

Fragment-Based *De Novo* Design of Cyclin-Dependent Kinase 2 Inhibitors

Sunil Kumar Tripathi, Poonam Singh, and Sanjeev Kumar Singh

Abstract

Cyclin-dependent kinases (CDKs) are core components of the cell cycle machinery that govern the transition between phases during cell cycle progression. Abnormalities in CDKs activity and regulation are common features of cancer, making CDK family members attractive targets for the development of anticancer drugs. One of the main bottlenecks hampering the development of drugs for kinase is the difficulty to attain selectivity. A huge variety of small molecules have been reported as CDK inhibitors, as potential anticancer agents, but none of these has been approved for commercial use. Computer-based molecular design supports drug discovery by suggesting novel new chemotypes and compound modifications for lead candidate optimization. One of the methods known as *de novo* ligand design technique has emerged as a complementary approach to high-throughput screening. Several automated *de novo* software programs have been written, which automatically design novel structures to perfectly fit in known binding site. The *de novo* design supports drug discovery assignments by generating novel pharmaceutically active agents with desired properties in a cost as well as time efficient approach. This chapter describes procedure and an overview of computer-based molecular *de novo* design methods on a conceptual level with successful examples of CDKs inhibitors.

Key words Cyclin-dependent kinases, *De novo* design, Genetic algorithm, High-throughput screening, X-ray crystallography

1 Introduction

In drug designing process the search for drug molecules with computational methods is frequently performed by molecular docking or to a lesser extent by *de novo* drug design approaches. While virtual screening relies on pre-existing compounds, *de novo* design approaches generate novel molecules out of building blocks consisting of single atoms or fragments [1]. The *de novo* design involves the design of novel structures based on the structure of binding site with which they are meant to interact, which can be identified from an X-ray crystallographic study of the receptor protein with bound ligand. By definition, *de novo* means “from the new” and involves the design of complex new molecules from

smaller constituent parts. The creation of any new chemical entity could be considered *de novo* design, while in practice the *de novo* design will be defined as the creation of novel chemical entities that are substantially different from any of the starting materials [2].

De novo design involves the design of novel structures based on the structure of binding site with which they have to interact. A drug can be designed which will have the correct size and shape to fit the space available with required functional groups to interact in binding sites [3]. The operator can carry out each of these procedures manually. Automated *de novo* design of bioactive molecules is one of the key aims in computational chemistry. To approach this idea, algorithms have to provide two principal tasks: First, the search space, i.e., the set of all algorithmically tractable molecules must be structured into regions of higher and lower quality to allow for prediction of desired properties (protein binding), and second, systematic search to make easy navigation in a high-dimensional chemical space [4].

One of the principal attractions of designing drug-like molecules starting from simpler building blocks is the potential to cover a larger fraction of the available compounds within the size range of interest. A relatively small number of fragments can be compiled to cover most of the possible shapes, features, and properties contained in a much larger set of drug-like compounds. This is in contrast to the very small amount of the accessible space that can be covered in a screening library for molecules of drug-like size [2, 5, 6]. Conceptually, Schneider and Fechner focused on three basic questions, which must be addressed by a *de novo* design program: (1) how to assemble the candidate compounds; (2) how to evaluate their potential quality; (3) and how to sample the search space effectively. All algorithmic decisions of a *de novo* design program are noticeably appraised by the quality of their outcome, which in turn crucially depends on a meaningful reduction of the search space [7].

In this chapter, we describe strategies and methods used in computer-based molecular *de novo* design methods on a conceptual level. The application of this approach is exemplified with a detailed description of the design of CDK2 and CDK4 inhibitors obtained by fragment-based *de novo* design program to identify promising scaffold candidates.

2 Materials

1. GANDI (Genetic Algorithm-based *de Novo* Design of Inhibitors).
2. Protein Data Bank (PDB) for protein structures.
3. Molinspiration Cheminformatics (www.molinspiration.com) to construct molecules and library of fragments design.

- CORINA and Babel for conversion from SMILES strings to MOL2 format.
- SEED software to dock fragment library into the receptor binding site.
- LEGEND *de novo* design program.
- MODELLER for homology or comparative modeling of protein three-dimensional structures.
- CHARMm force field for model minimization.

