Expert Opinion

informa healthcare

The quest for novel chemical matter and the contribution of computer-aided *de novo* design

Bernard Pirard
Novartis Institute for BioMedical Research, Global Discovery Chemistry, Computer-Aided Drug Discovery, Basel, Switzerland

Identifying novel chemical matter is the focus of many drug discovery efforts. Through these efforts, computer-based *de novo* design of drug-like molecules, which aim to build an entire molecule 'from scratch', has emerged as a valuable approach to identify novel chemical matter. In this paper, the author discusses the recent research efforts that aim to build, *in silico,* more chemically accessible molecules, sample more efficiently the chemical space and rank the proposed molecules. The author reviews *de novo* design algorithms developed between 2008 and 2010 and the issue of validation, and highlights some recent successful applications of *de novo* design to drug discovery projects. Although research has addressed the lack of synthetic accessibility of the molecules proposed by the first generation of *de novo* design tools, the lack of accurate scoring function remains a major limitation of structure-based *de novo* design. However, *de novo* design is a valuable approach to generate either chemical starting points or ideas.

Keywords: *de novo* design, fragment space, hit finding, multi-objective scoring, natural computing, scoring

Expert Opin. Drug Discov. (2011) **6**(3):225-231

1. Introduction

Lead identification is a key activity in small-molecule drug discovery. In a recent analysis of 60 lead/drug pairs, Perola observed that > 60% of the leads were based on molecules that were previously known to be related to the target [1]. Although a completely novel lead is not always a prerequisite for a successful drug discovery project, the identification of a novel starting point enables to secure intellectual property rights. In addition, novel small-molecules might be necessary to modulate targets which so far have not been the focus of drug discovery efforts. Over the last two decades, both experimental techniques such as high-throughput screening (HTS) or fragment-based screening (FBS) and computational approaches such as virtual screening and *de novo* design have emerged as valuable and complementary sources for novel starting points for small-molecule drug discovery projects [2-5]. Virtual screening encompasses a variety of computational approaches which aim to reduce a large collection of compounds to a short list of screening candidates by applying an information-dependent sequence of filters. On the other hand, *de novo* design aims to build *in silico* an entire molecule 'from scratch'. Like virtual screening, *de novo* design is an information-driven approach.

Herein, we review the recent developments in the field of *de novo* design with an emphasis on methods, applications and challenges. The first automated *de novo* design software packages appeared nearly two decades ago [5]. Nearly all of them focused on building 3D structures of candidate molecules within a binding site and on scoring each proposed molecule based on its fit with the binding site. These algorithms had the tendency to generate overly complex molecules that lack

10.1517/17460441.2011.554394 © 2011 Informa UK, Ltd. ISSN 1746-0441

synthetic accessibility. More recently, ligand-based *de novo* design programs have been described [5]. Methods to account for synthetic accessibility have also been developed [6]. *De novo* design is an iterative process which consists of three operations: design, sampling and evaluation. During the design step, molecules are generated using atoms or molecular fragments and a set of connection rules. Then, optimization algorithms are required to efficiently sample a huge chemical space. Finally, these molecules are scored, using a fitness function. If the 3D structure of the target is available, docking scores can serve as fitness function. The actual scoring function depends on the amount of information available within a project as well as on the objectives of the *de novo* design exercise. In the following paragraphs, these different steps are reviewed. Then, we describe recent methods for *de novo* design. We also outline validation studies as well as applications of *de novo* design.

2. Molecule construction

Building molecules from scratch requires a set of building blocks and connection rules. Building blocks are either single atoms or molecular fragments. Most of the earlier *de novo* design methods utilized atom-based building blocks [7]. These methods built molecules sequentially, adding one atom after the other. In principle, they could produce any ligand, meeting specified connection rules and design constraints. However, this was often achieved at the expense of the limited synthetic accessibility of the designed molecules. On the other hand, fragment-based approaches which are utilized by most modern *de novo* design methods tend to capture the synthetic accessibility of the designed molecule at the expense of a significant reduction of the size of the chemical space to be searched.

