JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

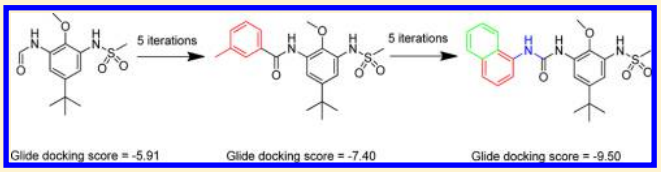
Evaluation of a Semi-Automated Workflow for Fragment Growing

Bernard Pirard* and Peter Ertl

Novartis Institutes for BioMedical Research, Novartis Campus, CH-4056 Basel, Switzerland

Supporting Information

ABSTRACT: Intelligent Automatic Design (IADE) is an expert system developed at Novartis to identify nonclassical bioisosteres. In addition to bioisostere searching, one could also use IADE to grow a fragment bound to a protein. Here we report an evaluation of IADE as a tool for fragment growing. Three examples from the literature served as test cases. In all three cases, IADE generated close analogues of the published compounds and reproduced their crystallographic binding modes. This exercise validated the use of the IADE system for fragment growing. We have also gained experience in optimizing the performance of IADE for this type of application.



1. INTRODUCTION

Fragment-based drug discovery (FBDD)[1,2] is a valuable hit-finding approach for well-characterized targets[3−6] as well as for emerging and more challenging ones.[7−10] Briefly, this approach seeks to identify small (MW ≤ 300) but efficient ligands using high-concentration screening and sensitive biophysical methods. Computational chemistry techniques assist the design of fragment libraries,[11,12] the selection of fragments for screening on a given target,[8,13,14] and the elaboration of fragment hits.[7,15] Fragment elaboration, which encompasses fragment morphing, growing and linking, can serve multiple purposes, depending on the outcome of a fragment screening campaign and on the aim of the project. Fragment morphing and analogue selection can help to explore the activity landscape around a fragment hit.[16−18] One can also optimize potency and physicochemical properties by fragment linking[7,19−21] and growing.[7,15,21,22]

Fragment morphing involves transformations such as heterocycle replacement, ring opening or closure, and modifications of substituents, to name a few. Two or several fragments that are observed to overlap in their binding to the protein target often serve as a source of inspiration for merging in a single molecule the structural features of overlapping binders. When a limited amount of structural data is available, computational methods for bioisostere identification and scaffold hopping[23−26] can generate valuable ideas for fragment morphing.

Fragment linking utilizes the concomitant binding of two fragments at two different but closely located subsites of a protein binding site. Using structural information, one can rationally design a linker connecting the two fragments without disrupting their interactions with the protein target. Current computational tools for fragment linking are based on the same concepts as scaffold replacement to construct and score molecules. Thompson and colleagues developed the CONFIRM algorithm to extract from a prepared library of linkers those matching search criteria derived from modes of fragment binding to a binding site,[27] and then hits of the database search are automatically connected to the bound fragments. The resulting molecules are docked in the binding site and assessed by docking scores and root-mean-square deviations from the known position of the starting fragments.

Dey and Caflisch combined a genetic algorithm and a tabu search to select poses of predocked fragments for linking with a user-supplied list of fragments.[28] They performed a retrospective validation study on cyclin-dependent kinase 2 (CDK2), considering a library of about 1400 fragments as building blocks and the binding mode of a known CDK2 inhibitor to bias the design. This study produced solutions forming the same hydrogen-bonding pattern as known CDK2 inhibitors. Furthermore, one of the top-scoring solutions shares the same substructure elements as a potent CDK2 inhibitor that was not present in the reference compound selected to bias the design.

In an ideal case, the binding affinity of a compound resulting from linking of two fragments should be greater than the sum of the binding energies of the two fragments, a phenomenon called superadditivity or positive cooperativity.[29] However, this is a rare event.

In contrast to fragment linking, more successful applications of fragment growing have been reported in the literature.[1,2,29] Visual analysis of crystal structures usually guides the fragment growing process. Docking of a virtual library built from commercially available building blocks helps to prioritize the selection of groups to add in the preferred direction(s) for growing. However, the outcome of these fragment growing approaches might depend on subjective decisions. Obviously, visual inspection of crystal structures is a subjective process. The scientist's training and experience also influence the selection of both a synthetic route and building blocks for fragment growing. Computer-based de novo design methods[30,31] provide a less biased alternative to the fragment growing approaches mentioned above. De novo design programs such as LUDI,[32−34] FlexNovo,[35] BIBuilder,[36] and LiGen,[37] to name a few, proceed by incremental addition of fragments to a seed fragment bound to a protein. Recently, Hoffer and colleagues described S4MPLE,[38]

Received: October 22, 2014
Published: December 16, 2014

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

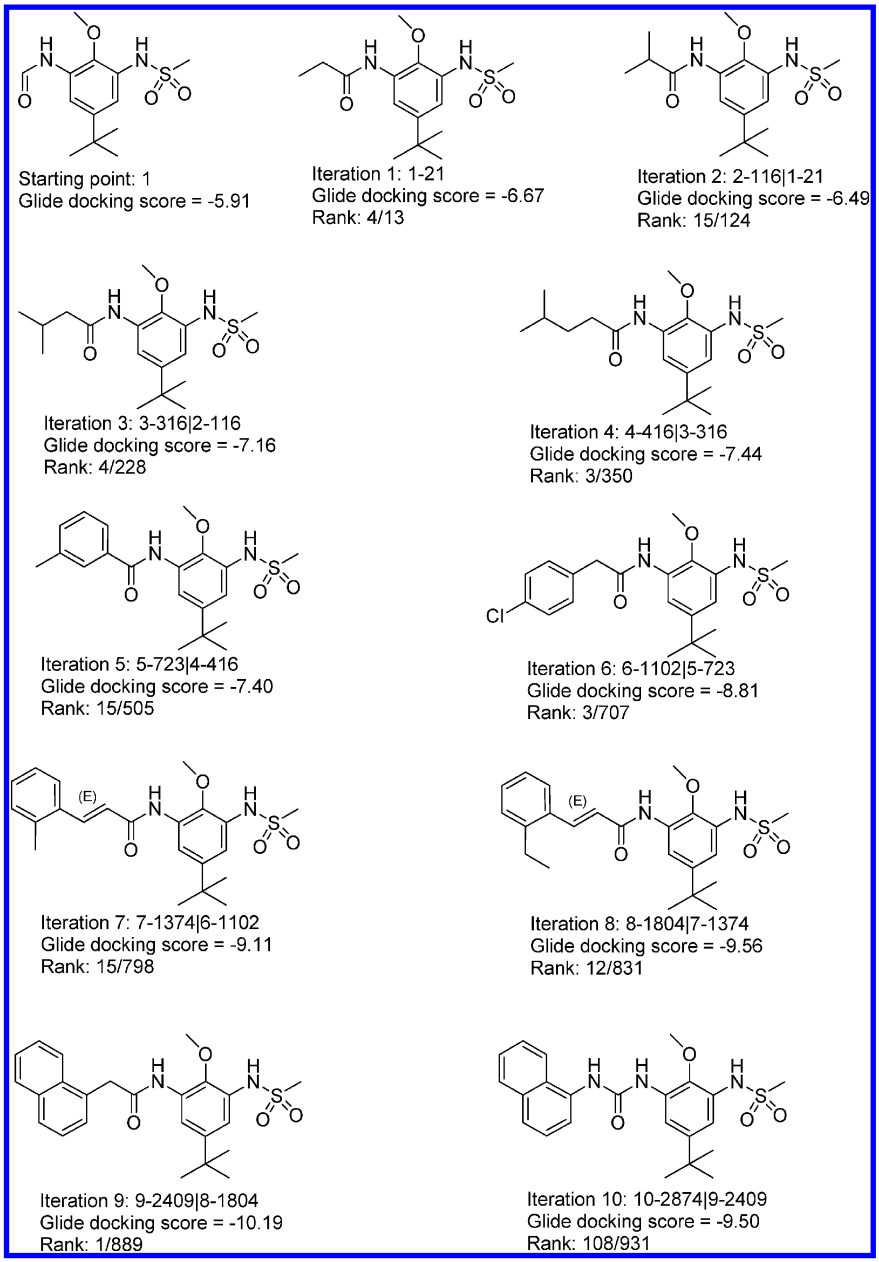


Figure 1. Example of fragment growing with IADE. Intermediate structures between the starting point **1** and the product 10-2874|9-2409 are shown. Each intermediate structure is characterized by an identifier, a Glide docking score, and a rank in the list of structures generated at a given iteration. The identifier consists of two parts. The first part refers to the iteration where this solution was generated and to a unique molecule identifier. The second part refers to the parent molecule. For instance, compound 10-2874|9-2409 was produced in the 10th iteration. Globally, it is the 2874th molecule and was generated from solution 9-2409. This identification system allows the user to find out which sequence of transformations were applied to generate a given molecule from the starting point. The molecules were ranked on the basis of their Glide scores.

a computational protocol for fragment growing or linking that uses a Lamarckian genetic algorithm for docking and conformational sampling. They carried out a validation study on four examples from the FBDD literature. In all cases, S4MPLE could automatically generate the expected ligands (and/or close analogues) from starting fragments, rank them among the top-scoring solutions, and reproduce their experimental binding modes.