3 Methods

The methods described here allow to the top ranking molecules to be involved in one to three hydrogen bonds with the backbone polar groups in the hinge region of CDK2, an interaction pattern observed in potent kinase inhibitors [1]. Crystallography has revealed that the ATP-binding site of CDK2 can accommodate a number of diverse molecular frameworks, exploiting various sites of interaction. In addition, residues outside the main ATP-binding cleft have been identified that could be targeted to increase specificity and potency [8]. For simplicity of reference the ATP-binding site is generally described by means of the ATP-binding mode as: The adenine pocket, the ribose pocket, and the phosphate groove. Figure 1 shows a schematic of the naming scheme incorporating two additional regions. An important feature in terms of drug design is that ATP does not occupy the total volume of the cleft and there are nonconserved regions that can be exploited in the development of inhibitors [9–11].

The search for drug molecules with computational methods is often performed by high-throughput docking or to a lesser extent by *de novo* drug design approaches [7]. The *de novo* procedures can also be done manually [4]; however this procedure would not be a good idea (*see* Notes 1 and 2). The *de novo* design approaches

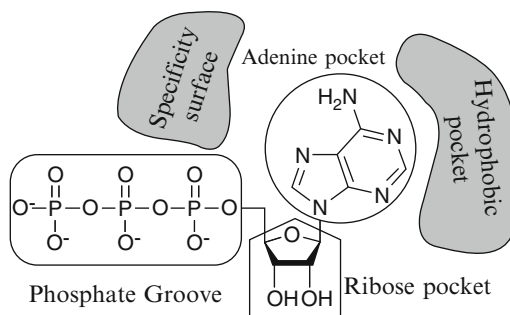


Fig. 1 Schematic presentation of ATP-binding site categorized according to ATP-binding mode in CDK2 with additional nonconserved regions

generate novel molecules out of building blocks consisting of single atoms or fragments. Because of the difficulty to predict synthetic accessibility, de novo drug design tools often generate molecules that demand synthesis [12]. Other important points could be also taken into consideration in de novo design (*see* **Notes 3–7**). Alternatively, there are many software packages available which will carry out the process automatically; some of de novo design programs are given in Table 1.

Table 1
Different de novo design programs with their scoring function

| De novo design programs | Scoring function | References |
|-------------------------|--|------------|
| GANDI | Force field | [1] |
| HSITE/2D Skeletons | Steric constraints and hydrogen bonds | [13] |
| Diamond Lattice | Steric constraints and hydrogen bonds | [14] |
| 3D skeletons | Steric constraints and hydrogen bonds | [15] |
| MCDNLG | Potential energy | [16] |
| GrowMol | Simple empirical scoring function | [17] |
| MCSS & HOOK | Simplified van der Waals potential of nonpolar interactions | [18] |
| SPROUT | Solvent accessible surface, hydrogen bonds, electrostatic and hydrophobic interactions | [19] |
| GroupBuild | Force field | [20] |
| CONCEPTS | Empirical scoring function | [21] |
| GenStar | Steric constraints and ligand-protein contact | [22] |
| SPLICE | Pharmacophore based and steric constraints | [23] |
| NEWLEAD | Steric constraints | [24] |
| LUDI | Empirical scoring function | [25, 26] |
| LEGEND | Force field | [27] |
| Chemical Genesis | Combined score of shape, grid based and scalar constraints | [28] |
| PRO_LIGAND | Empirical scoring function | [29] |
| DLD | Potential-energy function without electrostatic interactions | [30] |
| SMoG | Knowledge-based scoring function | [31] |
| BUILDER | Steric constraints | [32] |
| CONCERTS | Force field | [33] |
| RASSE | Force field | [34] |

(continued)

Table 1
(continued)

| De novo design programs | Scoring function | References |
|-------------------------|---|------------|
| PRO_SELECT | Empirical scoring function | [35] |
| SkelGen | Geometric, connectivity, and chemical constraints | [36] |
| Nachbar | Target-specific QSAR model based on topological connectivity descriptor | [37] |
| Globus | Molecular similarity based on all-atom-pairs-shortest-path descriptor | [38] |
| DycoBlock | Force field and solvent accessible surface | [39] |
| LEA | Target-specific QSAR model based on three-dimensional descriptors | [40] |
| LigBuilder | Empirical scoring function | [41] |
| TOPAS | Molecular similarity based on topological pharmacophore and substructure fingerprints | [3] |
| F-DycoBlock | Force field and solvent accessible surface | [42] |
| ADAPT | Weighted sum of DOCK score, clogP, MM, number of rotatable bonds and hydrogen bonds | [43] |
| CoG | Molecular similarity based on fingerprint descriptor | [44] |
| BREED | – | [45] |
| CCLD | Force field | [46] |