Mauser and Stahl introduced the concept of chemical fragment spaces, defined as combinations of molecular fragments and connection rules [8]. They developed an approach to generate a fragment space of 2000 – 4000 members. Their strategy consists of three steps: retrosynthetic fragmentation of a database of drug-like molecules, stepwise filtering to remove large fragments and undesirable structural motifs, and selection of a subset of fragments to maximize the coverage of the known chemical space. A set of compatibility rules that are encoded in each fragment definition determines which fragments can be recombined to generate new molecules. Degen *et al.* developed a similar approach known as breaking of retrosynthetically interesting chemical substructures (BRICS) [9]. BRICS applies all retrosynthetic cuts simultaneously to a given molecule to avoid the generation of redundant fragments, includes substructure filters into the shredding procedure and incorporates more elaborate medicinal chemistry concepts. This leads to the definition of 16 chemical environments, characterized by link atoms of different types. These link atoms serve to define the connection rules for the recombination of fragments.

The fragment spaces described above are based on a set of disconnection rules that describe a restricted number of simple and widely applicable reactions such as amide bond formation or Suzuki coupling [10]. To overcome this limitation, Patel *et al.* developed a knowledge-based approach to molecule generation based on reaction vectors [11]. Briefly, reaction vectors describe the structural changes taking place at the reaction centers as well as the environment in which the reaction occurs. The reaction vectors are derived automatically from a large collection of single-step reactions (up to the limit of two reactant and two product reactions). The transformations derived from the reaction database can be applied to a starting molecule to generate a new molecule.

3. Sampling the chemical space

Although fragment-based *de novo* design approaches significantly reduce the size of the search space, the number of possible combinations of fragments remains huge. Therefore, exhaustive construction and evaluation of all the possible solutions, that is, molecular structures as well as their configurational degrees of freedom are not usually feasible within a reasonable time frame.

Structure-based *de novo* design methods handle fragments docked in a binding site, which provide the so-called target constraints. Several procedures have been developed to assemble molecules by connecting predocked fragments. These include fragment linking, fragment growing and structure sampling [5,7]. Fragment linking approaches connect building blocks docked in spatially close binding pockets either directly or through linker fragments. Linking fragments has a number of limitations. First, small fragments can exhibit alternative arrangements in a binding site, which can result in a large number of alternative binding modes for the constructed molecule. Second, linking tightly fitting fragments might lead to a strained conformation for the overall molecule. Finally, in some cases, fragment linking might be difficult or even impossible because the fragments are both too far to be joined directly and too close to add a linker without geometrical distortions at the junction. On the other hand, growing approaches consider a single fragment optimally placed in a binding site as a starting point for an iterative and stepwise construction phase. Conformational flexibility of the fragments is one of the main problems with growing approaches. Structure sampling also considers a single molecule randomly constructed in a binding site as a starting point for transformations. These include fragment addition, subtraction and replacement as well as whole molecule translation, rotation or conformational changes. Ligand-based *de novo* design methods which mostly operate on molecular graphs can also use structure sampling approaches such as fragment addition, subtraction or substitution. One can also apply simulation techniques such as molecular dynamics or Monte Carlo methods to generate new

molecules from a set of fragments randomly positioned in a binding site. However, these simulation techniques offer less control over the generated chemical structures than the other sampling techniques.

Over the last decade, a class of techniques, often referred to as 'natural computing' techniques, have emerged as robust sampling methods for fragment-based *de novo* design [12]. Representatives of natural computing techniques that are most frequently used in *de novo* design include evolutionary algorithms (EA) such as genetic algorithms and particle swarm optimization.

EAs mimic biological evolution. Starting from a population randomly selected within the search space, they iteratively repeat a sequence of scoring, selection and variation. To begin with, each individual (a ligand candidate in the case of *de novo* design) is scored by a fitness function. Structural similarity as measured by molecular descriptors such as topological pharmacophores often serves as fitness function in ligand-based *de novo* design, while structure-based approaches use scoring functions used in docking. As a result, fitter individuals (e.g., molecules exhibiting the highest pharmacophore similarity to a reference molecule) are more likely to be selected to produce offspring for the next generation. Then, genetic operators are applied to generate variations of selected parents' solution. Common genetic operators include mutation and crossover. Mutation usually produces individuals that differ marginally from their parents, but it can result in the introduction of totally new features into the population. In particular, a mutation operator can select a fragment from a parent molecule and replace it by a fragment from a database to create a new molecule. In contrast, crossover only involves the recombination of features that are already present within the population. Typically, a crossover procedure works on two parental structures. It dissects both structures, selects one compatible fragment from each parent, swaps them and produces two children, representing mixtures of fragments present in their parents. This iterative process is repeated until a termination criterion is reached. Because each new generation consists of variations of the fittest individuals from the previous one, the process is expected to bring up optimized individuals over time.

