factors to quickly target compounds with the highest chance of downstream success. MPO can be combined with 'de novo design' methods to automatically generate and assess a large number of diverse structures and identify strategies to optimise a compound's overall balance of properties.

Areas Covered

We will review MPO methods and recent developments in the methods and opinions in the field. We will describe advances in de novo design that improve the relevance of automatically generated compound structures and integrate MPO. Finally we will discuss a recent case study of the automatic design of ligands to polypharmacological profiles.

Expert Opinion

Recent developments have reduced the generation of chemically infeasible structures and improved the quality of compounds generated by de novo design methods.

There are concerns about the ability of simple drug-like properties and ligand efficiency indices effectively to guide the detailed optimisation of compounds.

De novo design methods cannot identify a perfect compound for synthesis, but can identify high quality ideas for detailed consideration by an expert scientist.

1. Introduction

The process of drug discovery involves the optimisation of many compound properties in the search for a successful drug. It is now widely recognised that potency against a valid therapeutic target is not sufficient for a compound to progress in drug discovery. Even from an early stage, a compound must also exhibit acceptable physicochemical, absorption, distribution, metabolism and elimination (ADME) properties and as low a risk as possible of causing toxic effects. Ultimately, a development candidate with a high chance of reaching the market must satisfy multiple, often conflicting, criteria relating to efficacy, pharmacokinetics (PK) and safety.

The evidence to date suggests that the industry has not been very successful in achieving this challenging goal. In 2000, the success rate of candidate drugs entering development in the years 1991 to 2000 was found to be 11% [1] and in an analysis published in 2010 this was essentially unchanged at 12% for the period 2000 to 2007 [2]. Recent data from CMR International suggests this may have fallen even further, with a success rate of only 4% between 2006 and 2010 [3]. While many of the clinical failures can be attributed to lack of efficacy, in many cases due to the exploration of previously unvalidated therapeutic targets, a significant proportion was due to poor PK or safety issues. Interestingly, the causes for failure have changed over time; in 1991, 39% of failures in development were attributed to poor PK [4], but by 2000 this had fallen to an estimated 10% [1]. Conversely, in the same period, failures due to toxicity and safely issues increased from approximately 14% to 30%.

The reduction in failure rate in development due to poor PK was driven by the introduction of early screening for related ADME properties, using both high-throughput experimental assays [5] and *in silico* predictive models [6] [7]. However, the reduction in failure rate due to PK issues in development was accompanied by an increase in attrition of compounds and projects in the earlier discovery phases and in the length and number of iterations taken in lead optimisation. It has also been argued that these negative trends have been exacerbated because the 'low hanging fruit' of easily 'druggable' targets have been exhausted and more recent drug discovery targets present a greater challenge for drug discovery. These effects have contributed to the spiralling cost of pharmaceutical R&D, now estimated to be over $1.8 B per new chemical entity reaching the market [2] and, as we have seen, has not resulted in an improvement in the overall success rate.

7221 Cambridge Research Park
Beach Drive, Cambridge
CB25 9TL, UK

Tel: +44 1223 815900
Fax: +44 1223 815907

Email: info@optibrium.com
Website: www.optibrium.com

Many recent developments have focused on methods to aid the simultaneous optimisation of multiple factors required in a successful drug, targeting compounds with the highest chance of downstream success early in the discovery process [8]. These approaches are described by various terms, such as multi-parameter optimisation (MPO), multi-dimensional optimisation (MDO), multi-objective optimisation (MOOP) or multi-criteria decision-making (MCDM). For convenience, herein we will refer to all such methods as MPO. By quickly focusing efforts on high quality compounds and helping to simultaneously consider multiple optimisation criteria, MPO can reduce the number of design-make-test iterations, and hence the time and cost of drug discovery, when compared with sequential optimisation of individual properties. MPO methods can also help to quickly identify when the chemistry being explored is unlikely to yield a compound with the balance of properties required in a successful candidate, helping to "fail fast, fail cheap" and reduce the number of expensive late stage failures.

A further challenge for MPO is to identify strategies for modifying a compound to improve on its overall balance of properties. Unfortunately, modifications that result in the improvement of one property, e.g. potency, often have a negative impact on other important properties, e.g. lipophilicity. Here, computational methods can also help to address this challenge by increasing the diversity of optimisation strategies that can be explored. Techniques described as 'de novo design' can propose new compound structures, based on one or more starting points, for detailed consideration by a expert chemist. By drawing on a wealth of analogue design techniques, computers can generate and assess a far greater number of compounds than could be considered unaided by a single chemist or even a project team.

In this review, we will consider the state of the art of both MPO and de novo design methods and describe their combination in platforms that help to guide the design and selection of high quality compounds. The next section describes methods for MPO and recent developments in the field. This is followed by a discussion of recent advances in de novo design techniques and a case study that illustrates their joint application. Finally the Expert Opinion section draws some conclusions and assesses the future directions for these technologies.

2. Multi-Parameter Optimisation: Progress and Practice

MPO challenges are not unique to drug discovery; other fields, including engineering, quality control and economics, have provided methods that have been adapted to the requirements of drug discovery. However, one notable difference between drug discovery and other disciplines, such as engineering, is the confidence in the available data; while simulations of engineering problems can yield predictions with uncertainties of fractions of a percent, the complexity of biological systems means that predicted values often have uncertainties of an order of magnitude and even experimental values may vary by factors of between 2 and 5.

A previous review [9] provided a detailed description of MPO methods that have been applied in drug discovery. In this section, we will give an overview of the main methods and discuss more recent developments. Also, while MPO methods may be applied equally to experimental and calculated data, in this review we will focus on applications to calculated parameters because this is relevant to de novo design which, by definition, explores virtual compounds.

2.1. Filtering

The most common method for MPO is to apply multiple property filters to reject compounds that don't meet all of the property criteria. While the simplicity of filtering is very attractive, this approach has a number of notable drawbacks. 'Hard' filters draw artificially harsh distinctions between compounds with similar properties; if we choose to accept compounds with a molecular weight (MW) less than 500 Daltons, does this mean that a compound with MW of 501 has a significantly lower chance of success than one with MW of 499?

This harsh distinction is further compounded by the uncertainty in the data to which a filter may be applied. For example, calculated octanol:water partition coefficient (clogP) values typically have a standard error of prediction of 0.5 log units. This means that the probability of a compound with a clogP of 5.0 actually being more lipophilic than one with clogP of 4.5 is only 76% (assuming a normally distributed error). Therefore, does it make sense to apply a hard cut-off of clogP less than 5?

These uncertainties accumulate when we apply multiple filters in sequence. For example, if we have a series of 10 filters that are each 90% accurate, the probability of a perfect compound passing all of the filters, even if one were present in the first place, is only 35%. In other words, the filters would be more likely to reject an ideal compound than to pass it.

For this reason, filters should be treated with caution. In situations where good possibilities are abundant, it may be appropriate to apply multiple filters. However, where the cost of a missed opportunity is high, as is frequently the case in drug discovery, the risk of incorrectly rejecting good compounds may be too great.

2.2. Desirability Indices

One approach to avoid the artificial harshness of simple filters is a method that relates the value of a property to the 'desirability' of that outcome, using a 'desirability function' [10]. This is a mathematical function that translates the value of a property into a number between 0 and 1, representing how desirable that outcome would be; a desirability of 1 indicates that the property value is ideal, while 0 corresponds to a completely unacceptable outcome. Desirability functions can take many forms (Figure 1 shows some illustrative examples of simple, linear desirability functions) and map the property values onto a continuous scale of desirability, in contrast with the binary pass/fail outcome of a filter. This allows more subtle distinctions to be made between compounds.

The individual desirabilities of multiple properties can be combined to calculate an overall 'desirability index' representing the quality of the compound against a profile of multiple properties. The most common approaches for combining the individual property desirabilities use additive or multiplicative approaches:

Additive: $D(x_1, x_2, \dots, x_N) = \sum_{i=1}^{N} c_i d_i(x_i)$

Multiplicative: $D(x_1, x_2, \dots, x_M) = \prod_{i=1}^{N} d_i(x_i)^{c_i}$

where x_i are the values of N compound properties, d_i are the desirability functions for the properties and c_i are optional coefficients that can be used to define the importance of each individual property. These are sometimes normalised by the (weighted) number of properties by dividing by $\sum_{i=1}^{N} c_i$, in the case of an additive desirability index or by taking the $(\sum_{i=1}^{N} c_i)$'th root of the product in the multiplicative approach.