3.1 *Fragment-Based De Novo Ligand Design Strategy of CDK2 Inhibitors*

In this section a novel approach for Genetic Algorithm-based *de novo* Design of Inhibitors (GANDI) is presented [1]. GANDI is a fragment-based method that generates molecules by joining pre-docked fragments with linkers with parallel genetic algorithm [47], employing the simultaneous evolution of multiple populations used in GANDI to search for feasible solutions. Only the pre-docked fragments are encoded by the genetic algorithm, while suitable linker fragments are efficiently evaluated with a tabu search [48–50]. GANDI is evaluated on CDK2, and it is able to suggest molecules with new scaffolds or substituents which preserve the main binding interaction motifs of known inhibitors of CDK2 [1].

3.1.1 *GANDI Procedure for Design of CDK2 Inhibitors*

In order to connect pre-docked fragments with linker fragments (*see* Fig. 2), GANDI uses a combination of two stochastic search procedures, namely (1) a genetic algorithm [51, 52] and a tabu search [48–50], which are as follows.

1. The genetic algorithm in GANDI is an island model, using the simultaneous, noninteracting evolution of multiple populations at the same time. The exchange of genetic material is

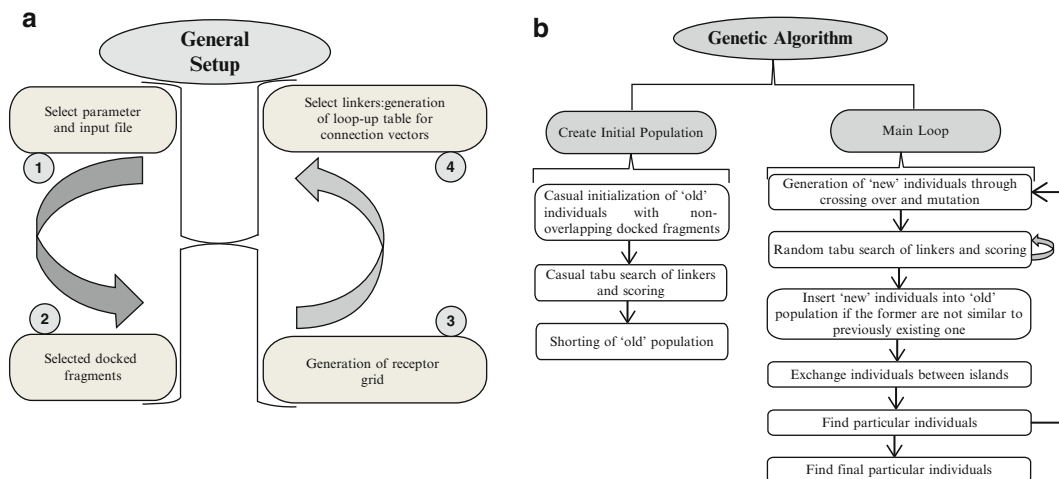


Fig. 2 Schematic illustration of GANDI with general procedure (a) from top to bottom including two iterative procedures, which are the main loop of the genetic algorithm and the random tabu search (b)

performed after a certain amount of optimization iterations by swapping individuals (i.e., molecules) between neighboring, all, or randomly picked islands as reported by Dey and Caffisch [1]. Every individual contains a single chromosome consisting of multiple genes. Contrary to classic genetic algorithms (*see Note 8*) [7, 53], the implementation in GANDI uses integers as gene values encoding indexes of docked fragments.

- The next step involves linking the encoded fragments for each individual by a tabu search [48–50]. GANDI builds a look-up table containing all distances and angles of all pairs of linker fragment connection vectors. All possible connections of fragment pairs of an individual are generated with cutoff values and the look-up table. A connection solution is randomly picked, and the two fragments are joined with the linker [1].

3.1.2 3D-Similarity and Selection of Fittest Individuals

The detection of 3D-similarity between molecules may be measured as reported by several researchers [1, 54–56]. The scoring function implemented in GANDI is a linear combination, and force field-based energy function consists of van der Waals and electrostatic terms. The potentials of the receptor are calculated and stored on a grid [56] and used only for the linkers. The next term in the scoring function is a fingerprint-based 2D measure of similarity (2D similarity) as Dey and Caffisch have used the crystallographic structure of CDK2 with an oxindole-based inhibitor [1, 57] from the protein data bank (PDB ID: 1KE5). The crystallographic ligand can be prepared as described for the protein but minimized without any constraints inside the rigid protein binding site.