Particle swarm optimization mimics the behavior of biological swarms. They consist of individuals also known as particles which work in parallel to optimize a given problem. Each of them produces a solution (i.e., a molecule in the case of *de novo* design) which is subsequently scored. Each particle stores its best solution found so far in its own memory. In addition, it has access to a so-called social memory where the best solution found so far by the whole swarm is kept. Individual and collective memories which are constantly updated during the search guide the swarm through the search space to regions featuring good solutions. At the same time, the swarm tends to ignore less promising areas of the search space because no attracting point is located there.

4. Scoring

Sampling algorithms rely on a scoring function to efficiently explore the search space. Structure-based *de novo* design algorithms use the same kind of scoring functions as docking programs [13]. Some of these scoring functions take solvation into account, while others do not. In ligand-based *de novo* design, candidate molecules are ranked using similarity-based fitness function or quantitative structure–activity relationship (QSAR) models [14,15]. Similarity-based scoring requires the choice of both molecular descriptors and a similarity index. Descriptors such as pharmacophore fingerprints, which are known to facilitate the discovery of new chemotypes, are frequently selected for similarity-based scoring. On the other hand, QSAR-based scoring requires highly predictive models. This prompted Erickson *et al.* to carry out an extensive evaluation of descriptors and statistical techniques [15]. As a result, they selected support vector machine on Merck atom pair derived fingerprints. Obviously, QSAR-based scoring is limited to target classes for which a large amount of activity data is already available.

In addition to affinity, one can also estimate the ease of synthesis (synthetic accessibility) of molecules generated by *de novo* design methods. Several computational approaches have been developed to assess synthetic accessibility [6]. Complexity-based approaches use a set of rules to estimate the complexity of the target molecule. They consider the number of stereocenters and the presence of features such as spiro-rings or non-standard ring fusion. Another group of methods are based on a retrosynthetic analysis of the target molecule, which can be time consuming [16,17]. These methods rely on reaction databases and on reagents lists. Recently, Ertl and Schuffenhauer developed a scoring scheme function based on a combination of fragment contributions and a complexity penalty [18]. The fragment contribution which is based on the analysis of 1 million of representative molecules from the PubChem database captures the historical synthetic knowledge. The complexity term takes into account unusual structural features such as large rings and stereocomplexity. Results from this scoring function show very good agreement with assessment of synthetic accessibility by experienced medicinal chemists.

Traditionally, *de novo* design methods have considered a single objective, mainly related to binding affinity to rank the proposed molecules. However, a suitable starting point for a drug discovery project must fulfill more criteria than just binding affinity. Therefore, the scores of molecules proposed by *de novo* design software can be calculated as weighted sum of several terms that assess the desirable features of the candidate molecule. In particular, Dey and Caflisch have implemented a linear combination of a force field-based binding energy (i.e., the sum of van der Waals and Coulomb terms from the CHARMm force field) and similarity to a reference ligand structure selected by the user [19]. One challenge with a weighted sum of several terms is to set the weights of the

different term, as it is not always clear how they should be ranked. Pareto-based methods are an alternative to linear combinations of several terms. Briefly, Pareto-based methods aim to identify multiple solutions that represent different compromises among the various and sometimes conflicting properties (also referred to as objectives) under consideration. Recently, Nicolaou *et al.* described the multiobjective evolutionary graph algorithm which considers binding affinity, molecular similarity and physico-chemical scorers as input for Pareto ranking of the proposed molecules [20].