One disadvantage of an additive approach is that when large numbers of properties are being combined in an assessment of the overall desirability, a very low desirability for a single property will only have a small impact on the desirability. However, if a compound has an unacceptable value of a critical property the compound should be rejected, e.g. a compound with 100 μM IC_{50} is not of interest, even if it has ideal ADME properties. This limitation is overcome by the use of a multiplicative approach, in which it is possible to define one or more properties for which a very poor outcome is sufficient to reject a compound outright. Conversely, unless the desirability functions for the individual properties are carefully considered, this behaviour of a multiplicative scheme can lead to an overly harsh penalty being applied to a compound where a poor outcome for one property can be mitigated by a good outcome in another.

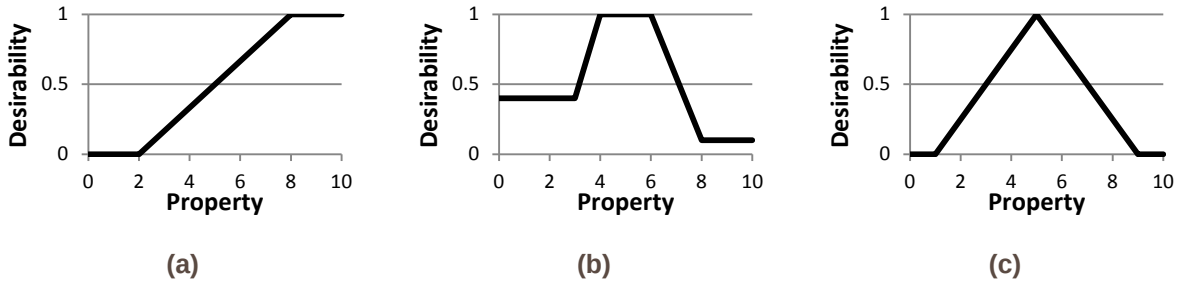


Figure 1. Three examples of linear desirability functions. (a) represents an ideal property criterion of >8, with linearly increasing desirability between 2 and 8. (b) represents an ideal property range of 4 to 6, with lower desirability above this range than below and linearly decreasing desirability above and below the ideal range. (c) represents an ideal property value of 5 with linearly decreasing desirability above and below this value and a desirability of 0 for property values below 1 or above 9.

A further limitation of a simple additive scheme is that it can lead to inappropriate biases where data are missing for some properties of one or more of the compounds being compared. The maximum possible score increases with the number of properties for which data are available, which can lead to a bias towards compounds that have been progressed further and hence have been studied to a greater extent. This drawback is particularly relevant when desirability functions are based on experimentally measured properties – but is less of a concern in the context of *de novo* design, which relies on calculated compound properties – and can be partially mitigated by normalising the desirability index by the number of properties for which data are available.

A recent paper by Nissink and Degorce [11] proposed an alternative approach to address some of these issues, by treating the desirabilites of the individual properties as coordinates in an N-dimensional space *(d_1, d_2, ..., d_N)* and using the distance to the point representing the perfect compound, *(1, 1, ..., 1)*, as a measure of the quality, to give:

$$D = 1 - \sqrt[2]{\frac{1}{N}\sum_{i=1}^{N}(1-d_i)^2}\,.$$

If the value of *N* used to calculate the score for each compound is the number of properties for which data are available for that specific compound, the overall score is scaled by the maximum possible distance from the ideal compound and therefore this metric is more robust to missing data than a simple additive scheme. The authors also note that this can easily be adapted to weight the individual properties in a similar manner to the multiplicative and additive schemes.

Desirability functions provide much greater flexibility than filters in defining the property requirements for a successful compound, the importance of each individual property to the overall objective of a project and the acceptable trade-offs if an ideal compound cannot be identified. However, to define the desirability functions and their weights for a specific project objective, requires an *a priori* knowledge of the ideal compound property values and acceptable compromises. The complexity of the data now generated in drug discovery means that this may not always be clear, even to an experienced scientist.

2.3. Ligand Efficiency Indices

Ligand Efficiency (LE) was first proposed by Kuntz *et al.* [12] and further popularised by Hopkins *et al.* [13], who suggested the binding energy (ΔG) per heavy atom as a metric for selection of leads:

$$LE = \frac{\Delta G}{HAC} = \frac{RT\ln K}{HAC} \approx \frac{1.4\times pIC_{50}}{HAC},$$

where HAC is the count of heavy (i.e. non hydrogen) atoms in the compound, pIC_{50} = -log(IC_{50}) and the IC_{50} is expressed in molar concentration. This is based on the observed correlations of increasing compound size with poor physicochemical and ADME properties. Therefore, of two equipotent compounds, the smaller will typically have a lower risk.

The LE concept has led to the definition of an increasing array of ligand efficiency indices (LEIs), including the Ligand Lipophilicity Efficiency (LLE),

$$LLE = pIC_{50} - logP,$$

reflecting the increased risk of high lipophilicity, and others combining potency with measures of polarity, MW and other simple compound characteristics [14].

These LEIs may provide useful metrics to track the optimisation of compounds, to ensure that increased potency is not being achieved simply by increasing compound size or lipophilicity. Increasing compound bulk and/or lipophilicity can increase potency by increasing the entropy of binding through displacement of coordinated water molecules from the protein binding pocket, instead of forming specific interactions with the protein, which decreases the enthalpy of binding [15]. However, such non-specific interactions increase the risk of off-target binding and non-specific toxicity. A recent review by Hopkins *et al.* [16] provides an excellent overview of the application of ligand efficiency metrics in drug discovery.

However, the simple properties on which these LEIs are based represent only a subset of the factors that influence the success of a compound and have only limited correlation with the biological properties of a compound. Furthermore, the data from which the LEIs are calculated often have significant uncertainty, which means that the values of LEIs will, themselves, be uncertain. Therefore, the values of LEIs should not be over-interpreted or used as hard filters in the selection of compounds.

Concerns have been raised regarding the correlation of composite parameters, such as LEIs, with the quality or success rate of compounds. An analysis by Shultz, based on experimental data and theoretical 'thought experiments', suggests that empirically derived composite parameters based on normalisation of potency by HAC or MW, such as LE, may be misleading. This is partly due to the asymptotic behaviour of the quotient, such that the LE becomes essentially independent of potency for low HAC [17] [18]. Shultz's analysis suggests that LLE may provide a more reliable metric due to its correlation with enthalpy-driven binding [19]. Size corrected ligand efficiency indices have also been proposed to mitigate this effect, such as the Fit Quality [20] and Size Independent Ligand Efficiency [21].

Despite these concerns, LEIs may provide a qualitative approach to understanding the relationship between compound potency and other relevant characteristics. For example Abad-Zapatero has proposed that LEIs representing polarity and size can be used to map 'Chemico-Biological Space' to chart trends across different target classes or the progress of optimisation in a drug discovery project [22] [23]. LEIs may also be useful when combined with other parameters in the context of another, more general MPO method, such as desirability functions or probabilistic scoring (see below), where the LEI can be given appropriate weight in the selection of compounds and the uncertainties in the values can be explicitly taken into account [24]

2.4. Pareto Optimisation

Pareto Optimisation is based on the principle that there may not be a single, optimal solution to an optimisation problem, but a family of possible outcomes that represent different, 'optimal' balances of properties [25]. A Pareto optimal solution (a compound in the context of drug discovery) is one for which there is not another solution that is better in all other properties. These solutions are often described as 'non-dominated' and form a 'Pareto front' in the property space. An illustration of this concept is shown in Figure 2.

Pareto optimisation is best applied in situations where an ideal compound cannot be found and the acceptable trade-offs between properties is not known *a priori*. The Pareto-optimal compounds sample different property combinations, which can be studied further to determine the best compromise.

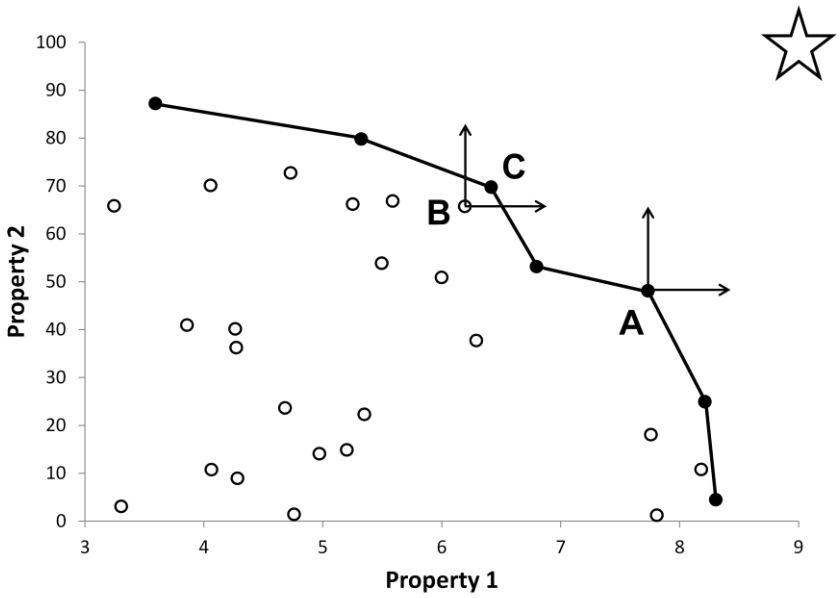


Figure 2. Illustration of the concept of Pareto optimality for compounds represented by points in a plot of Property 1 versus Property 2. The ideal goal corresponds to the top right corner of the plot, as indicated by the star. Solid points are Pareto optimal or 'non-dominated'; in the case of point A, there are no points with a higher value for both parameters. However, open circles are not Pareto optimal; for example point B is 'dominated' by point C.