3.1.3 Selection of Fragment Library and Docking with SEED

In this step, library of fragments from Molinspiration Cheminformatics (*see Note 9*) is the source of constructed molecules (www.molinspiration.com). The library consisted of 20,000 fragments with one and 20,000 fragments with two connection points occurring in bioactive molecules. Both sets are converted from SMILES strings to MOL2 format with CORINA [58, 59] and Babel [60]. Then, all are docked into the receptor binding site with SEED, a program for docking mainly rigid fragments with evaluation of protein-fragment energy and electrostatic desolvation [56] (*see Note 10*).

3.1.4 Quality Assessment of Binding Modes

The quality of the binding modes generated by GANDI is further assessed with an in-house developed suite of programs for flexible ligand docking. First GANDI molecules were decomposed into fragments by DAIM [61]. DAIM prioritizes also the resulting fragments according to their suitability as anchors for the docking program FFLD [62]. In the present case study applied to CDK2, 2D similarity is used only for analysis (and not during GANDI runs) by Tanimoto coefficient based on normalized DAIM fingerprints [1, 61].

3.2 Generation of Potent CDK4 Inhibitors with *De Novo* Design Program LEGEND

This section presents the strategy of *de novo* design program LEGEND with structure-based design of novel and potent CDK4 inhibitors. To obtain the CDK4 protein structure, homology model is build, based on monomeric form of CDK2. It has been implemented a new *de novo* design strategy which combines the *de novo* design program LEGEND with in-house structure selection supporting system SEEDS to generate new scaffold candidates [63].

3.2.1 Homology Model Building of CDK4 Protein

According to the level of sequence homology between CDK4 and CDK2 (45 %), the CDK4 homology model has been constructed based on the CDK2 Protein (*see Note 11*). The model is minimized with the CHARMM force field [64] with the exception of the conserved region in CDK4 and CDK2. This minimized structure is used for scaffold design in the *de novo* design strategy.

3.2.2 Identification of New Scaffolds with Structural Requirements

In CDK2, the NH group in Leu83 is the most important in binding of the inhibitors because it serves as a hydrogen-bonding donor in every structure reported in earlier studies [11, 65]. The main chain carbonyl groups of Glu81 and Leu83 also seemed to be important because most inhibitors form hydrogen bonds with them. In CDK4, most of the residues that are important for hydrogen bonds interactions are conserved. Among the altered amino acid residues, those between Leu83 in CDK2 and Val96 in CDK4 would not be critical since only the main chain is used for hydrogen bond(s). Therefore, this information concerning structural requirements is also useful to find out new class of CDK4

inhibitors and is used to identify new scaffolds that satisfy these structural requirements [63]. LEGEND program is based on the atom-by-atom approach and is suitable for generation of drug molecules in cavity of the ATP-binding site [66]. The output structures from primary step could not be commercially available (*see Note 12*). In the case of CDK4, a diarylurea scaffold has been identified as appropriate for rapid construction of structurally diverse libraries [63]. The data obtained from LEGEND de novo design suggested that neighboring NH and CO in the diarylurea scaffolds form hydrogen bond(s) with the main chain(s) of Glu94 and Val96, and aromatic rings are supposed to be located in the hydrophobic regions of CDK4. According to these insights, Honma and his group also provided further strategy for constructing diarylurea informer libraries based on the structural requirements of Cdk inhibitors in the ATP-binding pocket of the Cdk4 mode. Further, docking study is performed to investigate the binding mode of Diarylurea in CDK4. Moreover, design and synthesis can be also executed based on the predicted binding mode. Finally, X-ray analysis could be done for confirmation of binding mode from de novo design.