5. Recent *de novo* design methods

A list of *de novo* design tools described between 2008 and 2010 is reported in **Table 1**. For methods described before 2008, we refer the interested readers to earlier reviews [5,21]. In addition, **Table 1** contains information about the type of building blocks, sampling technique, scoring function and validation of each method. All these methods use fragments as building blocks. However, the source of fragments and the fragment generation protocol vary from one algorithm to the other. Natural computing techniques are most frequently used for sampling (five cases out of seven). The two remaining *de novo* design methods perform exhaustive sampling. Finally, most of the algorithms listed in **Table 1** use single objective scoring functions to rank the proposed molecules.

6. Validation of *de novo* design methods

Validating a computational tool for *de novo* design is not an easy task. In contrast to other computational chemistry techniques such as similarity searching, there is no single quality metrics for *de novo* design algorithms. Ultimately, a proper validation requires the synthesis and biological testing of the proposed molecules. For structure-based *de novo* design, one should also add to these basic requirements an experimental confirmation of the proposed binding mode hypothesis, that is, a crystal structure of the suggested molecule in complex with its target. However, this is time- and cost-intensive. Therefore, nearly all validation studies are retrospective. The authors of the NovoFLAP program have defined three criteria for a successful validation: i) the generation of chemically sensible and relevant ideas; ii) the generation of new ideas within a reference series and iii) the generation of a new chemical series, starting with at least one example from a reference series [22]. We should also mention that nearly all these validation studies are limited to a few selected examples (**Table 1**). In particular, Dey and Caflisch tested the performance of the SEED/GANDI algorithm on cyclin-dependent kinase 2 (CDK2), using a library of about 1400 fragments as building blocks and the binding mode of a known CDK2 inhibitor to bias the design [19]. As a result, some of the top-scoring SEED/GANDI hits were found to form the same hydrogen bond pattern as known CDK2 inhibitors. It is also noteworthy that one of the top-scoring molecules proposed by SEED/GANDI shares the same substructure element as a very potent CDK2 inhibitor. This substructure element is not present in the reference compound selected to bias the design.

In contrast to other researchers in the field, Zaliani *et al.* tested a comprehensive workflow for *de novo* design on 188 publically available crystal structures of protein–ligand complexes, corresponding to eight different targets from four protein families [23]. To start with, they generated a fragment space. Then, they investigated the ability of the fragment-based *de novo* design software FLEXNOVO to re-build known ligands of these proteins. They also analyzed the ability of FLEXNOVO to generate top-scoring solutions showing the same binding mode as the native ligand. For five out of eight proteins, FLEXNOVO could produce solutions identical to the reference ligand. In four of these five cases, the reference ligand appeared in the top five solutions, while the binding mode of the reference ligand could only be reproduced for three of them. During this validation study, it turned out that re-building highly flexible ligand binding to extended cavities remained challenging for FLEXNOVO.

It is also worth mentioning that retrospective validation of approaches to generate molecules from building blocks has also been reported. The aim of these validation studies was either to re-build a known molecule using a fragment space [8,9] or to reproduce the products of reactions stored in a reaction knowledge-base [11].

7. Applications of *de novo* design

Here, we present some recent successful applications of *de novo* design to drug discovery projects.

Application of the EA fragment-based *de novo* design tool TOPAS led to the identification of a novel chemical series of cannabinoid-1 receptor (CB1R) inverse agonists [24]. To start with, retrosynthetic fragmentation of 1381 GPCR modulators served to produce a collection of building blocks. These were subsequently recombined to generate molecules enriched with GPCR privileged motifs. The generated molecules were ranked based on their topological pharmacophore similarity to a reference inverse CB1 agonist as well as other parameters, including chemical tractability and patentability. One core structure from this *de novo* design exercise was selected for follow-up. Nine libraries were prepared, leading to the identification of a single digit micromolar CB1 ligand, which was subsequently optimized to a nanomolar lead compound. It is noteworthy that the *de novo* design and hit identification processes took only four months. In other words, application of ligand-based *de novo* design resulted in the rapid identification of a novel starting point in the crowded field of CB1R ligands. Recently, the ligand-based *de novo* design software SQUIRREL (**Table 1**) was applied to identify bioisosteric replacements for two key pharmacophoric features of the α subtype of a PPAR

Table 1. List of recent *de novo* design methods.