Herein we explore the possibility of growing a fragment bound to a protein using Intelligent Automatic Design (IADE), an in-house iterative expert system that was initially developed to identify nonclassical bioisosteres.[25] In addition, slight modifications of the IADE settings could allow one to use it for fragment growing applications. To evaluate IADE for fragment growing, we selected three test cases from the FBDD literature: mitogen-activated protein kinase p38,[39] heat-shock protein 90 (HSP90),[4] and peroxisome proliferator-activated receptor γ (PPARγ).[40] In each case, the binding modes of the target protein in complex with both the starting fragment and at least one of the results from fragment growing are reported. Furthermore, growing the starting fragments produced compounds forming additional interactions with the target protein. This is good test case for IADE, as the Glide scoring function tends to reward solutions forming interactions with the binding site.

We considered the following criteria to assess the performance of IADE:

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

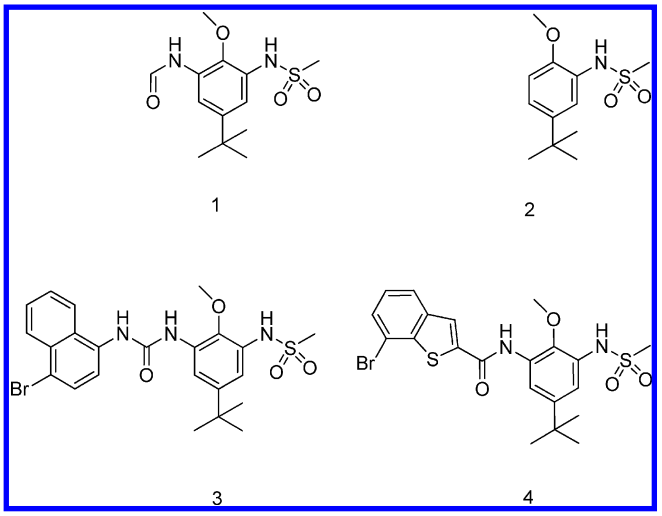


Figure 2. Molecular structures of starting points for fragment growing in the p38 binding site.

- ability to construct the final product (and/or its close analogues) of a fragment growing exercise starting from the initial fragments
- ability to generate molecules sharing substructure elements with known ligands of a target or target family
- ability to generate chemically sensible and relevant ideas

Table 1. Features of the Three Binding Sites Selected for the Evaluation of IADE

descriptor[a]	p38	HSP90	PPARγ
Size	489	223	374
Volume (Å^3)	1171	535	573
Enclosure	0.69	0.73	0.83
Exposure	0.57	0.54	0.33
Hydrophobic	0.81	1.02	1.36
Hydrophilic	0.94	0.94	0.99
Balance	0.86	1.08	1.37

[a]We refer to Methods for details on the SiteMap descriptors.

- ability to reproduce the crystallographic binding mode of the final product
- ability to generate molecules forming the same interactions with the target as known ligands.

These criteria are very similar to those defined by others to validate de novo design tools.[38,41]

2. METHODS

We describe below the computational protocol for proteins structures preparation and characterization, fragment growing with IADE and postprocessing of IADE solutions.

2.1. Protein Structure Preparation and Characterization. Atomic coordinates were extracted from Protein Data Bank (PDB) entries 3GI3 (p38), 3HZ1 (HSP90), and 3ET3

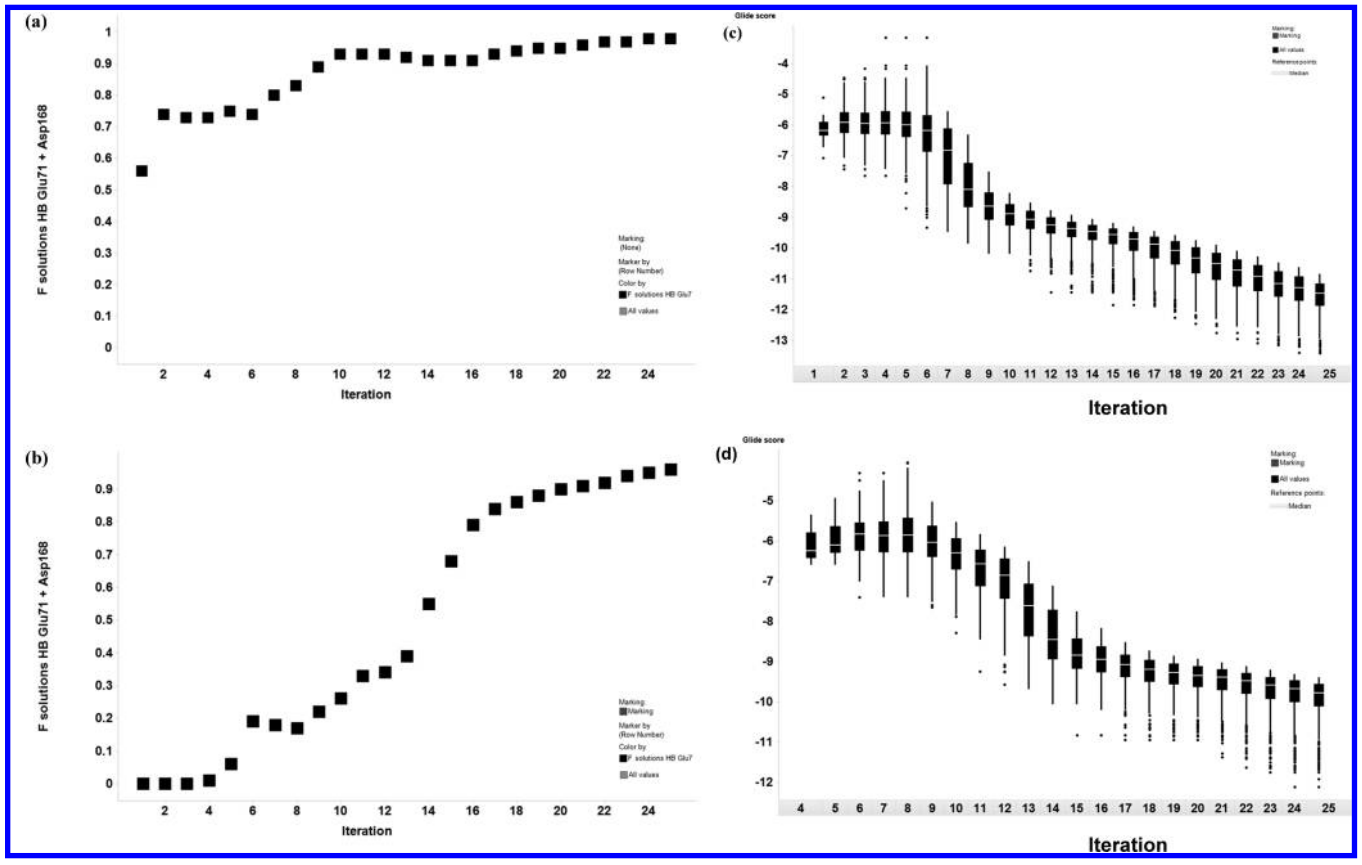


Figure 3. (a, b) Fractions of solutions forming a pair of H-bonds with the side chain of Glu71 and the backbone NH of Asp168 as functions of the number of iterations for growing runs starting from (a) fragment **1** and (b) fragment 2 (Figure 2). (c, d) Box plots of the Glide docking scores of solutions forming a pair of H-bonds with the side chain of Glu71 and the backbone NH of Asp168 as functions of the number of iterations for growing runs starting from (c) fragment **1** and (d) fragment **2** (Figure 2). For the growing starting from fragment **2**, IADE generated the first solutions forming the above-mentioned pair of H-bonds at iteration **4**.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