One limitation of Pareto optimisation is that the number of optimal compounds increases exponentially with the number of properties being considered. Therefore, in practice, the number of optimal compounds becomes too large to be useful when considering more than approximately 4 properties. This problem is further compounded if the uncertainty in the data is taken into account, because the number of compounds with a significant probability of being non-dominated will often be considerably larger. Therefore, Pareto optimisation is often combined with other MPO methods to combine multiple factors into a smaller number of dimensions in which the optimisation is conducted [26].

2.5. Considering Uncertainty

There is increasing recognition of the importance of taking account of the uncertainty in the underlying data in the selection and optimisation of compounds. As noted above, the statistical uncertainties in predictions by *in silico* models of compound properties remains high. Therefore, unless MPO methods explicitly reflect the impact of combining multiple, uncertain data points into an overall assessment of compound quality, there is a high risk of incorrectly rejecting good compounds due to uncertain predictions.

A recent paper by Debe *et al.* [27] described the application of their ALOHA method that combines the results of multiple binary classification models of properties including cell permeability, solubility, fraction absorbed and unbound clearance in humans. However, unlike a simple filtering approach, each model predicts the *probability* that a compound will pass the corresponding property criterion and the authors calculate a score for each compound corresponding to the probability that the compound will pass all of the classification models. Thus, the uncertainties in the property classifications of a compound are explicitly taken into account and a compound will not be unduly penalised by an uncertain classification.

A method published by Nissink and Degorce [11], based on desirability functions, also considers the potential for errors in the overall desirability of a compound due to the uncertainty in the underlying predicted or experimental compound data. They consider the probability that the desirability of each compound property is greater or less than the value assigned and combine these into an overall confidence parameter for the compound score. While this is not a rigorous estimate of the uncertainty in the score, it provides an indication of cases where a compound's score should be treated with caution.

The 'Probabilistic Scoring' method of Segall *et al.* [8] is also based on the foundation of desirability functions to define the profile or properties required for a successful compound. In this case, the compound property values and their uncertainties are used to calculate the probability of a compound achieving the ideal profile, defined by desirability functions for the individual properties, taking into account the importance of each criterion and the acceptable trade-offs. Furthermore, the uncertainty in the score for each compound is estimated, providing a clear indication when compounds may be confidently distinguished based on the available data, as illustrated in Figure 3.

2.6. Drug-like Properties

Numerous simple compound characteristics or physicochemical properties have been used to define the properties of 'drug-like' compounds. These include: MW, number of hydrogen bond donors (HDB) and acceptors (HBA), clogP, the octanol:water partition coefficient at pH 7.4 (clogD), polar surface area (PSA), acid dissociation coefficient (pKa), the number of rotatable bonds (ROTB), counts of structural alerts of reactive or potentially toxic functionalities (ALERT), fraction of sp^3 Carbons (FSP3) and the number of aromatic rings (AROM).

Numerous rules have been developed for the selection of drug like compounds based on these properties. The first of the well-known rules was Lipinski's 'Rule of Five' (RoF) [28] which has been followed by many more, such as: the rules of 4/400 [29] and 3/75 [30]; rules relating ROTB and PSA to oral bioavailability [31]; FSP3 to clinical success [32]; AROM [33] to 'developability'; MW and logD to *in vivo* clearance and oral absorption [34]; and many others [35]. This explosion in the number of these 'rules of thumb' to guide the selection of compounds has been termed "Ro5 envy" [36].

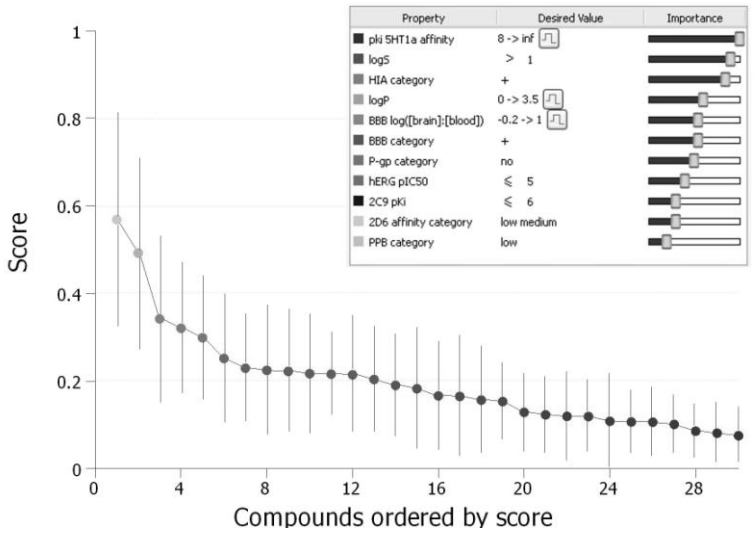
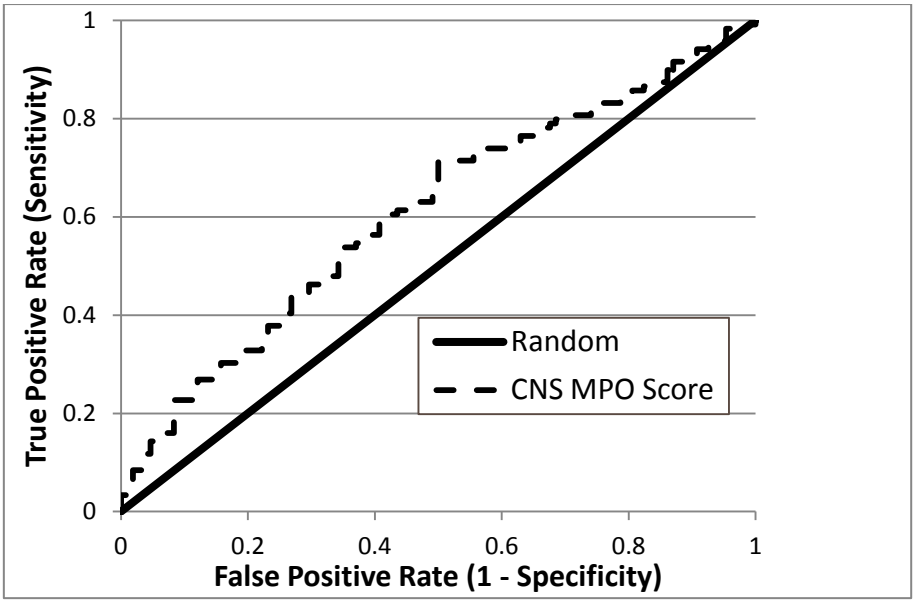


Figure 3 . An example output from probabilistic scoring for 30 compounds. The compounds are ordered from left to right along the x-axis in order of their score and overall score for each compound is plotted on the y-axis. The overall uncertainty in each score (one standard deviation), due to the uncertainty in the underlying data, is shown by error bars around the corresponding point. From this it can be seen that approximately the bottom 50% of compounds may be confidently rejected, as their error bars do not overlap with that of the top-scoring compound.