4 Notes

1. In general, the structure of the binding site is identified from X-ray crystallography and it is possible that the designed molecule may not bind to the binding site exactly as predicted. If the projected fit is too rigid, a minor modification in the binding mode may prevent the molecule binding. It would be better to have a flexible structure in the first instance [4].
2. It is significant to leave possibility for variation and elaboration in the molecule, and this agrees on fine modification of the molecule's binding affinity and also on required pharmacokinetics.
3. Flexible molecules are better than rigid molecules, because the former are more likely to find an alternative binding conformation. This allows modification to be carried out based on the actual binding mode. On other hand, if rigid molecules fail to bind as predicted, it may not bind at all.
4. It is pointless designing molecules that are difficult or impossible to synthesize.
5. Similarly, it is pointless designing molecules that need to adopt an unstable conformation in order to bind.
6. Consideration of the energy losses involved in water desolvation should be taken into account.

7. There may be subtle differences in structure between receptors and enzymes from different species.
8. Evolutionary algorithms are based on the ideas described by Charles Darwin, which are population-based optimization algorithms that mimic biological evolution with the genetic operators. Mutation introduces new information into a population, whereas recombination exploits the information inherent in the population [7].
9. Molinspiration offers a broad range of cheminformatics software tools supporting molecule manipulation and processing, including SMILES and SDfile conversion, normalization of molecules, generation of tautomers, molecule fragmentation, calculation of various molecular properties needed in QSAR, molecular modeling and drug design, high quality molecule depiction, molecular database tools supporting substructure and similarity searches.
10. It is significant to note that in the Dey and Caflisch studies, 6882 fragments with two connection vectors were used also as linker fragments in GANDI. They have pointed out the protocol of GANDI which integrated three different parameter settings with 400 individuals in a single island, 4 islands of 100 individuals each, and 4 islands with 100 individuals exchanging 5 % of all individuals every 20th iteration with a randomly selected island. The minimized oxindole-based inhibitor co-crystallized with the protein (PDB code 1KE5) was used as a target structure [1].
11. De Bondt et al. in 1993 have solved the first X-ray structure and analysis of CDK2. Although the structures of the CDK6-p16 complex and CDK6-p19 complex were already reported, their cyclin binding region and the ATP-binding pocket in CDK6 were both largely altered by the endogenous inhibitors (p16 and p19). Therefore, despite a very high sequence identity between CDK4 and CDK6 (70%), CDK6 structural information could not be used for designing CDK4 inhibitors [63].
12. This difficulty appears to apply to most *de novo* design programs that build up structures sequentially. To overcome this disadvantage, one can use an in-house program SEEDS [63].

Acknowledgments

The authors thank the Department of Science and Technology (SR/FT/CS-66/2010), New Delhi, for financial support. One of the authors SKT gratefully acknowledges CSIR, New Delhi, for SRF.