Name (year)	Type	Building blocks	Sampling technique	Scoring function	Validation (target)	Ref
SEED/GANDI (2008)	Structure-based	Fragments	GA + Tabu search	Weighted sum of force field-based scoring function and similarity terms	Retrospective (CDK2)	[19]
COLIBREE (2008)	Ligand-based	Fragments	PSO	Similarity on topological pharmacophoric descriptors	Retrospective (PPAR)	[32]
MEGA (2009)	Ligand- or structure-based	Fragments	EA on molecular graphs	Pareto on multiple objectives	Retrospective (ER)	[20]
SQUIRREL (2009)	Ligand-based	Fragments	Exhaustive	Shape and pharmacophore similarity	Prospective (PPAR-α and -γ)	[25]
PHDD (2010)	Ligand-based	Fragments	Exhaustive	Pharmacophore fit, computed physico-chemical properties and synthetic accessibility	Retrospective (HDAC, CDK2, HIV integrase)	[33]
NovoFLAP (2010)	Ligand-based	Fragments	EA	Shape alignment and matching of pharmacophoric features	Retrospective (NK1, CRF, AII, H_1) Prospective (5-HT_{1B})	[22]
GARLig (2010)	Structure-based	Fragments	GA	Scoring function of docking algorithms	Retrospective (serine proteases)	[34]

AII: Angiotensin II receptor; CDK2: Cyclin-dependent kinase 2; CRF: Corticotrophin releasing factor; EA: Evolutionary algorithm; ER: Estrogen receptor; GA: Genetic algorithm; H_1: Histamine I receptor; NK1: Neurokinin 1 receptor; PSO: Particle swarm optimization.

agonist [25]. One compound containing the replacements suggested by SQUIRREL was prepared and turned out to be inactive. Docking of this compound in the PPAR-α binding site suggested a potential reason for its inactivity. To test this hypothesis, an analog of the inactive compound was manually designed. This latter compound turned out to show submicromolar activity on the target. This latter example shows the importance of the role of molecular designer to analyze the list of potential molecules generated by a *de novo* design algorithm.

Over the last three years, several successful applications of structure-based *de novo* design have also been published [26-29]. In two cases, a fragment was grown into an unoccupied pocket of the binding site [26,27]. A *de novo* design algorithm (LigBuilder [30]) was applied in only one of the four applications [28]. The authors of this latter work insisted on the importance of implementing an accurate solvation model to score the proposed molecules. At this point, we would also like to highlight the efforts of Guo *et al.* to identify novel inhibitors of human Pin1, a member of the *cis-trans* peptidyl-prolyl isomerase, which offer some potential as an anticancer target [29]. In this project, both HTS and target family focused screening campaigns failed to identify any inhibitor. Using the crystal structure of the *apo* human PIN1, Guo *et al.* started a manual *de novo* design exercise. They designed one compound which was subsequently synthesized and determined to be a low micromolar PIN1 inhibitor. This compound served as starting point for a chemical optimization program which resulted in more potent analogs. One of them was crystallized with PIN1. Its crystal structure revealed a binding mode different from the initial model.

8. Expert opinion

The first automated *de novo* design techniques were developed about 20 years ago. These algorithms had the tendency to generate complex molecules lacking synthetic accessibility. Over the last couple of years, a few research groups have focused on the development of approaches that capture more synthetic chemistry knowledge into the *de novo* generation of candidate molecules. All of these use fragments as building blocks. Hence, computational *de novo* design approaches and FBS techniques are conceptually similar. Furthermore, an FBS campaign can produce unexpected binding modes which can serve as starting points for a *de novo* design exercise. The growing number of crystal structures of protein–ligand complexes provides another source of potential starting points for *de novo* design. So far this source has remained mostly underexploited. However, we expect the development of computational tools for large-scale analysis of binding pockets to enable the identification of common features between regions of binding pockets of unrelated proteins [31]. This would offer the opportunity to utilize a fragment binding to a pocket of a well-known protein as a starting point for the design of ligands for an unrelated target that shares some local similarity with the former. These techniques for large-scale analysis of binding sites could also help to design libraries of pocket-specific fragments.