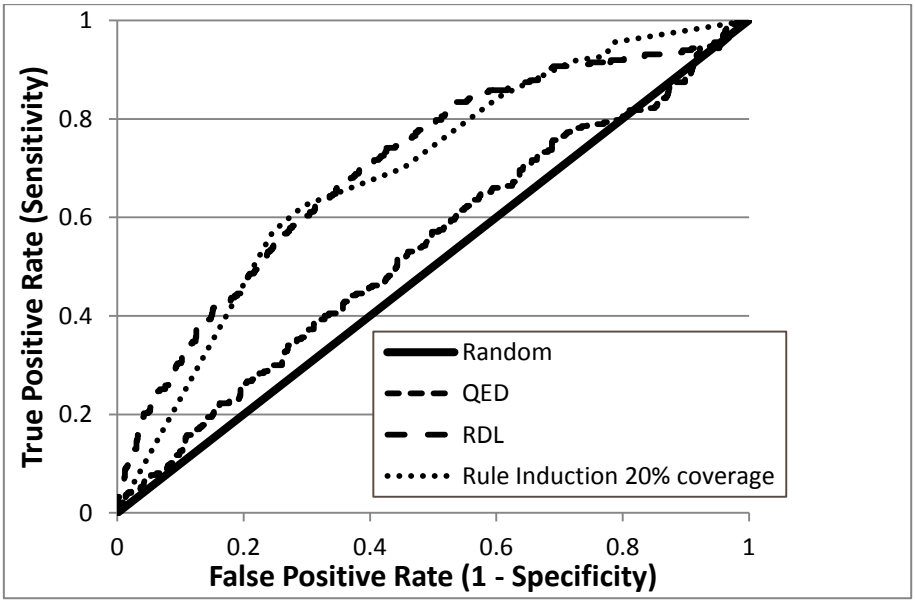
The scores have been calculated against the inset scoring profile, showing the property criteria and importance of each criterion to the overall project objective. Underlying each criterion is a desirability function.

While some of these rules may provide useful guidelines, they should be applied with a number of important caveats in mind. First, many of these rules are based on observations of the properties that successful drugs have in common, not the properties that make successful drugs different from unsuccessful compounds explored in drug discovery [37]; simply because a compound is drug-like does not mean that it is likely to be a drug. The rules are derived with specific objectives in mind, typically a small molecule, orally administered drug. However, there is a tendency to treat these rules as general definitions of 'drug-likeness' which may be misleading if applied in the context of a project with a different objective. In addition, these rules are often applied as pass/fail filters which, as discussed above, often make inappropriately harsh distinctions between similar compounds, particularly in light of the poor correlation of the simple properties on which they are based with the *in vivo* disposition of a compound. Finally, the very simplicity and memorability of these rules may lead to unconscious biases in decision-making in drug discovery, a psychological effect known as 'anchoring' [18] [38].

More sophisticated approaches, avoiding rules or filters, have been used to combine the values of drug-like properties into a composite score, using desirability functions. For example, Wager *et al.* defined desirability functions for six commonly used properties (MW, logP, logD, PSA, HBD and pKa of the most basic nitrogen) into a score for selection of compounds intended for central nervous system (CNS) indications [39]. This 'CNS MPO' score was calculated by summing the values of the desirability functions to give a value between 0 and 6. The authors found that from a data set of 119 marketed CNS drugs and 108 failed Pfizer candidates for CNS indications, 74% of the drugs had a CNS MPO > 4, but only 60% of the failed candidates. The author's also found that a high CNS MPO score corresponded to an increased chance of a good outcome in a variety of *in vitro* assays for permeability, cytotoxicity, metabolic stability and inhibition of the hERG ion channel. However, one needs to be careful not to over interpret such results. For example, in the same data set, 75% of drugs have a MW < 350 while only 44% of failed candidates meet this threshold; a better discrimination than provided by the rule CNS MPO > 4. Yet, we would not advocate selection of compounds for CNS indications based on a criteria of MW < 350. A more rigorous assessment of the performance of a classification method is provided by a receiver operating characteristic (ROC) curve, such as that shown in Figure 4(a) for the selection of CNS drugs over failed candidates from the data set published by Wager *et al.*. This shows that the performance of CNS MPO in this case is not much better than random selection. A similar outcome was found by Debe *et al.* when they applied the CNS MPO score in an attempt to discriminate between 250 marketed neuroscience drugs and a background set of 'leads' (compounds with micromolar or better inhibition of a drug target) [27].



(a)



(b)

Figure 4. ROC plots of the true positive rate (TPR (sensitivity)) against the false positive rate (FPR (1 - specificity)) for the classification of compounds. A perfect classifier would be represented by the point in the top left and a performance below the identity line indicates worse performance than a random classification. A greater area under the curve (AUC) for a classifier indicates higher performance, a random selection will have an AUC of 0.5 and the maximum AUC is 1.

(a) ROC plot for selection of CNS drugs using the CNS MPO score for a data set containing 119 marketed CNS drugs and 108 failed Pfizer candidates derived from reference [34]. The AUC for CNS MPO score is 0.61.

(b) ROC plot for classification of compounds as orally absorbed drugs or otherwise using RDL, QED, and a profile generated by rule induction and shown in Error! Reference source not found.**. In this case, a set of 247 orally administered drugs was differentiated from 1,000 randomly selected compounds from ChEMBL; the AUC for QED is 0.52, RDL is 0.70, and rule induction is 0.69.**

Another application of desirability functions to MPO of drug-like properties is the Quantitative Estimate of Drug-likeness (QED), as described by Bickerton *et al.* [40]. In this case, the authors fitted desirability functions to the distributions of eight properties (MW, clogP, HBD, HBA, PSA, ROTB, AROM and ALERT) for a data set of 771 oral drugs. The highest desirability in each case was assigned to the property value corresponding to the largest proportion of the marketed drugs. The overall QED was calculated by taking the geometric mean of the individual property desirabilities. Therefore, a compound with a high QED will have similar properties to the majority to oral drugs. However, as noted above, a compound that is similar to known oral drugs does not necessarily have a higher chance of being an oral drug. To illustrate this, Figure 4(b) shows the results of applying QED in an attempt to distinguish an independent set of 247 oral drugs from 1000 randomly selected non-drug compounds from the ChEMBL database [41] of compounds published in medicinal chemistry journals [42]. From this, we can see that selection on the basis of the QED does not differ significantly from random selection, suggesting that QED is unlikely to provide a powerful metric with which to guide the design of compounds with a higher chance of success. A low enrichment was also found by Debe *et al.* when QED was applied to a set of compounds comprising 250 marketed neuroscience drugs and 250 'leads' [27].

To demonstrate the difference between metrics defined by similarity to known drugs and approaches based on identifying the differences between drugs and non-drugs, Yusof and Segall derived a metric called the Relative Drug Likelihood (RDL) [37]. This method fitted 'likelihood functions' for the same properties as QED that indicate property values where the likelihood of a drug being found is highest, relative to the background of non-drugs. These functions were fitted by comparison of the set of 771 oral drugs from the QED paper with 1000 randomly selected non-drugs from the ChEMBL database [41]. The overall RDL was calculated by taking the geometric mean of the individual property likelihoods. The results of applying the RDL to distinguish an independent set of 247 oral drugs from 1000 different non-drug compounds from the ChEMBL database is also shown in Figure 4(b), indicating that the performance of this approach for this problem is significantly better than QED. However, Yusof and Segall also noted that the form of the likelihood functions vary significantly for drugs intended for different target classes or therapeutic indications, casting doubt on the viability of any general definition of drug-likeness.

A final, important observation regarding the use of multiple drug-like properties for the selection of compounds is that many of these properties are correlated; for example, clogP, clogD and PSA. Including multiple, correlated properties in the calculation of a metric can lead to 'overcounting' of a single factor, artificially biasing the selection of compounds. This commonly results from considering each property individually and combining the resulting criteria post-hoc to calculate an overall score. To address this, Yusof *et al.* introduced a 'rule induction' method that considers multiple properties simultaneously to identify compound selection rules based on property criteria that *in combination* select compounds with a higher chance of success [42]. Furthermore, this method can be applied to any objective for which successful and unsuccessful compounds are known to create a property profile tailored to that specific objective. To illustrate this, Figure 4(b) also shows the ROC curve for an independent test of the discrimination of oral drugs from non-drugs, based on the property criteria shown in Figure 5. This profile uses only 5 of the 8 properties employed by QED and RDL and the results indicate that the additional 3 properties add little value in making this selection.

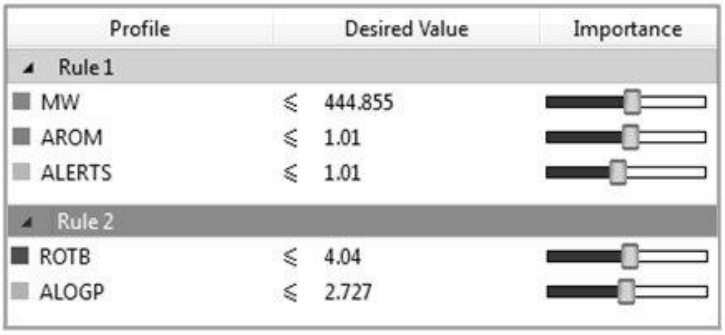


Figure 5 A profile of property criteria derived using rule induction for the selection of oral drugs based on simple drug-like properties. This profile is based on two separate rules based on multiple property criteria and each criterion is represented by a desirability function. The score for a compound will be the highest derived from either of the rules. An example of the performance of this profile on an independent test set is shown in Figure 4(b).