References

1. Dey F, Caffisch A (2008) Fragment-based de novo ligand design by multi objective evolutionary optimization. *J Chem Inf Model* 48: 679–690
2. Loving K, Alberts I, Sherman W (2010) Computational approaches for fragment-based and de novo design. *Curr Top Med Chem* 10:14–32
3. Schneider G, Lee ML, Stahl M et al (2000) De novo design of molecular architectures by evolutionary assembly of drug-derived building blocks. *J Comput Aided Mol Des* 14: 487–494
4. Patrick GL (2009) An introduction to medicinal chemistry. Oxford University Press, New York
5. Fink T, Bruggesser H, Reymond JL (2005) Virtual exploration of the small-molecule chemical universe below 160 Daltons. *Angew Chem Int Ed* 117:1528–1532
6. Martin YC (1981) A practitioner's perspective of the role of quantitative structure-activity analysis in medicinal chemistry. *J Med Chem* 24:229–237
7. Schneider G, Fechner U (2005) Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov* 4:649–663
8. Davies TG, Pratt DJ, Endicott JA et al (2002) Structure-based design of cyclin-dependent kinase inhibitors. *Pharmacol Ther* 93:125–133
9. Toledo LM, Lydon NB, Elbaum D (1999) The structure-based design of ATP-site directed protein kinase inhibitors. *Curr Med Chem* 6:775–805
10. Cherry M, Williams DH (2004) Recent kinase and kinase inhibitor X-ray structures: mechanisms of inhibition and selectivity insights. *Curr Med Chem* 11:663–673
11. Furet P (2003) X-ray crystallographic studies of CDK2, a basis for cyclin-dependent kinase inhibitor design in anti-cancer drug research. *Curr Med Chem Anticancer Agents* 3:15–23
12. Böhm HJ, Banner DW, Weber L (1999) Combinatorial docking and combinatorial chemistry: design of potent non-peptide thrombin inhibitors. *J Comput Aided Mol Des* 13:51–56
13. Danziger DJ, Dean PM (1989) Automated site-directed drug design: a general algorithm for knowledge acquisition about hydrogen-bonding regions at protein surfaces. *Proc R Soc Lond B* 236:101–113
14. Lewis RA (1990) Automated site-directed drug design: approaches to the formation of 3D molecular graphs. *J Comput Aided Mol Des* 4:205–210
15. Gillett VA, Johnson AP et al (1990) Automated structure design in 3D. *Tetrahedron Comput Methodol* 3:681–696
16. Gelhaar DK, Moerder KE, Zichi D et al (1995) *De novo* design of enzyme inhibitors by Monte Carlo ligand generation. *J Med Chem* 38: 466–472
17. Bohacek RS, McMartin C (1994) Multiple highly diverse structures complementary to enzyme binding sites: results of extensive application of a *de novo* design method incorporating combinatorial growth. *J Am Chem Soc* 116:5560–5571
18. Eisen MB, Wiley DC, Karplus M et al (1994) HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site. *Proteins* 19:199–221
19. Gillett VJ, Myatt G, Zsoldos Z et al (1995) SPROUT, HIPPO and CAESA: tools for *de novo* structure generation and estimation of synthetic accessibility. *Perspect Drug Discov Des* 3:34–50
20. Rotstein SH, Murcko MA (1993) GroupBuild: a fragment-based method for *de novo* drug design. *J Med Chem* 36:1700–1710
21. Pearlman DA, Murcko MA (1993) CONCEPTS: new dynamic algorithm for *de novo* design suggestion. *J Comput Chem* 14:1184–1193
22. Rotstein SH, Murcko MA (1993) GenStar: a method for *de novo* drug design. *J Comput Aided Mol Des* 7:23–43
23. Ho CMW, Marshall GR (1993) SPLICE: a program to assemble partial query solutions from three-dimensional database searches into novel ligands. *J Comput Aided Mol Des* 7:623–647
24. Tschinke V, Cohen NC (1993) The NEWLEAD program: a new method for the design of candidate structures from pharmacophoric hypothesis. *J Med Chem* 36:3863–3870
25. Böhm HJ (1992) The computer program LUDI: a new simple method for the de-novo design of enzyme inhibitors. *J Comput Aided Mol Des* 6:61–78
26. Böhm HJ (1992) LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J Comput Aided Mol Des* 6:593–606
27. Nishibata Y, Itai A (1991) Automatic creation of drug candidate structures based on receptor structure. Starting point for artificial lead generation. *Tetrahedron* 47:8985–8990
28. Glen RC, Payne AWR (1995) A genetic algorithm for the automated generation of molecules within constraints. *J Comput Aided Mol Des* 9:181–202