Despite the development of FBS techniques, the field of structure-based *de novo* design has remained static over the last couple of years. The lack of a major breakthrough in the accurate prediction of binding free energies continues to hamper the development of structure-based *de novo* design.

In contrast, several ligand-based *de novo* design methods have been developed in the last few years. The development of both descriptors for similarity searching and methods to compute molecular shape has had an impact on ligand-based *de novo* design.

Over the last decade, another field, known as natural computing, has provided valuable approaches to tackle the sampling problem, that is, the efficient exploration of the chemical space to produce solution within a reasonable time frame. Most of the recently developed *de novo* design tools utilize natural computing techniques to sample the chemical space. We expect development in the field of natural computing to continue to influence *de novo* design.

Scoring is another key component of any *de novo* design software. Traditionally, *de novo* design tools have considered a single objective, mainly related to binding affinity to rank the proposed molecules. However, candidate molecules should meet other requirements including reasonable physico-chemical properties, novelty, synthetic accessibility and potential for further chemical modifications. Pareto-based methods offer a promising approach to identify solutions representing different compromises among different and sometimes conflicting properties. At this point, one should also mention that development of metrics for desirable properties such as novelty and potential for further chemical modifications would be helpful. Assessment of these desirable features relies mostly on manual scoring by experienced medicinal chemists.

In principle, one could conceive a *de novo* design algorithm as a combination of any of the available pieces of algorithms for building molecules, sampling the chemical space and scoring the solutions. Hence, there is a wide variety of combinations that could be tested. As far as we know, a large-scale evaluation of all the possible combinations of the different components of *de novo* design software is still missing. Most of the existing software packages have been validated retrospectively on limited data sets. Furthermore, criteria for a successful validation vary from one case study to the other.

Despite all these limitations, *de novo* design methods are a useful piece in the computational chemist's toolbox. They can either generate valuable starting points for further chemistry exploration activities or serve as an idea generator. However, the output of *de novo* design software should always be carefully and critically reviewed by a scientist to avoid computer-generated artifacts.

Declaration of interest

B Pirard is an employee of Novartis. He declares no conflict of interest and has received no payment in the preparation of this manuscript.

Bibliography

1. Perola E. An analysis of the binding efficiencies of drugs and their leads in successful drug discovery programs. J Med Chem 2010;53:2986-97
2. Walters WP, Namchuk M. Designing screens: how to make your hits a hit. Nat Rev Drug Discov 2003;2:259-66
3. Congreve M, Chessari G, Tisi G, et al. Recent developments in fragment-based drug discovery. J Med Chem 2008;51:3661-80
4. Bleicher KH, Bohm H-J, Muller K, et al. Hit and lead generation: beyond high throughput screening. Nat Rev Drug Discov 2003;2:369-78
5. Schneider G, Fechner U. Computer-based de novo design of drug-like molecules. Nat Rev Drug Discov 2005;4:649-63
6. Baber JC, Feher M. Predicting synthetic accessibility: application in drug discovery and development. Mini Rev Med Chem 2004;4:681-92
7. Loving K, Alberts I, Sherman W. Computational approaches for fragment-based and de novo design. Curr Top Med Chem 2010;10:14-32
8. Mauser H, Stahl M. Chemical fragment spaces for de novo design. J Chem Inf Model 2007;47:318-24
9. Degen J, Wegscheid-Gerlach C, Zaliani A, et al. On the art of compiling and using 'drug-like' chemical fragment spaces. ChemMedChem 2008;3:1503-7
10. Lewell XQ, Judd DB, Watson SP, et al. Recap retrosynthetic combinatorial analysis procedure, a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. J Chem Inf Comput Sci 1998;38:511-22
11. Patel H, Bodkin MJ, Chen B, et al. Knowledge-based approach to de novo design using reaction vectors. J Chem Inf Model 2009;49:1163-84
12. Hiss JA, Hartenfeller M, Schneider G. Concepts and applications of "natural computing" techniques in de novo drug and peptide design. Curr Pharm Des 2010;16:1656-65
13. Gohlke H, Klebe G. Approaches to the description and prediction of the binding affinity of small-molecules ligands to macromolecular receptors. Angew Chem Int Ed 2002;41:2644-76
14. Stahura FL, Bajorath J. New methodologies for ligand-based virtual screening. Curr Pharm Des 2005;11:1189-202
15. Erickson JA, Mader MM, Watson IA, et al. Structure-guided expansion of kinase fragment libraries driven by support vector machine models. Biochimica et Biophysica Acta 2010;1804:642-52
16. Gasteiger J, Ihlefnfeldt WD. Computer-assisted planning of organic syntheses: the second generation of programs. Angew Chem Int Ed Eng 1996;34:2613-33
17. Gillet VJ, Myatt G, Zsoldos Z, et al. SPROUT, HIPPo and CAESA: tools for de novo structure generation and estimation of synthetic accessibility. Perspect Drug Disc Design 1995;3:34-50
18. Ertl P, Schuffenhauer A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J Cheminformatics 2009;1:8