3. Developments in *de novo* design

The term 'de novo design' describes the application of computational methods to automatically generate new compound structures in the search for an optimal compound. The original de novo design methods were based on the optimisation of binding affinity against a target, using structure-based approaches to generate molecules that optimise the fit of a compound to a protein binding pocket, in terms of shape and interactions with binding residues [43]. These methods typically proceed by 'growing' a small fragment known to bind weakly within the binding pocket or linking two or more fragments that bind in different regions of the pocket. The success of this first generation of structure-based de novo design methods has been limited, due to their tendency to generate compounds that are synthetically intractable or have poor ADME or physicochemical properties; although the results can be improved by post-filtering of compounds [44].

In order to address these limitations, many recent developments in the field of de novo design have focused on ligand-based methods for the generation of relevant compound structures, guided by MPO methods, in an attempt to balance the optimisation of potency with other properties required in a successful compound.

3.1. Medicinal Chemistry Transformation Rules

The application of compound transformation rules, derived from medicinal chemistry experience, to generate new compound structures based on an initial input compound, was pioneered by the 'Drug Guru' (drug generation using rules) platform developed by Stewart *et al.* of Abbot Laboratories [45]. These 'medicinal chemistry transformation rules' correspond to typical changes explored by chemists in the optimisation of a compound or series. While they do not correspond to specific reactions, they represent tractable steps in chemistry space and therefore increase the likelihood of the resulting compounds being synthetically feasible. In tests, typically 90-95% of the compound structures generated were acceptable to medicinal chemists and the authors described how new compound structures, generated by Drug Guru, can be sorted or prioritised by calculated physicochemical properties.

Ekins *et al.* [46] combined the Drug Guru approach with Pareto optimisation to create 'Pareto Ligand Designer'. This platform automatically 'evolves' new compound structures, by iteratively applying medicinal chemistry transformation rules and selecting compounds using property filters and Pareto optimisation from each 'generation' as the basis for the next. In one example, the authors described the application of Pareto Ligand Designer to a known CCK antagonist [47], with poor blood-brain barrier (BBB) penetration and poor aqueous solubility. Predictions of these properties from quantitative structure-activity relationship (QSAR) models were used as the basis for the selection of Pareto optimal compounds from each generation and filters were also applied to remove compounds with MW > 500, clogP > 5, undesirable substructures or a Tanimoto similarity of < 0.35 with the initial compound. The authors observed an improvement in both predicted BBB penetration and solubility for the resulting compounds, indicating a better balance of properties, while retaining sufficient similarity with the initial compound to give a reasonable likelihood of retaining acceptable target potency.

The Nova™ module of the StarDrop™ platform [48] also iteratively applies medicinal chemistry transformation rules to explore compound optimisation strategies, guided by Probabilistic Scoring to target a project-specific profile of predicted properties [49]. In a retrospective example, the authors described the application of this method to the lead compound that led to the discovery of the serotonin reuptake inhibitor Duloxetine, guided by QSAR models of properties including inhibition of the serotonin transporter, BBB penetration, aqueous solubility, logP, efflux by P-glycoprotein, inhibition of the hERG ion channel, plasma protein binding and human intestinal absorption. The resulting top-scoring compounds included the drug Duloxetine and also identified a compound with very high similarity to another clinical candidate serotonin reuptake inhibitor, Litoxetine.

As discussed, the molecular transformations used to generate new compounds in these approaches are typically derived from the practical experience of medicinal chemists. They often represent relatively small modifications, e.g. substitution of functional groups, addition of heteroatoms and replacement of isosteric fragments. The number of these common transformations is typically in the hundreds and they have the advantage that the resulting compounds are more likely to be synthetically tractable. The range and number of transformations can be extended by automatically mining databases of existing compounds from in-house collections, the literature or patents to identify pairs of compounds that differ only in the replacement of one substructure with another, while preserving all attachment points [50]. One potential downside of this

approach is that it can generate many transformations that occur rarely and are unlikely to be relevant, but this can be mitigated by considering the frequency of occurrence to assign a confidence to each transformation.

A balance between these two approaches is represented by the BIOSTER database [51] that comprises over 29,000 pairs of bioanalogous compounds, manually curated from the chemistry literature. In each case, the substructure replacement relating these compounds has been identified and, in approximately 85% of cases, this can be represented as a transformation that can be automatically applied in the context of de novo design. Each transformation is associated with the primary references from which it was derived and BIOSTER therefore represents a chemically and synthetically validated database of transformations, encompassing a broad range of structure modification and replacement techniques that have been used in analogue design, including: bioisosteric replacements, linker replacements, homologization, introduction of conformational constraints and reversible derivatizations (e.g. prodrugs). An example transformation derived from BIOSTER is shown in Figure 6.

3.2. Evolutionary Algorithms

An alternative approach to generation of new compound structures uses the principles underlying the theory of evolution to 'evolve' a population of compounds towards a goal [52]. These Evolutionary Algorithms (EAs) 'mutate' the compound structures in a population (e.g. breaking or forming bonds, introducing heteroatoms or substituting new atoms) or 'cross' different compounds by combining their structural features. The 'fittest' of the resulting compounds are then selected as the basis for the next generation and the process is repeated until one or more members of the population achieve the objective.

EAs have been most commonly applied to the optimisation of target binding affinity, using both structure- and ligand-based methods for the assessment of potency (see, for example, references [53] [54] [55]). However, the combination of EAs with MPO methods to evolve new compound structures, with the goal of simultaneously optimising multiple properties, is being increasingly explored [56] [57].

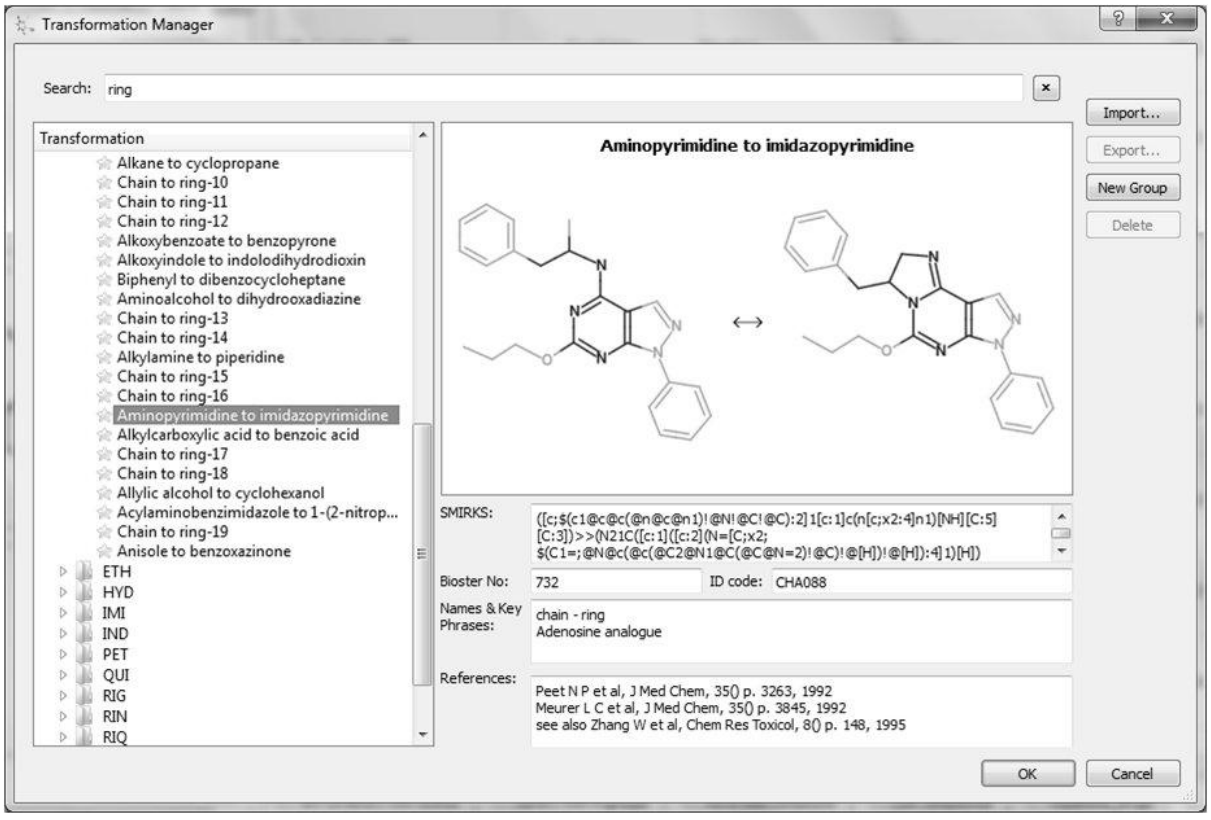


Figure 6 An example transformation from the BIOSTER database. The pair of molecules identified from the literature is shown, with the replacement region highlighted in bold. The corresponding transformation is shown below in SMIRKS format [64], with supporting information including the literature references from which the transformation was derived.