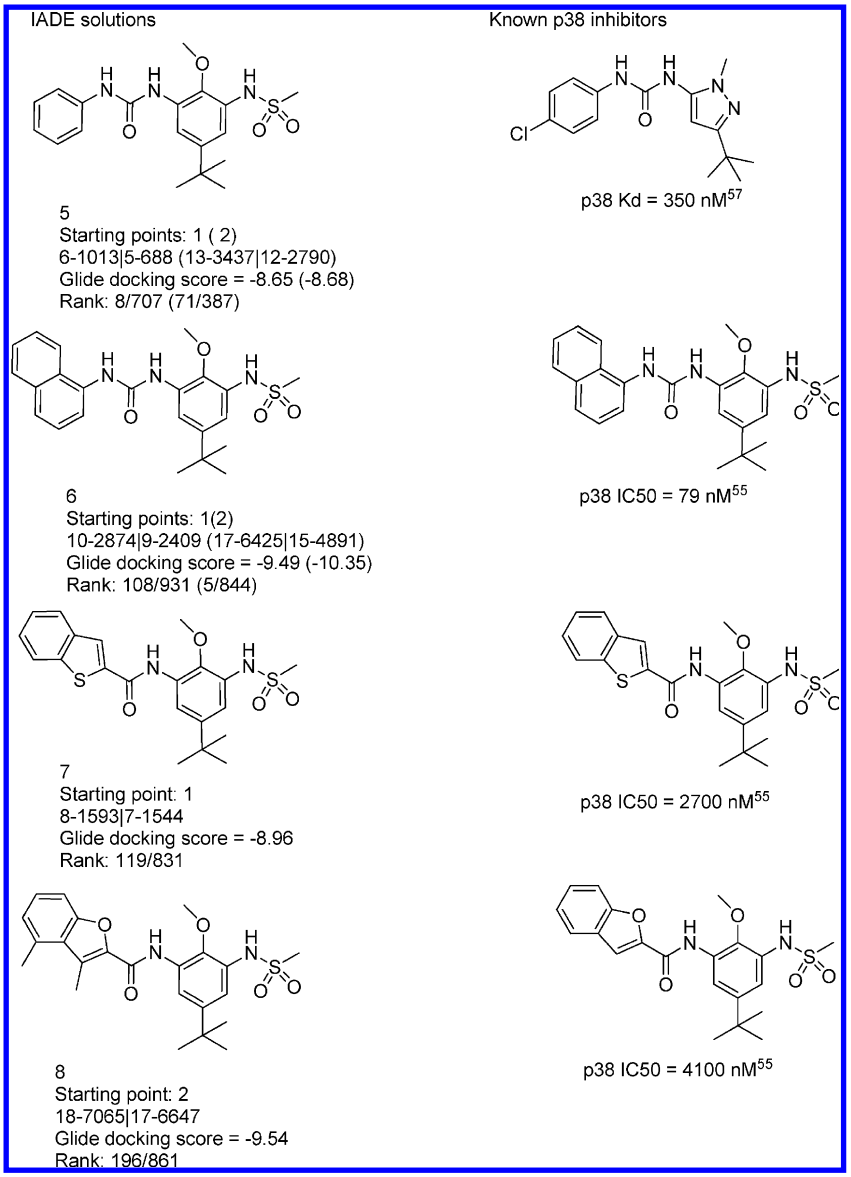


Figure 4. Molecular structures of selected IADE solutions and known p38 inhibitors. For each IADE solution, we specify the starting point (Figure 2), its identifier, the Glide docking score, and its rank in the list of structures generated at a given iteration. We refer to Figure 1 for an explanation of the naming scheme for the IADE solutions.

(PPARγ).[42] They served as input to Schrödinger's Protein Preparation Wizard module, considering default settings.[43] We removed all of the water molecules from 3GI3 and 3ET3, while after visual inspection we kept three of the conserved water molecules of the HSP90 binding site: HOH243, HOH327, and HOH333.[44] The structures produced by the Protein Preparation Wizard module were visually inspected to check the hydrogen-bonding patterns of polar side chains and conserved water molecules (HSP90) as well as the protonation states of side chains with ionizable groups. As a result, we did not make any manual modifications of the protein structures.

Schrödinger's SiteMap module was also run to characterize the binding sites of each protein structure.[45] Briefly, SiteMap proceeds in three stages: identification of binding sites, mapping of the sites, and calculation of descriptors. A site is defined by a set of points located on a grid. These points are subsequently filtered and grouped. Binding site mapping involves the calculation of van der Waals and electric field grids between a probe (water molecule) placed at each grid point and the protein. These calculations lead to the definition of the so-called phobic and philic potentials, which serve to partition the accessible space in each site into hydrophobic, hydrophilic, and neither hydrophobic nor hydrophilic regions. The hydrophilic regions are further subdivided into H-bond donor, H-bond acceptor, or metal binding areas. The output of the first two stages serves to evaluate several properties for each site. A weighted sum of some of the properties discussed below assesses the ability of a site to bind small organic molecules.

In each of the three test cases, SiteMap identified the site binding the fragments as the top-scoring one. Table 1 shows selected SiteMap descriptors for the top-scoring binding site of each target. Size is measured by the number of site points that make up the site. Volume for a protein site is well-defined when the site is fully enclosed by the protein. However, most protein sites are open to the solvent on one or more sides. In such a case, one needs to decide where to stop counting. A special procedure has been implemented in SiteMap to exclude regions that protrude too far into the solvent. Enclosure and Exposure are two

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

Table 2. p38 Test Case: Summary of Fragment Growing with IADE

	run 3	run 4	run 5
starting point[a]	**3**	**4**	**1**
no. with H-bond to NH of Met109[b]	106	200	501
no. with H-bond to NH of Met109 selected[c]	30	26	64
no. of hinge binding fragments[d]	20	23	10
no. of known hinge binding fragments[e]	8	6	2
no. of selected known hinge binding fragments[f]	15	11	17
median MW	570.2	567.9	517.7
median number of heavy atoms	40	39	35

[a]Fragments are shown in Figure 2. [b]Numbers of IADE solutions forming a hydrogen bond with the backbone NH of Met 109 (hinge). [c]Numbers of IADE solutions forming a hydrogen bond with the backbone NH of Met109 that were selected after visual inspection of the protein−ligand complex. [d]Numbers of different hinge binding fragments among the selected IADE solutions. [e]Numbers of known kinase hinge binders among the selected IADE solutions. The list of known kinase hinge binders was derived from ref 58. [f]Numbers of selected IADE solutions with a known kinase hinge binder.

different measures of how open a site is to solvent. Enclosure is the fraction of rays drawn from each site point that strike the protein surface within a distance of 10 Å, averaged over the site points. Higher Enclosure values are better, with an average of 0.78 for a tight binding site. To evaluate Exposure, one adds site points to the original ones. These additional site points must lie within a given distance from an original site point (default 3 Å) and form good contacts with the protein or lie within 4 Å of the nearest protein atom. Exposure is defined as the ratio of additional points to the total number of points. A shallow open site would allow the addition of many points, resulting in a high Exposure value. The average Exposure score is 0.49 for tight binding sites. Hydrophobic and Hydrophilic measure the relative hydrophobic and hydrophilic characters of a site, respectively. Balance is the ratio of these two properties. Hydrophobic and Hydrophilic have been calibrated to produce an average score of 1 for a tight binding site.

2.2. Fragment Growing. The IADE algorithm is described in detail in the original publication.[25] Therefore, we provide only a brief overview of the process. The iterative procedure starts from an X-ray structure of the ligand (template) bound to the target protein. The ligand is broken into fragments along all of the activated single nonring bonds. Every resulting fragment serves as a query for a database search aiming to identify replacements with similar properties (size, shape, pharmacophoric features, hydrophobicity, polarity,[46] and electronic properties[47]). After identification of a user-defined number of most similar potential replacements, molecules are constructed by combinatorially combining the identified fragments. The generated molecules are subsequently filtered to discard candidates with unwanted features, such as potential toxic substructures, nondruglike properties,[48] and low synthetic accessibility.[49] Then a set of representative conformations is generated for molecules passing these filters. These conformations serve as input for field-based alignment software that seeks the best overlay of the molecular interaction field extrema between each generated molecule and the template.[50] At this stage, one can use also the target protein to define an excluded volume. The best conformation for each molecule produced by the alignment software is further optimized in the protein binding site using the Glide docking score.[51] As a result, the generated molecules are ranked on the basis of their Glide docking scores. At the end of the whole process, the 10−20 best-ranking molecules are passed as starting structures to the next iteration. Program options allow the user to focus the search by defining substructure elements to be kept or substituents to be exclusively modified.

To support fragment growing, we made two modifications to the original IADE algorithm. In fragment growing applications, the top-scoring molecule becomes the template for the next iteration. This contrasts with searches for bioisosteres, which consider the same template for the whole process. This allows molecules to grow and accommodate to the protein binding cavity. In addition, the search procedure to identify bioisosteric replacements is modified to favor fragments slightly larger than the original one. According to our test, this modification considerably speeds the whole iterative procedure. Figure 1 illustrates the process of fragment growing with IADE.

2.3. Postprocessing of IADE Solutions. For each of the IADE solutions, we computed protein−ligand interaction

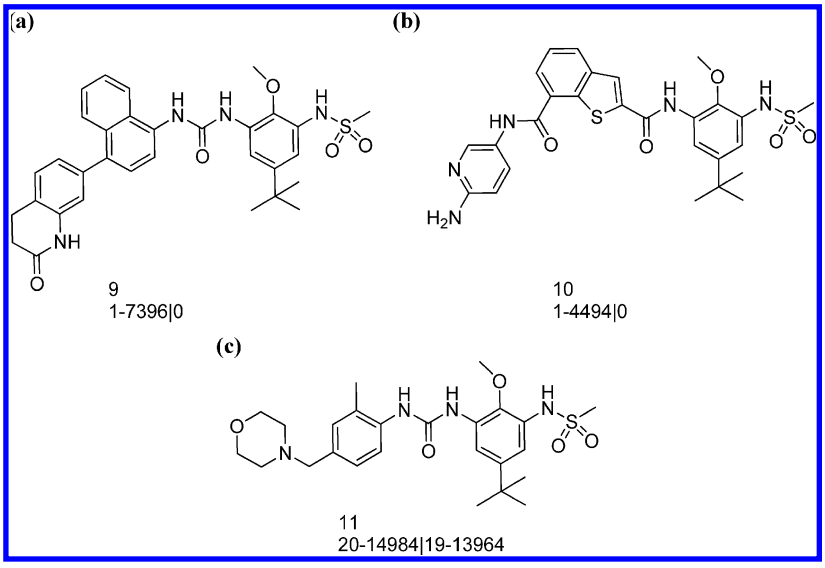


Figure 5. Molecular structures of representative IADE solutions generated by (a) run 3, (b) run 4, and (c) run 5 on p38. We refer to Figure 1 for an explanation of the naming scheme for the IADE solutions.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

fingerprints (PLIFs) using the Chemical Computing Group implementation with default settings.[52] PLIFs served to select solutions with defined interaction patterns. IADE solutions that did not match these interaction patterns were discarded at this stage. A list of the required interaction patterns for each target is available in Table S1 in the Supporting Information.