19. Dey F, Caflisch A. Fragment-based de novo design by multiobjective evolutionary optimization. J Chem Inf Model 2008;48:679-90
20. Nicolaou C, Apostolakis J, Pattichis CS. De novo drug design using multiobjective evolutionary graphs. J Chem Inf Model 2009;49:295-307
21. Mauser H, Guba W. Recent developments in de novo design and scaffold hopping. Curr Opin Drug Discov Devel 2008;11:365-74
22. Damewood Jr 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;50:1296-303
23. Zaliani A, Boda K, Seidel T, et al. Second-generation de novo design: a view from a medicinal chemist perspective. J Comput Aided Mol Des 2009;23:593-602
24. Alig L, Alsenz J, Andjelkovic M, et al. Benzodioxoles: novel cannabinoid-1 receptor inverse agonists for the treatment of obesity. J Med Chem 2008;51:2115-27
25. Proschak E, Sander K, Zettl H, et al. From molecular shape to potent bioactive agents II: fragment-based de novo design. ChemMedChem 2009;4:45-8
26. Lo HY, Bentzien J, White A, et al. 2-Aminobenzimidazoles as potent ITK antagonists: de novo design of a pyrrole system targeting additional hydrogen bonding interaction. Tetrahedron Lett 2008;49:7337-40
27. Li W-W, Chen J-J, Zeng R-L, et al. Taking quinazoline as a general support-nog to design potent and selective kinase inhibitors: application to FMS-like tyrosine kinase 3. ChemMedChem 2010;5:513-16
28. Park H, Bahn YJ, Ryu SE. Structure-based de novo design and biochemical evaluation of novel cdc25 phosphatase inhibitors. Bioorg Med Chem Lett 2009;19:4330-4
29. Guo C, Hou X, Dong L, et al. Structure-based design of novel human Pin1 inhibitors. Bioorg Med Chem Lett 2009;19:5613-16
30. Wang R, Gao Y, Lai L. LigBuilder: a multi-purpose program for structure-based drug design. J Mol Model 2000;6:498-516
31. Perot S, Sperandio O, Miteva MA, et al. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. Drug Discov Today 2010;15:656-67
32. Hartenfeller M, Proschak E, Schuller A, et al. Concept of combinatorial de novo design of drug-like molecules by particle swarm optimization. Chem Biol Drug Des 2008;72:16-26
33. Huang Q, Li L-N, Yang S-Y. PHDD: a new pharmacophore-based de novo design method of drug-like molecules combined with assessment of synthetic accessibility. J Mol Graphics Model 2010;28:775-87
34. Pfeffer P, Fober T, Hullermeier E, et al. GARLig: a fully automated tool for subset selection of large fragment spaces via a self-adaptative genetic algorithm. J Chem Inf Model 2010;50:1644-59

Affiliation

Bernard Pirard
Novartis Institute for BioMedical Research,
Global Discovery Chemistry,
Computer-Aided Drug Discovery,
Forum 1, CH-4002 Basel,
Switzerland
Tel: +41 61 32 45 620;
E-mail: bernard.pirard@novartis.com