In a recent example, van der Horst *et al.* [58] described the de novo design of adenosine receptor (AR) ligands to simultaneously optimise affinity for the A_1 receptor, selectivity against other AR subtypes and multiple ADME- and toxicity-related properties (MW, PSA, Aqueous solubility, HBA and HBD, and Ames mutagenicity). Desirability functions were used to combine the ADMET properties into an overall ADMET desirability index. A desirability index was also calculated for selectivity against three AR subtypes (A_{2A}, A_{2B} and A_3), predicted using support vector machine models, and the affinity for A_1 was calculated using a pharmacophore model. Pareto optimisation was then used for the selection of compounds in each generation, following mutations and crosses, to simultaneously optimise the three parameters: pharmacophore score for AR A_1, for selectivity against other AR subtypes and ADMET properties. The resulting set of 3946 compounds were filtered according to pharmacophore score and novelty and grouped by scaffold. From these, 6 compounds representing different scaffolds were selected on the basis of ease of synthesis and tested for inhibition of the four ARs. Of these, 2 compounds showed low micromolar affinity for A_1, although these did not achieve sufficient selectivity over the other AR subtypes. Further optimisation, guided by the scoring scheme described above, yielded analogues with improved potency and selectivity.

Evolutionary algorithms can provide a wide diversity of potential new compounds. However, the compound modifications corresponding to the mutation and combination operations often result in synthetically intractable structures. Therefore, as illustrated in the previous example, the resulting compounds are usually heavily filtered to identify those that can be synthesised in practice.

4. Case Study

For some therapeutic indications, an ideal drug will interact with a single specific target to minimise the chance of toxicity or side effects; however, in other cases, for example some psychiatric indications, it is necessary for a drug to interact with multiple targets in order to achieve efficacy. Besnard *et al.* described a comprehensive study investigating the de novo design of novel compounds with specific polypharmacological profiles and appropriate ADME properties for a CNS indication [59].

In this study, the de novo design process was initialised with the structure of the drug Donepazil, an acetylcholinesterase inhibitor approved for cognitive enhancement in Alzheimer's patients. In addition to its primary activity, Donepazil also exhibits dopamine D4 receptor activity (K_i = 614 nM) and minimal dopamine D2 activity. New structures were generated by iterative application of medicinal chemistry transformations derived from mining of the ChEMBL database [41].

Besnard *et al.* applied an MPO scheme employing Bayesian target activity models, generated by the authors for 784 proteins, and predictions of BBB penetration and other ADME properties using the StarDrop software platform [48]. The ADME properties were combined into a single score which was optimised simultaneously with the predicted target activities. Compounds were selected from each iteration based on their distance from an optimal point in this MPO space, representing the required profile of target activities and good ADME properties. In each generation, the 10,000 highest scoring compounds and 500 random structures from the remaining population were selected as the input for the next iteration.

This approach was first applied to the improvement of D2 activity and achieving good blood-brain barrier penetration. In this case, a series of indoles was evolved, from which 8 compounds were synthesised and tested. All showed D2 receptor affinities in the range K_i = 156 – 1,700 nM. The most active compound (compound 3 shown in Figure 7) also exhibited a good *in vivo* brain-blood ratio (BBR) of 0.5.

In a second experiment, the isoindole series was further evolved in an attempt to identify compounds with activity against a polypharmacological profile (5-HT_{1A} serotonin receptor, D2-, D3- and D4-dopamine receptors) and selectivity over the α_{1A}-, α_{1B}- and α_{1C}-adrenoreceptors, while maintaining BBB penetration. The resulting benzolactam series, including compound 9a shown in Figure 7, exhibited a good profile of target activities (e.g. K_i for compound 9a: 5HT1A = 1.5 nM, D2 = 14 nM, D3 = 1 nM, D4 = 13 nM, α_{1A} = 117 nM, α_{1B} =649 nM, α_{1C} = 78 nM) and good brain penetration (BBR for compound 9a = 5.9).

The authors next explored the potential to find highly selective, potent D4 ligands with good BBB penetration, starting again from Donepezil. This was achieved in two steps, first optimising for D4 potency and BBB penetration, followed by further optimisation of D4 selectivity, resulting in a 2,3-dihydro-indol-1-yl chemotype, including compound 13, shown in Figure 7. Compound 13 had a potency against D4 of K_i = 8.9 nM with good

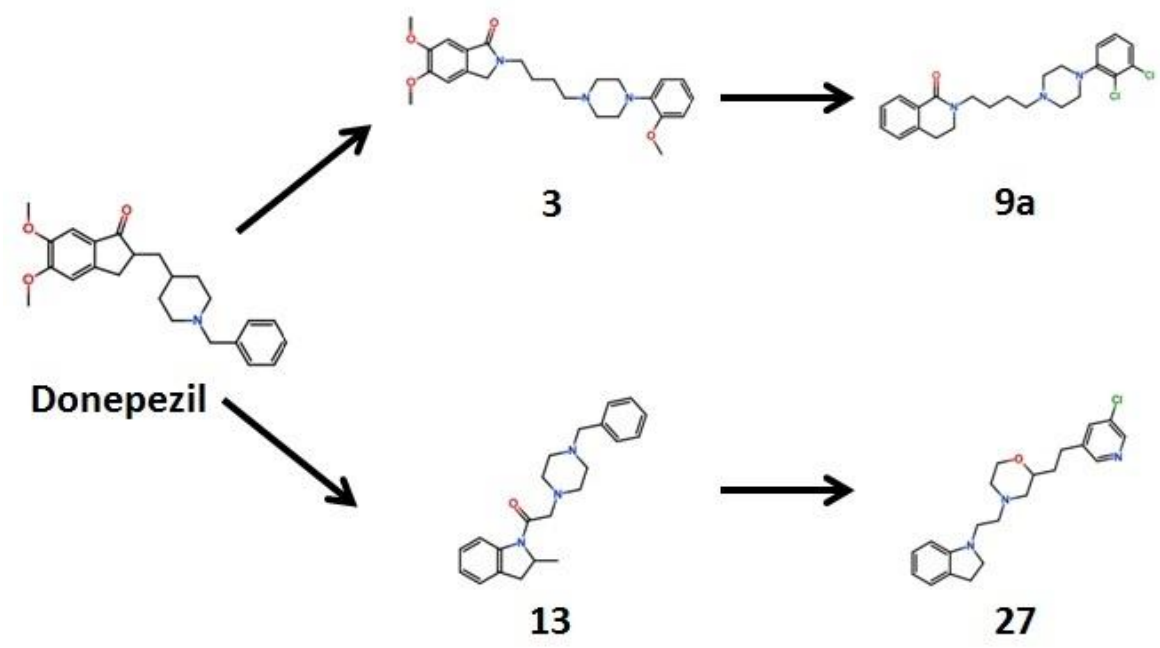


Figure 7 Key compounds in the automatic design of compounds to specified polypharmacological profiles beginning with the drug Donepezil, described in [54] and summarised in the Case Study herein.

selectivity over the panel of other G-protein-coupled receptors (GPCRs) and BBR = 7.5. This compound was further tested in an *in vivo* behavioural model in wild-type and D4 receptor knock-out mice and showed good evidence of target engagement.

Finally, compound 13 was used as the basis to evolve highly selective D4 dopamine antagonists with good BBB penetration and representing a novel chemotype. These resulted in a series of isoindol-1-yl-ethyl-morpholino analogues, including compound 27 (see Figure 7), which had D4 Ki = 90 nM, "exquisite" sensitivity for the D4 receptor over other GPCRs and BBR = 2. Most known D4 ligands are 1,4-distributed aromatic piperidines and piperazines, in common with many biogenic amine GPCRs, and therefore this series represented a new D4 chemotype.

5. Conclusion

As we have seen, developments in the application of MPO have resulted in dramatic improvements in the quality and relevance of the compound structures generated by de novo design. However, it should be emphasised that the objective of these methods is not to automatically design a final, perfect compound for synthesis, but to explore a diverse range of optimisation strategies and identify high quality ideas for detailed consideration by an expert scientist. The uncertainties in *in silico* model predictions mean that a computer cannot identify a compound with absolute confidence that it will achieve all of the required properties; however, as we have seen, the use of MPO methods can guide this exploration to identify new chemistries with a high probability of achieving the required property profile.