Solutions passing the PLIF filters were minimized in the binding site. Amino acids within 5 Å of the IADE solutions were allowed to move during the energy minimization. We computed the energy using Schrödinger's implementation of the molecular mechanics/generalized Born surface area (MM/GBSA) approach, known as the VSGB 2.0 energy model.[53]

Each minimized complex was visually inspected to discard solutions involved in nonideal intermolecular H-bonds or forming steric clashes with the protein. The other solutions were kept and considered as potential candidates for synthesis, as one would do in a drug discovery project.

3. RESULTS

In each of the three test cases, fragment growing took place in an enclosed binding site (Table 1). One could also use IADE to grow a fragment toward more solvent-exposed regions of a binding site. However, it may be more difficult to consider the evaluation criteria listed above for the assessment of IADE solutions binding to a solvent-exposed site. Two of the three test cases, p38 and HSP90, involved growing fragments bound to an allosteric pocket toward a region of the protein forming H-bonds with the ligands. In the third case, PPARγ, a fragment involved in H-bonds with the receptor served as the starting point for growing toward a hydrophobic region of the protein.

3.1. p38. The mitogen-activated protein kinase p38 plays a key role in the biosynthesis of some pro-inflammatory cytokines.[54] Goldberg and colleagues utilized a fragment linking algorithm to generate ideas for linkers between a *tert*-butylphenylsulfonamide fragment bound to an allosteric pocket created by a conformational change of the DFG motif and a pyridine involved in a hydrogen bond with the backbone NH of Met109 (hinge) of p38.[39] The same *tert*-butylphenylsulfonamide fragments (**1** and **2**; Figure 2) served as starting points for testing the ability of IADE to generate molecules that bind to the hinge of p38.[55] This enclosed binding site with mixed hydrophilic and hydrophobic properties (Table 1) is a nearly ideal test case for IADE. However, one needs to add fragments spanning a distance of about 15 Å to grow fragment **2** into a molecule that binds to the hinge.

We considered two different settings of IADE to grow fragments **1** and **2**. The first setting consisted of three steps: (1) growing of fragments **1** and **2** (runs 1 and 2, using 25 iterations of IADE with default parameters); (2) analysis of the results and selection of starting points for the next round of growing; and (3) growing of fragments **3** and **4** (Figure 2) (runs 3 and 4, using one iteration of IADE with modification of the Br atom only). This three-step approach mirrors a sequential fragment growing exercise involving multiple cycles of design, synthesis, and screening. On the other hand, the second setting (run 5) allowed larger modifications of fragment **1** to generate molecules filling the whole binding site after 25 iterations of a single IADE job.

Runs 1 and 2 did not produce any solution involved in a hydrogen bond with the backbone NH of Met109 (hinge). However, each iteration of runs 1 and 2 led to an increased fraction of solutions forming a pair of H-bonds with the side chain of Glu71 and the backbone NH of Asp168 (Figure 3A,B). This pair of H-bonds is a key feature of inhibitors binding to the

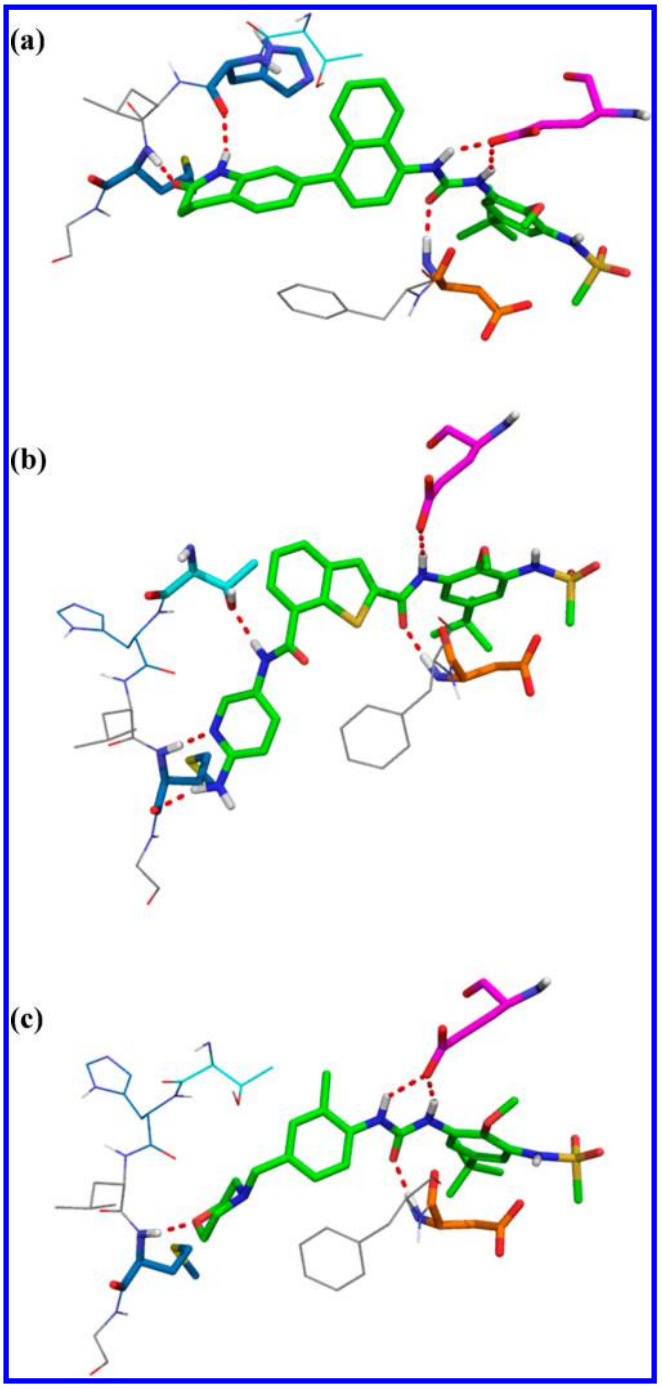


Figure 6. Representative IADE solutions docked in the p38 structure from PDB entry 3GI3. For clarity, only selected amino acids are shown. Amino acids involved in H-bonds are displayed as sticks, and each H-bond is represented as a red dotted line between the donor and acceptor atoms. We have colored in cyan the C atoms of the gatekeeper (Thr106), in blue the C atoms of the hinge amino acids able to form H-bonds with an inhibitor (His107 and Met109), in magenta the C atoms of Glu71 from helix C, and in orange the C atoms of Asp168 from the DFG motif. The three representative solutions donate at least one H-bond to the side chain of Glu71 and accept one H-bond from the backbone of Asp168. However, they form different H-bond patterns with the hinge. (a) Compound **9** (run 3) accepts a hydrogen bond from the backbone of His107 and donates a hydrogen bond to the backbone of Met109; (b) compound **10** (run 4) forms a pair of H-bonds with the backbone Met109 and also donates a hydrogen bond to the side chain of the gatekeeper; (c) compound **11** accepts a single H-bond from the backbone of Met109.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

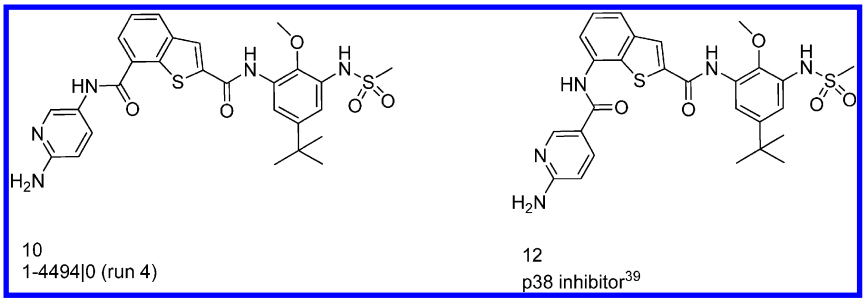


Figure 7. Molecular structures of compound **10** (IADE solution 1-4494|0, run 4) and compound **12**, one of the nicotinamide p38 inhibitors from ref 39. We refer to Figure 1 for an explanation of the naming scheme for the IADE solutions.

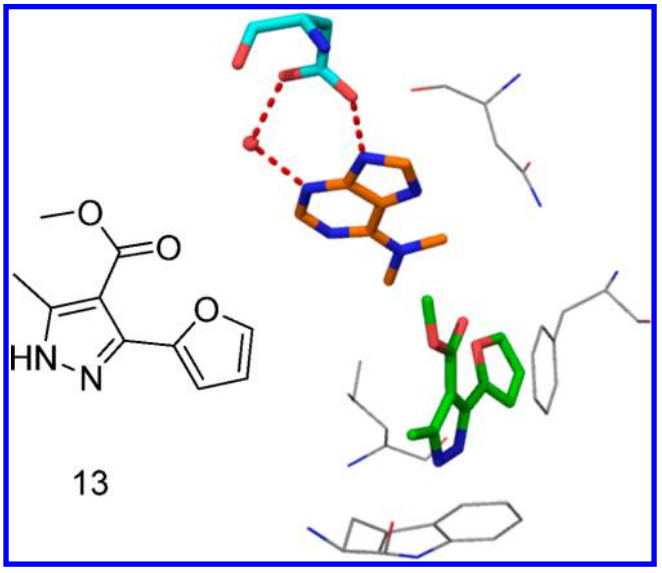


Figure 8. Molecular structure of second site binder fragment **13** and crystal structure of **13** in complex with the N-terminal domain of HSP90 (PDB entry 3HZ1). For clarity, only selected amino acids are shown. Amino acids involved in H-bonds are displayed as sticks; each H-bond is represented as a red dotted line between the donor and acceptor atoms. Fragment **13** binds to the second site and does not form any H-bonds. A purinergic fragment occupies the adenine site and forms H-bonds with both Asp93 and a conserved water molecule.