29. Clark DE, Frenkel D, Levy SA et al (1995) PRO-LIGAND: an approach to *de novo* molecular design. 1. Application to the design of organic molecules. *J Comput Aided Mol Des* 9:13–32
30. Miranker A, Karplus M (1995) An automated method for dynamic ligand design. *Proteins* 23:472–490
31. DeWitte RS, Shakhnovich EI (1996) SMOG *de novo* design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence. *J Am Chem Soc* 118:11733–11744
32. Roe DC, Kuntz ID (1995) BUILDERv.2: improving the chemistry of a *de novo* design strategy. *J Comput Aided Mol Des* 9: 269–282
33. Pearlman DA, Murcko MA (1996) CONCERTS: dynamic connection of fragments as an approach to *de novo* ligand design. *J Med Chem* 39:1651–1663
34. Luo Z, Wang R, Lai L (1996) RASSE: a new method for structure-based drug design. *J Chem Inf Comput Sci* 36:1187–1194
35. Murray CW, Clark DE, Auton TR et al (1997) PRO_SELECT: combining structure-based drug design and combinatorial chemistry for rapid lead discovery. 1. Technology. *J Comput Aided Mol Des* 11:193–207
36. Todorov NP, Dean PM (1997) Evaluation of a method for controlling molecular scaffold diversity in *de novo* ligand design. *J Comput Aided Mol Des* 11:175–192
37. Nachbar RB (2000) Molecular evolution: automated manipulation of hierarchical chemical topology and its application to average molecular structures. *Genet Program Evolvable Mach* 1:57–94
38. Globus A, Lawton J, Wipke WT (1999) Automatic molecular design using evolutionary algorithms. *Nanotechnology* 10:290–299
39. Liu H, Duan Z, Luo Q et al (1999) Structure-based ligand design by dynamically assembling molecular building blocks at binding site. *Proteins* 36:462–470
40. Douguet D, Thoreau E, Grassy G (2000) A genetic algorithm for the automated generation of small organic molecules: drug design using an evolutionary algorithm. *J Comput Aided Mol Des* 14:449–466
41. Wang R, Gao Y, Lai L (2000) LigBuilder: a multi-purpose program for structure-based drug design. *J Mol Model* 6:498–516
42. Zhu J, Fan H, Liu H, Shi Y (2001) Structure-based ligand design for flexible proteins: application of new F-DycoBlock. *J Comput Aided Mol Des* 15:979–996
43. Pegg SCH, Haresco JJ, Kuntz ID (2001) A genetic algorithm for structure-based *de novo* design. *J Comput Aided Mol Des* 15: 911–933
44. Brown N, McKay B, Gilardoni F, Gasteiger J (2004) A graph-based genetic algorithm and its application to the multiobjective evolution of median molecules. *J Chem Inf Comput Sci* 44:1079–1087
45. Pierce AC, Rao G, Bemis GW (2004) BREED: generating novel inhibitors through hybridization of known ligands application to CDK2, P38, and HIV protease. *J Med Chem* 47: 2768–2775
46. Caflich A (1996) Computational combinatorial ligand design: application to human R-thrombin. *J Comput Aided Mol Des* 10: 372–396
47. Venkatasubramanian V, Chan K, Caruthers JM (1994) Computer-aided molecular design using genetic algorithms. *Computers Chem Eng* 18:833–844
48. Glover F (1986) Future paths for integer programming and links to artificial intelligence. *Comput Oper Res* 13:533–549
49. Glover F (1989) Tabu Search-Part I. *ORSA J Comput* 1:190–206
50. Glover F (1990) Tabu Search-Part II. *ORSA J Comput* 1:4–32
51. Holland JH (1975) *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, MI
52. Goldberg DE (1989) *Genetic algorithms in search optimization and machine learning*. Addison-Wesley Longman Publishing, Boston
53. Darwin C (1859/1975) *On the origin of species: facsimile of the first edition*. Harvard University Press, Cambridge, MA
54. Cormen TH, Leiserson CE, Rivest RL et al (2001) *Introduction to algorithms*. MIT Press, Cambridge, pp 531–539
55. Kearsley S, Smith G (1990) An alternative method for the alignment of molecular structures. Maximizing electrostatic and steric overlap. *Tetrahedron Comput Methodol* 3:615–633
56. Majeux N, Scarsi M, Apostolakis J et al (1999) Exhaustive docking of molecular fragments on protein binding sites with electrostatic solvation. *Proteins Struct Funct Genet* 37:88–105
57. Bramson HN, Corona J, Davis ST et al (2001) Oxindole-based inhibitors of cyclin-dependent kinase 2 (CDK2): design, synthesis, enzymatic activities, and X-ray crystallographic analysis. *J Med Chem* 44:4339–4358
58. Momany FA, Rone R (1992) Validation of the general purpose QUANTA 3.2/CHARMm force field. *J Comput Chem* 13:888–900

59. Sadowski J, Gasteiger J (1993) From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. *Chem Rev* 93:2567–2581
60. Guha R, Howard M, Hutchison G et al (2006) The Blue Obelisk- interoperability in chemical informatics. *J Chem Inf Model* 46:991–998
61. Kolb P, Caffisch A (2006) Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. *J Med Chem* 49:7384–7392
62. Budin N, Majeux N, Caffisch A (2001) Fragment-based flexible ligand docking by evolutionary optimization. *Biol Chem* 382: 1365–1372
63. Honma T, Hayashi K, Aoyama T et al (2001) Structure-based generation of a new class of potent Cdk4 inhibitors: new de novo design strategy and library design. *J Med Chem* 44:4615–4627
64. Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamic calculation. *J Comput Chem* 4:187–217
65. Tripathi SK, Singh SK, Singh P et al (2012) Exploring the selectivity of a ligand complex with CDK2/CDK1: a molecular dynamics simulation approach. *J Mol Recognit* 25:504–512
66. Murcko MA (1997) An introduction to *de novo* ligand design. In: Practical application of computer-aided drug design. Marcel Dekker, New York, pp 322–325