6. Expert Opinion

Recent developments in the field of de novo design have helped to overcome many of the issues encountered in the first generation of these methods. In particular, the combination of de novo design with MPO methods has helped to provide strategies for generation of new compound ideas that balance the optimisation of potency against the therapeutic target with other requirements of a successful lead or candidate drug.

The use of medicinal chemistry transformation rules to define potential compound modifications has helped to reduce the generation of irrelevant or chemically infeasible structures by de novo design methods. Approximately 95% of compound structures generated using rules based on practical medicinal chemistry experience have been found to be acceptable to experienced chemists [45] [49]. This makes the output of de

novo design systems more palatable for medicinal chemists because it is no-longer necessary to disregard large numbers of irrelevant compounds to find interesting ideas to consider in more detail. This encourages more routine use.

Advances in MPO have been based on the adaptation of methods from other disciplines, such as quality control [10] and economics [25], to the circumstances of drug discovery. In particular, this has focussed on approaches to explicitly consider the impact of uncertainty on our ability to confidently prioritise compounds when combining data for multiple properties into an overall score [8] [27] [11]. There is significant uncertainty in both experimental and predicted data generated in drug discovery and, in the context of de novo design, we should be aware of the large statistical errors that remain in the predictive methods used to guide the optimisation process. The objective of de novo design is to aid the exploration of a wide range of optimisation strategies and we do not wish to inappropriately reject potentially valuable ideas based on an uncertain predicted property value. However, it should be emphasised that the output of a *de novo* design method will be only as good as the underlying models that guide the optimisation process; if the model predictions are not sufficiently accurate, the compounds proposed may be misleading. It is particularly important to note if the novel structures generated lie outside of the domain of applicability of one or more models, implying that the confidence in the corresponding predictions will be low.

There has also been considerable focus on the use of simple drug-like properties and LEIs to guide the optimisation of compounds. While these provide useful guidelines to avoid venturing into high risk property space, e.g. large, lipophilic compounds, concerns have recently been raised about the ability of these metrics effectively to guide the optimisation of compounds at a more subtle level [17] [37]. Indeed, the negative impact of simple 'rules of thumb' that can bias the decision-making process may outweigh their positive effects [18] [38]. In the context of de novo design, the application of simple filters based on properties such as MW, logP and PSA is being replaced by more sophisticated scoring methods to guide the optimisation of balanced compounds without introducing inappropriate bias [49] [59] [58].

6.1. Future directions

One important area of future development relates to further analysis of the ease of synthesis of compounds suggested by de novo design methods. As noted above, the structures generated by recent methods are now more acceptable from a medicinal chemistry perspective. However, the ease with which they may be synthesised depends on several factors, including the reagents and reactions available and the experience of the chemist evaluating them. Studies have shown that assessment of the ease with which a compound may be synthesised is subjective [60]. However, several methods have been developed to estimate and score synthetic feasibility or accessibility [61] [60] [62] [63] and these could be used as an input to an MPO algorithm, to further improve the likelihood of the proposed structures being synthetically feasible, in addition to achieving the other property requirements. Indeed, in the case study above, Besnard *et al.* applied a synthetic accessibility score [63] as a filter to remove compounds which were unlikely to be synthesisable. A further step might be to couple the output of a de novo design algorithm to a method for retrosynthetic analysis [64] [65]. However, retrosynthetic methods rely on the availability of up-to-date reaction and reagent databases and their computational cost grows exponentially with the number of required reaction steps, which may make their application intractable in this context, where large numbers of potential compounds must be quickly assessed.

A further source of molecular transformations for de novo design may be provided by the recent development of methods for matched molecular pair analysis (MMPA) [66] [67]. Matched molecular pairs are compounds that differ only by one well-defined structural transformation and MMPA of large databases of compounds with experimental data can identify transformations that, on average, have a significant impact on a compound property, e.g. potency or a physicochemical or ADME property. Thus, instead of applying a general set of transformations to generate a large number of possible structures and applying predictive models to select compounds post-hoc, transformations derived from MMPA could be used to bias the search towards generating compounds with an improved property value. Raymond *et al.* suggested this as a future direction when describing the development of an algorithm for automatically extracting molecular transformations from databases of diverse molecular structures [50].

As noted above, the success of de novo design is governed by the quality of the underlying models used to direct the search for high quality compounds. The corollary of this is that de novo design methods will benefit

from the ongoing research into improvements in predictive models, driven by larger, higher quality data sets, improved modelling methods and molecular descriptors [68]. As models become more accurate and cover greater chemical diversity, this will enable de novo design methods to more confidently focus on optimisation strategies with the highest chance of success.

7. Acknowledgements

The author would like to thank his colleagues at Optibrium, past and present, for their input on the ideas discussed herein. He would particularly like to acknowledge Ed Champness and Nick Foster for their helpful suggestions on this manuscript.

Bibliography

1 Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates. Nature Rev. Drug Discov. 2004;3:711-716.

2 Paul SM, Mytelka DS, Dunwiddie DT, Persinger CC, Munos BH, Lindborg SR, Schacht AL. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat. Rev. Drug Discov. 2010;9:203-14.

3 CMR International, a Thomson Reuters business. CMR International Global R&D Performance Metrics programme.

4 Kennedy T. Managing the drug discovery/development interface. Drug Discov. Today. 1997;2:436–444.

5 Tarbit MH, Berman J. High-throughput approaches for evaluating absorption, distribution, metabolism and excretion properties of lead compounds. Curr. Opin. Chem. Biol. 1998;2:411-416.

6 Ekins S, Boulanger B, Swaan P, Hupcey M. Towards a new age of virtual ADME/TOX and multidimensional drug discovery. J. Comp. Aided Mol. Design. 2001;16:381-401.

7 Van de Waterbeemd H, Gifford E. ADMET in silico modelling: towards prediction paradise? Nat. Rev. Drug Discovery. 2003;2:192-204.

8 Segall M, Beresford A, Gola J, Hawksley D, Tarbit M. Focus on Success: Using in silico optimisation to achieve an optimal balance of properties. Expert Opin. Drug Metab. Toxicol. 2006;2:325-337.

9 Segall MD. Multi-Parameter Optimization: Identifying high quality compounds with a balance of properties. Curr. Pharm. Des. 2012;18:1292-1310.

** A comprehensive review of MPO methods applied in drug discovery

10 Harrington EC. The desirability function. Ind. Qual. Control. 1965;21:494-498.

11 Nissink JWM, Degorce S. Analyzing Compound and Project Progress Through Multi-Objective-Based Compound Quality Assessment. Future Med. Chem. 2013;5:753-767.

12 Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. Proc. Natl Acad. Sci. USA. 1999;96:9997-10002.

13 Hopkins AL, Groom CR, Alexander A. Ligand Efficiency: a useful metric for lead selection. Drug Discov. Today. 2004;9:430-431.

14 Abad-Zapatero C. Ligand efficiency indices for effective drug discovery. Expert Opin. Drug Discov. 2007;2:469-488.

15 Ferenczy GG, Keseru GM. Thermodynamics guided lead discovery and optimization. Drug Discov. Today. 2010;15:919-932.

16 Hopkins AL, Keseru GM, Leeson PD, Rees DC, Reynolds CH. The role of ligand efficiency metrics in drug discovery. Nat. Rev. Drug Discov. 2014;13:105-121.

** A comprehensive review of ligand efficiency metrics and their application in drug discovery.

17 Shultz MD. Setting expectations in molecular optimizations: Strengths and limitations of commonly used composite parameters. Bioorg. Med. Chem. Lett. 2013;23:5980-5991.

* This paper highlights some important limitations of commonly used ligand efficiency indices.

18 Shultz MD. Improving the Plausibility of Success with Inefficient Metrics. ACS Med. Chem. Lett. 2013;5:2-5.

* A very thought provoking paper describing the potential impact of simple rules of thumb for drug-like properties

19 Shultz MD. The thermodynamic basis for the use of lipophilic efficiency (LipE) in enthalpic optimizations. Boorg. Med. Chem. Lett. 2013;23:5992-6000.

20 Reynolds CH, Rounge BA, Bembenek SD. Ligand Bindinf Efficiency: Trends, Physical Bias and Implications. J. Med. Chem. 2008;51:2432-2438.

21 Nissink JWM. Simple size-independent measure of ligand efficiency. J. Chem. Inf. Model. 2009;49:1617-1622.

22 Abad-Zapatero C. Ligand Efficiency Indices for Drug Discovery: Towards an Atlas-Guided Paradigm. Oxford: Elsevier; 2013.

23 Abad-Zapatero c, Metz JM. Ligand Efficiency Indices as Guideposts for Drug Discovery. Drug. Discov. Today. 2005;10:464-469.