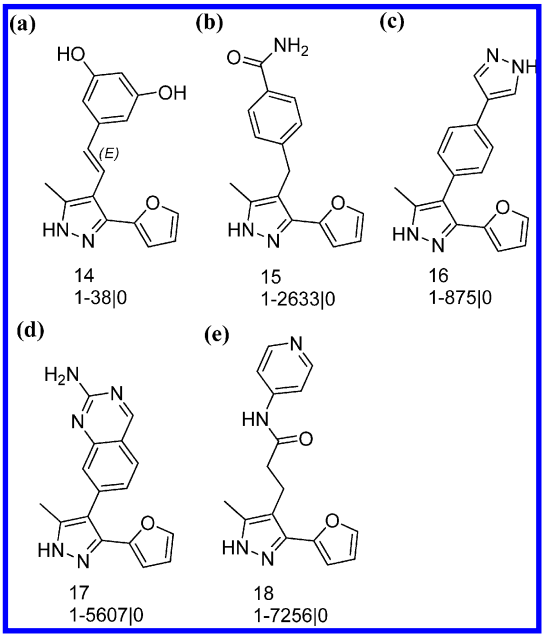


Figure 9. Molecular structures of representatives from the four classes of IADE solutions for HSP90: (a) phenol and isosteres, (b, c) benzamide and isosteres, (d) aminopyrimidine, and (e) others. We refer to Figure 1 for an explanation of the naming scheme for the IADE solutions. Molecular structures for all of the selected solutions are available in Figures S8−S12 in the Supporting Information.

DFG-out conformation of kinases.[56] In the meantime, the Glide docking scores of the generated molecules improved (Figure 3C,D). Furthermore, several of the IADE solutions were similar or even identical to known p38 inhibitors (Figure 4).[55,57] The nature of the starting point did not seem to influence the outcome of IADE (Figure 4). Depending on the starting point, six to 18 iterations out of a maximum of 25 were necessary to generate the molecules shown in Figure 4. Subsequent iterations resulted in the generation of close analogues of these molecules (Figure S1 in the Supporting Information).

Run 3 produced 106 solutions out of 1000 donating a hydrogen bond to the backbone NH of Met109 (hinge) (Table 2). After visual inspection of these 106 docking poses, we selected 30 compounds as potential candidates for follow-up (Figure S2 in the Supporting Information). Less than half of the 20 different hinge binders found in these compounds occur in known kinase inhibitors (Table 2, Figures 5a and 6a, and Figure S5 in the Supporting Information), while the remaining 11 are potentially novel hinge binding motifs.[58]

Fragment **4** (Figure 2) served as starting point for run 4, considering the same settings as for run 3. After visual inspection, we selected 26 (Table 2 and Figure S3 in the Supporting Information) of the 200 solutions involved in a hydrogen bond with the backbone NH of Met109. However, these selected solutions have more diverse hinge binding fragments compared with those of run 3 (23 vs 20 for run 3; Table 2). A quarter of these 23 fragments correspond to known hinge binders (Table 2, Figures 5b and 6b, and Figure S6 in the Supporting Information), while the other 17 are potentially novel hinge binders. Furthermore, solution 1-4494|0 is very similar to one of the most active compounds described by Goldberg and colleagues (Figure 7).[39] However, IADE failed to generate molecules with a nicotinamide substructure, as in those described in ref 39. At this point, we should also mention that we considered PDB entry 3GI3[55] as the target for this fragment growing exercise instead of one of the structures from ref 39, as the latter are not available in the PDB. Selecting a different p38 structure might lead to different IADE solutions. One could also consider the IADE solutions as a source of ideas for analogues. In this context, one would come up with the nicotinamide as replacement of the diaminopyridine generated by IADE (Figure 7).

Run 5 involved a simpler starting point (**1**; Figure 2) and allowed larger structural modifications at each iteration.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

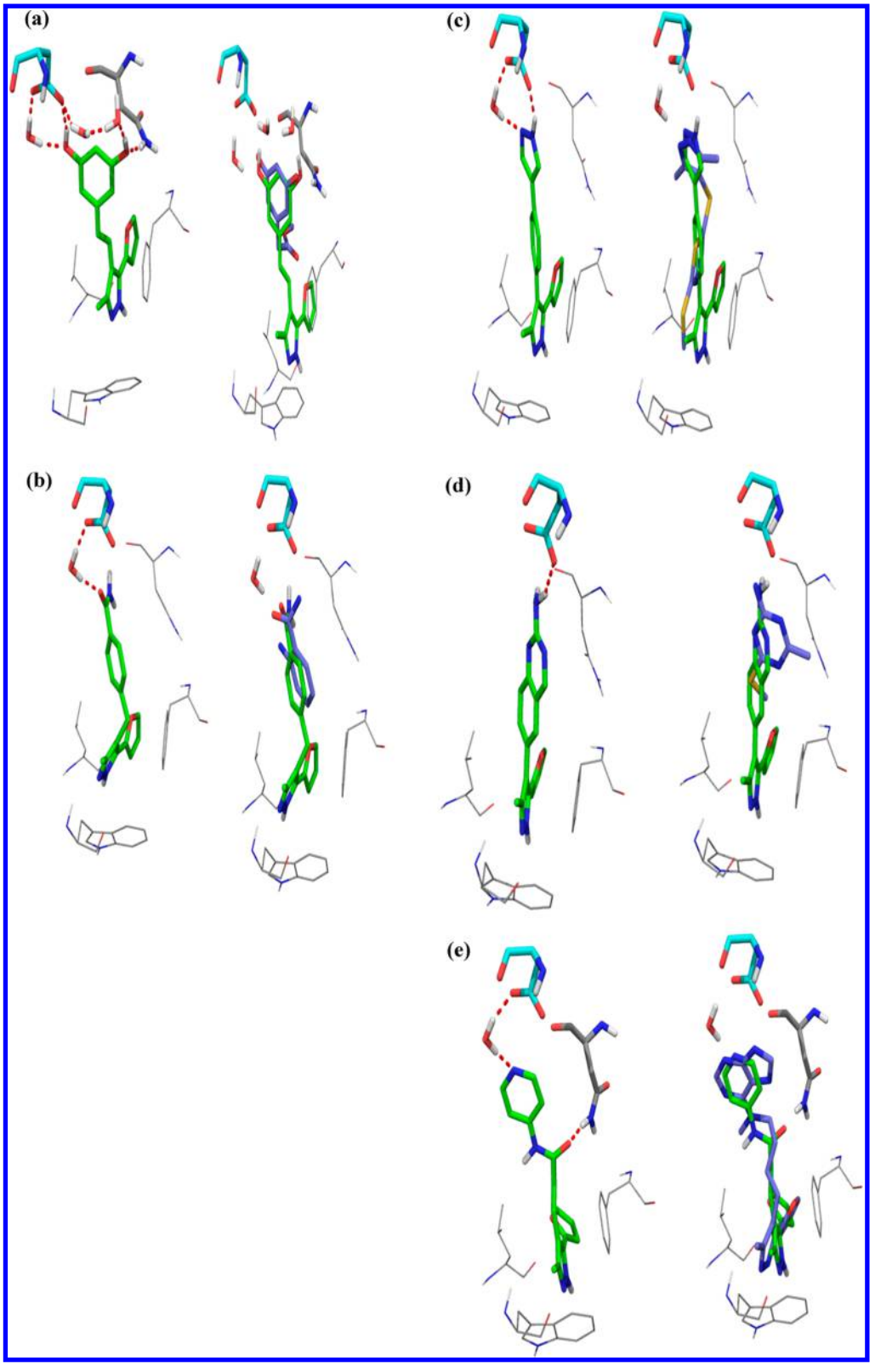


Figure 10. Representative IADE solutions for the four compound classes docked in HSP90. In each panel, the docking pose is presented on the left and an overlay of the representative IADE solution with the crystal structure of a related HSP90 inhibitor (Figure S13 in the Supporting Information) is shown on the right. For clarity, only selected amino acids and water molecules involved in H-bonds with the IADE solutions are displayed. Amino acids involved in H-bonds with the IADE solutions are displayed as sticks, and each H-bond is represented as a red dotted line between the donor and acceptor atoms. We have colored in cyan the C atoms of Asp103. (a) Compound **14** (phenol) overlaid with the fragment from PDB entry 2YE2; (b) compound **15** (benzamide) overlaid with the fragment from PDB entry 2YE6; (c) compound **16** (benzamide isostere) overlaid with the fragment from PDB entry 2YE7; (d) compound **17** (aminopyrimidine) overlaid with the fragment from PDB entry 2WI3; (e) compound **18** overlaid with the inhibitor from PDB entry 3HZ5.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

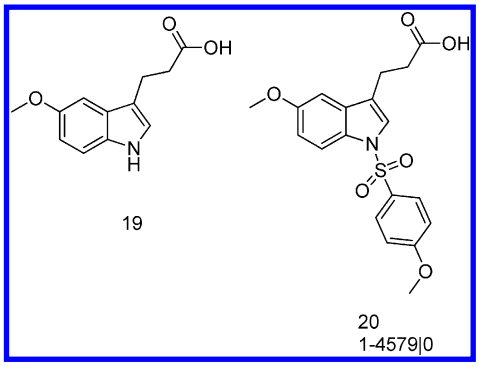


Figure 11. Molecular structures of two pan-PPAR modulators, fragment hit **19** and Indeglitazar (**20**) obtained by fragment growing.