24 Abad-Zapatero C, Champness EJ, Segall MD. Alternative variables in drug discovery: promises and challenges. Future Med. Chem. 2014;(in press).

25 Jaffe W. Pareto translated: A review article. J. Economic Literature. 1972;10:1190-1201.

26 Kruisselbrink J, Michael E, Back T, Bender A, Ijzerman A, van der Horst E, Ehrgott M, Fonesca C, Gandibleux X, Hao JK, et al. Combining aggregation with Pareto optimization: A case study in evolutionary molecular design. In: Proceedings of EMO '09: Lect. Notes Comput. Sci.; 2009; Berlin/Heidelberg. p. 453-467.

27 Debe DA, Mamidipaka RB, Gregg RJ, Metz JT, Gupta RR, Muchmore SW. ALOHA: a novel probability fusion approach for scoring multi-parameter drug-likeness during the lead optimization stage of drug discovery. J. Comput.-Aided Mol. Des. 2013;27:771-82.

28 Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 1997;23:3-25.

29 Gleeson MP. Generation of a set of simple, interpretable ADMET rules of thumb. J. Med. Chem. 2008;51:817-834.

30 Hughes JD, Blagg J, Price DA, Bailey S, Decrescenzo GA, Devraj RV, Ellsworth E, Fobian YM, Gibbs ME, Gilles RW, et al. Physicochemical drug properties associated with in vivo toxicological outcomes. Bioorg. Med. Chem. Lett. 2008;18:4872-4875.

31 Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. 2002;45:2615-2623.

32 Lovering F, Bikker J, Humblet C. Escape from flatland: increasing saturation as an approach to improving clinical success. J. Med. Chem. 2009;52:6752-6756.

33 Ritchie TJ, Macdonald SJF. The impact of aromatic ring count on compound developability – are too many aromatic rings a liability in drug design? Drug Discov. Today. 2009;14:1011-1020.

34 Johnson TW, Dress KR, Edwards M. Using the Golden Triangle to optimize clearance and oral absorption. Bioorg. Med. Chem. Lett. 2009;19:5560-5564.

35 Garcia-Sosa AT, Maran U, Hetenyi C. Molecular Property Filters Describing Pharmacokinetics and Drug Binding. Curr. Med. Chem. 2012;19:1646-1662.

36 Kenny PW, Montanari CA. Inflation of correlation in the pursuit of drug-likeness. J. Comput.-Aided Mol. Des. 2013;27:1-13.

* Discusses "Rule of 5 envy" and pitfalls in the development of drug-likeness metrics.

37 Yusof I, Segall MD. Considering the impact drug-like properties have on the chance of success. Drug Discovery Today. 2013;18:659-666.

38 Chadwick AT, Segall MD. Overcoming psychological barriers to good discovery decisions. Drug Discov. Today. 2010;15:561-569.

39 Wager TT, Hour X,VPR, Villalobos A. Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. ACS Chem. Neurosci. 2010;1:435-449.

40 Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. Nature Chemistry. 2012;4:90-98.

41 Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012;40:D1100-D1107.

42 Yusof I, Shah F, Hashimoto T, Segall MD, Greene N. Finding Rules for Successful Drug Optimization. Drug Discov. Today. 2014.

43 Schneider G, Fechner U. Computer-based de novo design of drug-like molecules. Nature Reviews Drug Discovery. 2005;4:649-663.

44 Hartenfeller M, Schneider G. Enabling future drug discovery by de novo design. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2011.

45 Stewart K, Shiroda M, James C. Drug Guru: a computer software program for drug design using medicinal chemistry rules. Bioorg. Med. Chem. 2006;14:7011-22.

** The seminal paper introducing the concept of 'medicinal chemistry transformation rules' for the generation of new compound structures.

46 Ekins S, Honeycutt J, Metz J. Evolving molecules using multi-objective optimization: applying to ADME/Tox. Drug Discov. Today. 2010;15:451-60.

47 Evans BK, Rittle KE, Bock MG, DiPardo RM,FRM,WWL., Lundell GF, Veber DF, Anderson PS. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. J. Med. Chem. 1988;31:2235–2246.

48 Optibrium. [Internet]. [cited 2013 May 28]. Available from: http://www.optibrium.com/stardrop.

49 Segall MD, Champness EJ, Leeding C, Lilien R, Mettu R, Stevens B. Applying medicinal chemistry transformations to guide the search for high quality leads and candidates. J. Chem. Inf. Model. 2011;51:2967–2976.

50 Raymond J, Watson I, Mahoui A. Rationalizing lead optimization by associating quantitative relevance with molecular structure modification. J. Chem. Inf. Model. 2009;49:1952-62.

51 Ujváry I, Hayward J. BIOSTER: A Database of Bioisosteres and Bioanalogues. In: Brown N, editor. Bioisosteres in Medicinal Chemistry. Vol 54. Weinheim: Wiley-VCH; 2012. p. 55-74.

52 Clark DE, editor. Evolutionary Algorithms in Computer-Aided Molecular Design. Weinheim: Wiley-VCH; 2000.

53 Schneider G, Lee ML, Stahl M, Schneider P. De novo design of molecular architectures by evolutionary assembly of drug-derived building blocks. J. Comput.-Aided Mol. Des. 2000;14:487-494.

54 Damewood JR, Lerman CL, Masek BB. NovoFLAP: A Ligand-Based De Novo Design Approach for the Generation of Medicinally Relevant Ideas. J. Chem. Inf. Model. 2010;50:1296-1303.

55 Nugiel DA KJHDDJJR, Bernstein PR, Sobotka-Briner CD, Liu J, Zacco A, Pierson ME. De novo design of a picomolar nonbasic 5-HT(1B) receptor antagonist. J. Med. Chem. 2010;53:1876-1880.

56 Nicolaou CA, Apostolakis J, Pattichis CS. De novo drug design using multiobjective evolutionary graphs. J. Chem. Inf. Model. 2009;49:295-307.

57 Nicolaou CA, Brown NA, Pattichis CS. Molecular optimization using computational multi-objective methods. Curr. Opin. Drug Discov. Devel. 2007;10:316-324.

58 van der Horst E, Marqués-Gallego P, Mulder-Krieger T, van Veldhoven J, Kruisselbrink J, Aleman A, Emmerich MT, Brussee J, Bender A, Ijzerman AP. Multi-objective evolutionary design of adenosine receptor ligands. J. Chem. Inf. Model. 2012;52:1713-1721.

59 Besnard J, Ruda GF, Setola V, Abecassis K, Rodriguiz RM, Huang XP, Norval S, Sassano MF, Shin AI, Webste LA, et al. Automated design of ligands to polypharmacological profiles. Nature. 2012;492:215-220.

** A comprehensive study demonstrating the prospective application of de novo design guided by MPO, with experimental validation.

60 Boda K, Seidel T, Gasteiger J. Structure and reaction based evaluation of synthetic accessibility. J. Comput.-Aided Mol. Des. 2007;21:311-325.

61 Baber JC, Feher M. Predicting synthetic accessibility: application in drug discovery and development. Mini Rev. Med. Chem. 2004;4:681-692.

62 Allu TK, Oprea TI. Rapid Evaluation of Synthetic and Molecular Complexity for in Silico Chemistry. J. Chem. Inf. Model. 2005;45:1237-1243.

63 Ertl P, Schuffenhauer A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J. Cheminfom. 2009;1:8.

64 Ihlenfeldt WD, Gasteiger J. Computer-Assisted Planning of Organic Syntheses: The Second Generation of Programs. Angew. Chem. Int. Ed. Engl. 1995;34:2613-2633.

65 Gillet VJ, Myatt G, Zsoldos Z, Johnson P. SPROUT, HIPPO and CAESA: Tools for de novo structure generation and estimation of synthetic accessibility. Perspect. Drug Disc. Design. 1995;3:34-50.

66 Leach AG, Jones HD, Cosgrove DA, Kenny PW, Ruston L, MacFaul P, Wood JM, Colclough N, Law B. Matched molecular pairs as a guide in the optimization of pharmaceutical properties; a study of aqueous solubility, plasma protein binding and oral exposure. J. Med. Chem. 2006;49:6672-6682.

67 Dossetter AG, Griffen EJ, Leach AG. Matched Molecular Pair Analysis in drug discovery. Drug Discov. Today. 2013;18:724-731.

68 Gola J, Champness E, Segall M. ADMET property prediction: The state of the art and current challenges. QSAR Comb. Sci. 2006;25:1127-29.

69 Daylight Chemical Information Systems Inc. Daylight Theory Manual. [Internet]. [cited 2013 May 28]. Available from: www.daylight.com/dayhtml/doc/theory/theory.smirks.html.