This setting produced more molecules forming a hydrogen bond with the backbone NH of Met109 (Table 2, Figure S4 in the Supporting Information, and Figures 5c and 6c). However, their hinge binding fragment is more conserved and occurs less frequently in known kinase inhibitors compared with the hinge binders of solutions from runs 3 and 4 (Table 2 and Figure S7 in the Supporting Information).

Interestingly, solutions from run 5 tend to exhibit lower MW and fewer heavy atoms than the products of runs 3 and 4 (Table 2 and Tables S2−S4 in the Supporting Information). The selection of **3** and **4** (Figure 2) as starting points biased the property distributions of solutions from runs 3 and 4.

Molecular structures for all of the selected solutions from runs 3 to 5 are available in Figures S2−S4 in the Supporting Information.

3.2. HSP90. The ATP-dependent molecular chaperone HSP90 plays a role in the conformational stability, maturation, and function of several client proteins involved in cancer progression.[59,60] HSP90 belongs to the GHKL family of ATPases and consists of three domains: an N-terminal domain with an ATP binding site characterized by a left-handed $\beta-\alpha-\beta$ (Bergerat) fold that is topologically remote from the kinase ATP binding site, a central catalytic domain harboring the hydrophobic cleft for recruitment of the client protein, and a C-terminal domain involved in dimerization and autophosphorylation of the protein.[61] Despite the multidomain structure of HSP90 and the complex chaperone cycle, the structure of the N-terminal domain has been sufficient to guide the design of inhibitors.[4] Some of these inhibitors have entered clinical trials.[60,62]

Several drug discovery campaigns have utilized fragment-based approaches to identify starting points for lead optimization of HSP90 inhibitors.[4] Most of the FBDD efforts have led to the identification of fragments that mimic the binding of the adenine of ADP lying in a deep pocket, where it forms an extensive network of H-bonds with Asp93 and four conserved water molecules.[4] Researchers have also discovered fragments binding to a lipophilic pocket remote from Asp93.[19,63,64] This pocket is formed by a substantial rearrangement of residues Ile110-Gly114, which become the central portion of a long helix. Crystal structures were obtained for fragments bound to the induced lipophilic pocket either on their own or in the presence of another fragment bound to the adenine site. Attempts to link fragments bound to these two sites have resulted in potent HSP90 inhibitors.[19,64]

The crystal structure of fragment **13** bound to the induced lipophilic pocket of HSP90 (Figure 8) served as a starting point to test the ability of IADE to generate molecules forming H-bonds with the side chain of Asp93 and the array of conserved water molecules. We decided to grow **13** from position 4 of the pyrazole (the C atom bearing the methyl ester; Figure 8). Only modifications of the ester were allowed. A single IADE run generated 1000 solutions, 210 of them interacting with either the side chain of Asp93 or the array of conserved water molecules. After visual inspection of the docking poses, we selected 54 molecules as potential candidates for follow-up.

Analysis of the groups binding to the adenine site led to the grouping of these 54 compounds into four classes: phenol and isosteres (17 compounds; Figure S8 in the Supporting Information), benzamides and isosteres (nine compounds; Figures S9 and S10 in the Supporting Information), amino-pyrimidine (one compound; Figure S11 in the Supporting Information), and others (27 compounds; Figure S12 in the Supporting Information). The first three classes correspond to known chemical classes of HSP90 inhibitors. In other words, half of the selected compounds show adenine binding groups similar to those of known HSP90 inhibitors. Figure 9 shows the molecular structures of representative IADE solutions from each class. Docking poses and overlays with crystal structures of related HSP90 inhibitors are displayed in Figure 10. The representative of the phenol class, **14** (Figure 10a), forms multiple H-bonds with the side chains of both Asn51 and Asp93 as well as with two of the conserved water molecules. Its 3-hydroxyphenol is almost perfectly overlaid with the resorcinol-containing fragment of PDB entry 2YE2.[4] On the other hand, benzamide **15** and the fragment from PDB entry 2YE6 form slightly different H-bonding patterns (Figure 10b).[4] The benzamide O of both complexes accepts a hydrogen bond from one of the conserved water molecules. The NH_2 of **15** does not form any H-bonds with HSP90, whereas in PDB entry 2YE6 the benzamide NH_2 is

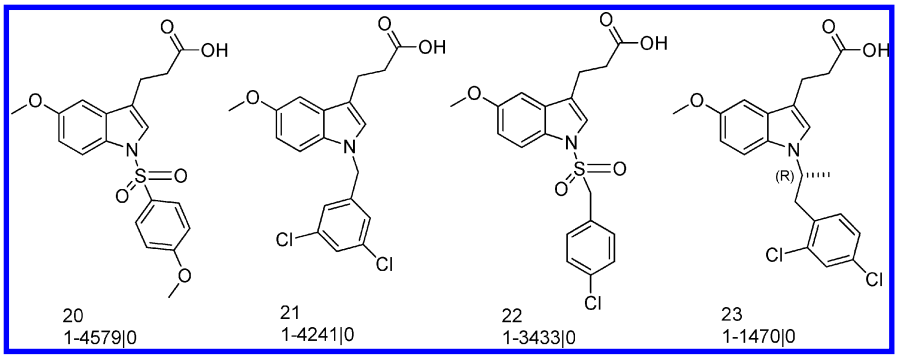


Figure 12. Molecular structures of representatives from the four classes of IADE solutions for PPARγ. We refer to Figure 1 for an explanation of the naming scheme for the IADE solutions. Molecular structures for all of the selected solutions are available in Figure 14 in the Supporting Information.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

involved in a direct H-bond with the side chain of Asp93. In this case, IADE failed to reproduce the fine details of the H-bond pattern. The benzamide isostere **16** forms the same H-bonds with the side chain of Asp93 and one of the conserved water molecules as the pyrazole-containing fragment from PDB entry 2YE7 (Figure 10c).[4] IADE produced only one solution with an aminopyrimidine binding to the adenine site (**17**; Figure 10d). Its NH_2 group donates a hydrogen bond to the side chain of Asp93, while none of its aromatic N accepts a hydrogen bond from any of the conserved water molecules. The aminopyrimidine of **17** shows a suboptimal overlay with the aminopyrimidine-containing fragment from PDB entry 2WI3 (Figure 10d).[65] Although IADE could generate solutions with known adenine site binders, it failed to reproduce the fine details of the interactions for some of these solutions. The largest group from the fourth class of IADE solutions (five out of 27) have a pyridine as the adenine site binder. The pyridine N of **18** (Figure 10e) accepts a hydrogen bond from one of the conserved water molecules, while its neighboring CH donates a nonclassical CH···O bond to the side chain of Asp93. Furthermore, the pyridine of **18** overlays well with the purine of the inhibitors from PDB entry 3HZ5 (Figure 10e).[19]

A single IADE run produced molecules mimicking the interaction of known HSP90 inhibitors with Asp93 and the conserved array of water molecules. The 54 selected solutions exhibit lead-like properties (Table S5 in the Supporting Information). They would be a valuable starting point for a further round of ideas for generation and design.

3.3. Peroxisome Proliferator-Activated Receptor γ. PPARγ, one of the three PPAR subtypes, regulates storage of fatty acids in adipocytes. It is the target of the glitazone class of insulin-sensitizing drugs. The PPARs belong to the thyroid hormone-like subfamily of nuclear receptors, which consist of six structural and functional regions. One of these harbors the ligand binding domain. This domain exhibits a fold comprising 12 α-helices arranged in three layers, forming an antiparallel "helical sandwich". Nearly half (five) of the 12 α-helices form the walls of a large and mainly hydrophobic cavity. This cavity has been the target of numerous drug discovery efforts aiming to identify PPAR modulators with different selectivity profiles.[66]

Artis and colleagues carried out an activity-based assay at high compound concentration to discover modulators of the three PPARS subtypes from a diverse fragment library.[40] As a result, they identified 5-methoxyindole-3-propionic acid (**19**) (Figure 11), a weak pan-PPAR agonist. Crystallographic characterization of **19** in complex with PPARγ revealed the indole core to be bound to a hydrophobic pocket of the ligand binding domain (pocket A) and the carboxylate to accept four H-bonds from the side chains of a triad of aromatic amino acids (two His and one Tyr) and a serine. This crystal structure provided a starting point for fragment growing. The indole N1−H offers an exit vector to an unoccupied hydrophobic pocket of the ligand binding domain (pocket B). After in silico evaluation of analogues of **19**, a small set of arylsulfonamides were prepared. This effort resulted in Indeglitazar (**20**) (Figure 11), which showed over 100-fold improvement in activity on the three PPAR subtypes.

We tested the ability of IADE to generate molecules filling pocket B of PPARγ by appending substituents to the indole N1−H of **19** (Figure 11). A single IADE run produced 1000 solutions. In nearly all of them (972), the carboxylate accepts H-bonds from the Ser289, His323, His449, and Tyr473 side chains. After visual inspection, we selected 82 solutions as potential candidates for follow-up (Figure 12 and Figure S14 in the Supporting Information). Nearly a quarter (18) of these 82 molecules have an arylsulfonamide substituent binding to pocket

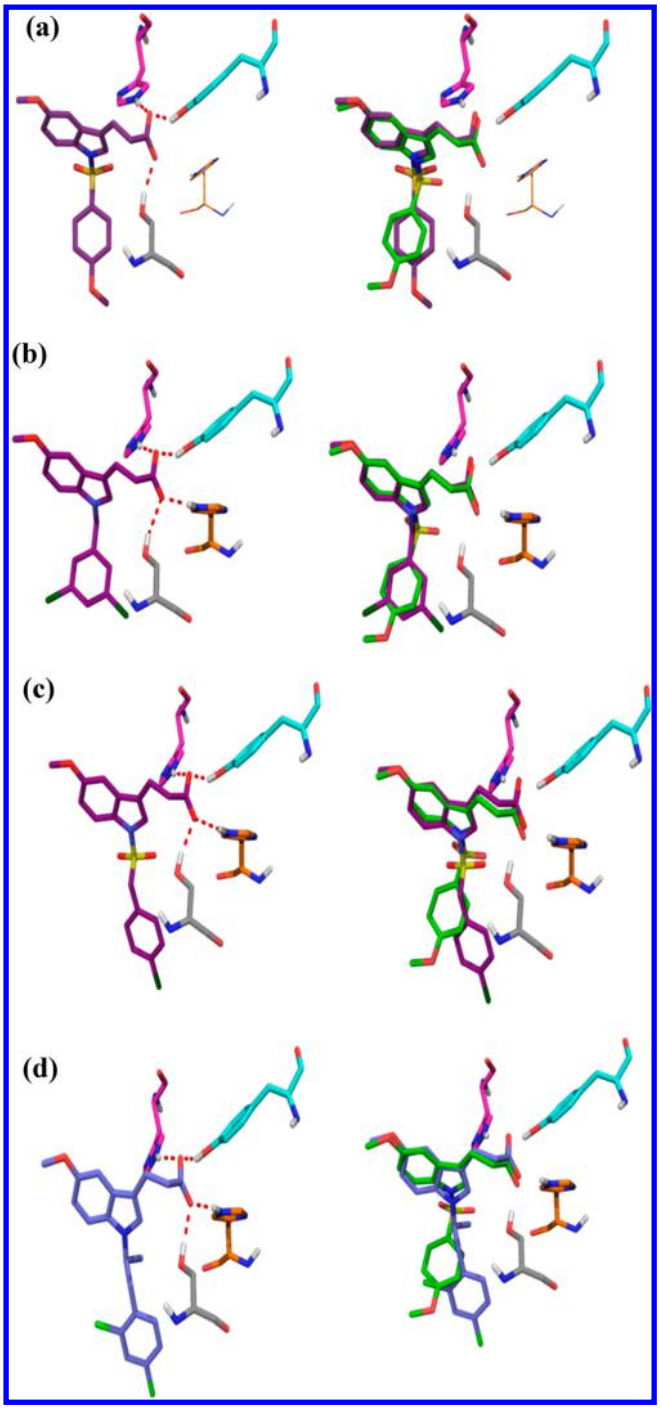


Figure 13. Representative IADE solutions docked in PPARγ. In each panel, the docking pose is presented on the left and an overlay of the representative IADE solution with the crystal structure of Indeglitazar bound to PPARγ (green sticks, PDB entry 3ET3) is shown on the right. For clarity, only selected amino acids are shown. Amino acids involved in H-bonds with the carboxylate of each IADE solution are displayed as sticks, and each H-bond is represented as a red dotted line between the donor and acceptor atoms. We have colored in gray the C atoms of Ser289, in orange the C atoms of His323, in magenta the C atoms of His449, and in cyan the C atoms of Tyr473. (a) Compound **20** (Indeglitazar); (b) compound **21** with a benzyl group replacing the sulfonamide; (c, d) compounds **22** and **23** with longer linkers between the indole and the phenyl ring binding to another area of pocket B.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

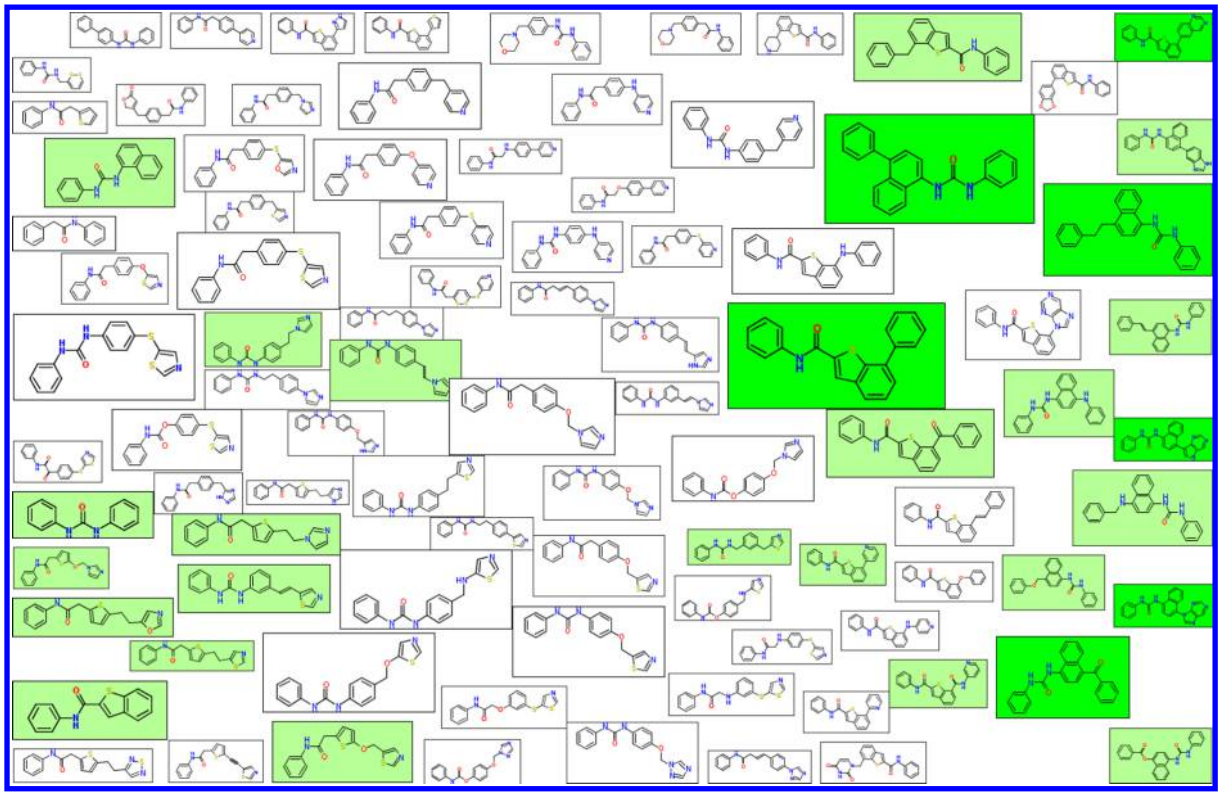


Figure 14. Scaffold map[67] of the 92 scaffolds found in at least two molecules. The size of each scaffold diagram is proportional to the number of molecules containing this scaffold. Each box is color-coded by the Glide docking score of the top-scoring molecule containing this scaffold. A dark-green background corresponds to a docking score better than −13 and a light-green background to a docking score between −13 and −12.

B. In particular, IADE generated Indeglitazar (**20**) (Figure 11). Although we did not include the water molecule interacting with both the acid and one of the sulfonamide O atoms in PDB entry 3ET3, IADE could reproduce the crystallographic binding mode of Indeglitazar (Figure 13a). Other selected solutions have a non-arylsulfonamide substituent binding to pocket B. A benzyl group turns out to be a perfect replacement for the arylsulfonamide of Indeglitazar (Figure 13b). On the other hand, IADE solutions with a longer spacer between the indole N1−H and the terminal aryl group bind to another region of pocket B (Figure 13c,d).

4. CONCLUSIONS

We have explored the possibility of growing a fragment bound to a protein using IADE, an in-house expert system that was initially developed to identify nonclassical bioisosteres. Three examples from the FBDD literature served as test cases. Starting from the initial fragments bound to their target, IADE could generate in each case close analogues of the published compounds or molecules sharing substructure elements with known ligands of a protein or protein family. Furthermore, this fragment growing protocol produced molecules adopting nearly the same binding mode or forming the same interactions with their target as known ligands of these proteins.

As a result of these three retrospective analyses, we also gained valuable insights into the preferred IADE settings for different binding sites. For small and enclosed binding sites as in HSP90 and PPARγ (Table 1), a single IADE run can produce chemically sensible solutions if the user defines the group to be exclusively modified. Because of the stochastic nature of IADE, it is worthwhile to run several jobs with different settings for larger binding sites as in p38 (Table 1). In this case, running IADE with the default parameters resulted in solutions filling only a portion of the binding site. We subsequently selected some of these solutions as starting points for separate rounds of fragment growing. This led to molecules filling the whole binding site. Although this incremental approach mirrors a real-life fragment growing project, our intervention affected the property distributions of the final products. Besides this three-step approach, we also modified the parameters to allow larger transformations. As a result, a single IADE job of 25 iterations generated solutions filling the whole p38 binding site. However, this latter job produced solutions with fewer different hinge binding fragments compared with the three-step approach (Table 2). In contrast to the single IADE job, the incremental approach afforded a more thorough sampling of the fragment space binding to a limited portion of the binding site. In other words, one can consider the two fragment growing strategies as complementary.

IADE was initially developed to generate analogues of a lead compound by adding or removing substituents, fusing rings, or changing atom types. Therefore, one may wonder how effectively IADE searches the available search space in fragment growing applications. The molecules produced for p38 are an ideal test case to assess the chemical space coverage of the IADE solutions. For this purpose, we applied an in-house method to analyze the scaffolds of the 807 molecules produced by runs 3 to 5 and passing the PLIF filters (Table S1 in the Supporting Information).[67] This analysis resulted in the identification of 489 unique scaffolds, 92 of them occurring in at least two molecules (Figure 14). This scaffold map illustrates the diversity of the IADE solutions.

Although IADE proceeds automatically, the user has to make several careful choices that will impact the outcome of a fragment growing exercise. To start with, one needs to select a protein structure, decide whether to keep some of the water molecules present in the binding site, and carefully prepare the selected

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

protein structure. IADE treats the protein as rigid. Obviously, the selection of a given protein conformation has a significant impact on the outcome of an IADE run. To handle protein flexibility, one could run IADE on multiple conformations of the target protein. During the preparation of a fragment growing run, the user can also set the search parameters, define substructure elements to be conserved or exclusively modified, and decide on the number of solutions to be kept. In this exercise, we kept the top 1000 solutions as ranked by the Glide scoring function. This retrospective study aimed to assess the ability of IADE to generate molecules forming interactions that are the hallmark of well-characterized ligand−protein systems. For this reason, we filtered these solutions on PLIFs. Finally, molecules passing these filters were minimized in the binding site and visually inspected. Because of the limitations of the current scoring functions, visual inspection of solutions remains an important step in the evaluation process. Clustering the IADE output on PLIFs could help the scientist to sift through a large number of solutions and visually inspect representatives of each cluster. One could also identify potentially interesting molecules that were poorly ranked by the Glide scoring function. Both deficiencies in the structure generation process and in the Glide scoring function can lead to poor ranking. In addition, visual inspection of solutions with a low score can help to generate ideas for structure modifications that can be further evaluated by computational techniques such as docking. Clustering the IADE output on PLIFs would be very useful in a prospective fragment growing exercise on a new or poorly characterized target. Finally, the modular nature of IADE affords the user the opportunity to consider alternative scoring functions. One could connect an IADE workflow to additional in silico models addressing on-target activity, selectivity, biopharmaceutical properties such as permeability, or liabilities such as hERG activity. Recently, Viklund and colleagues described such a multiparameter workflow for lead generation of β-secretase inhibitors.[68]

This evaluation has shown that IADE can provide useful ideas for fragment growing. In each case, it was able to generate analogues similar to known active compounds. We have also gained valuable experience to improve IADE and to optimize its performance for fragment growing.

■ ASSOCIATED CONTENT

Supporting Information

Figure S1: Examples of close analogues of **3** (Figure 2) generated after more than 10 iterations of IADE starting from **1** (Figure 2). For each solution, we provide the same information as in Figure 1. Figures S2−S4: Molecular structures of the 30 selected solutions from run 3, the 26 selected solutions from run 4, and the 64 selected solutions from run 5, respectively, for the p38 test case (Table 2). Figures S5−S7: Lists of known hinge binding fragments among the 30 selected IADE solutions from run 3, the 26 selected IADE solutions from run 4, and the 64 selected IADE solutions from run 5, respectively. The known hinge binders are from ref 58. Figures S8−S12: Molecular structures of the 17 selected solutions from the phenol and phenol isosteres class, the four selected solutions from the benzamide class, the five selected solutions from the benzamide isostere class, the solution from the aminopyrimidine class, and the 27 selected solutions from the class "others", respectively, for the HSP90 test case. Figure S13: Molecular structures of reference HSP90 inhibitors. Figure S14: Molecular structures of the 82 selected solutions for the PPARγ test case. Table S1: Required interaction patterns for filtering of the IADE solutions on PLIFs. Tables S2−S4: Distributions of computed physicochemical properties for the 30 selected solutions from run 3, the 26 selected solutions from run 4, and the 64 selected solutions from run 5, respectively, for the p38 test case. Table S5: Distribution of computed physicochemical properties for the 54 selected solutions for the HSP90 test case. Table S6: Distribution of computed physicochemical properties for the 82 selected solutions for the PPARγ test case. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: bernard.pirard@novartis.com.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

IADE, Intelligent Automatic Design; FBDD, fragment-based drug discovery; MW, molecular weight; CDK2, cyclin-dependent kinase 2; HSP90, heat shock protein 90; PPPARγ, peroxisome proliferator-activated receptor γ; PDB, Protein Data Bank; PLIF, protein−ligand interaction fingerprint; MM/GBSA, molecular mechanics/generalized Born surface area; ATP, adenosine triphosphate; ADP, adenosine diphosphate.

■ REFERENCES

(1) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent Developments in Fragment-Based Drug Discovery. *J. Med. Chem.* **2008**, *51*, 3661−3680.

(2) Scott, D. E.; Coyne, A. G.; Hudson, S. E.; Abell, C. Fragment-Based Approaches in Drug Discovery and Chemical Biology. *Biochemistry* **2012**, *51*, 4990−5003.

(3) Gill, A. New Lead Generation Strategies for Protein Kinase Inhibitors: Fragment-Based Screening Approaches. *Mini-Rev. Med. Chem.* **2004**, *4*, 301−311.

(4) Roughley, S. D.; Hubbard, S. D. How Well Can Fragments Explore Accessed Chemical Space? A Case Study from Heat Shock Protein 90. *J. Med. Chem.* **2011**, *54*, 3989−4005.

(5) Cheng, Y.; Judd, T. C.; Bartberger, M. D.; Brown, J.; Chen, K.; Fremeau, R. T.; Hickman, D.; Hitchcock, S. A.; Jordan, B.; Li, V.; Lopez, P.; Louie, S. W.; Luo, Y.; Michelsen, K.; Nixey, T.; Powers, T. S.; Rattan, C.; Sickmier, E. A.; St. Jean, D. J.; Wahl, R. C.; Wen, P. H.; Wood, S. From Fragment Screening to in Vivo Efficacy: Optimization of a Series of 2-Aminoquinolines as Potent Inhibitors of Beta-Site Amyloid Precursor Protein Cleaving Enzyme 1 (BACE1). *J. Med. Chem.* **2011**, *54*, 5836−5857.

(6) Andrews, S. P.; Brown, G. A.; Christopher, J. A. Structure-Based and Fragment-Based GPCR Drug Discovery. *ChemMedChem* **2014**, *9*, 256−275.

(7) Ward, R. A.; Brassington, C.; Breeze, A. L.; Caputo, A.; Critchlow, S.; Davies, G.; Goodwin, L.; Hassall, G.; Greenwood, R.; Holdgate, G. A.; Mrosek, M.; Norman, R. A.; Pearson, S.; Tart, J.; Tucker, J. A.; Vogtherr, M.; Whittaker, D.; Wingfield, J.; Winter, J.; Hudson, K. Design and Synthesis of Novel Lactate Dehydrogenase A Inhibitors by Fragment-Based Lead Generation. *J. Med. Chem.* **2012**, *55*, 3285−3306.

(8) Chung, C.-w.; Dean, A. W.; Woolven, J. M.; Bamborough, P. Fragment-Based Discovery of Bromodomain Inhibitors Part 1: Inhibitor Binding Modes and Implications for Lead Discovery. *J. Med. Chem.* **2012**, *55*, 576−586.

(9) Ferguson, F. M.; Fedorov, O.; Chaikuad, A.; Philpott, M.; Muniz, J. R. C.; Felletar, I.; von Delft, F.; Heightman, T.; Knapp, S.; Abell, C.; Ciulli, A. Targeting Low-Druggability Bromodomains: Fragment-Based Screening and Inhibitor Design against the BAZ2B Bromodomain. *J. Med. Chem.* **2013**, *56*, 10183−10187.

(10) Scott, D. E.; Ehebauer, M. T.; Pukala, T.; Marsh, M.; Blundell, T. L.; Venkitaraman, A. R.; Abell, C.; Hyvönen, M. Using a Fragment-Based Approach To Target Protein−Protein Interactions. *ChemBioChem* **2013**, *14*, 332−342.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

(11) Blomberg, N.; Cosgrove, D. A.; Kenny, P. W.; Kolmodin, K. Design of Compound Libraries for Fragment Screening. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 513−525.
(12) Vulpetti, A.; Hommel, U.; Landrum, G.; Lewis, R.; Dalvit, C. Design and NMR-Based Screening of LEF, a Library of Chemical Fragments with Different Local Environment of Fluorine. *J. Am. Chem. Soc.* **2009**, *131*, 12949−12959.
(13) Kawatkar, S.; Moustakas, D.; Miller, M.; Joseph-McCarthy, D. Virtual Fragment Screening: Exploration of MM-PBSA Re-Scoring. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 921−934.
(14) Zhu, T.; Lee, H.; Lei, H.; Jones, C.; Patel, K.; Johnson, M. E.; Hevener, K. E. Fragment-Based Drug Discovery Using a Multidomain, Parallel MD-MM/PBSA Screening Protocol. *J. Chem. Inf. Model.* **2013**, *53*, 560−572.
(15) Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, D. M.; Chung, C.-w. Fragment-Based Discovery of Bromodomain Inhibitors Part 2: Optimization of Phenylisoxazole Sulfonamides. *J. Med. Chem.* **2012**, *55*, 587−596.
(16) Talamas, F. X.; Ao-Leong, G.; Brameld, K. A.; Chin, E.; de Vicente, J.; Dunn, J. P.; Ghate, M.; Giannetti, A. M.; Harris, S. F.; Labadie, S. S.; Leveque, V.; Li, J.; Lui, A. S-T.; McCaleb, K. L.; Nájera, I.; Schoenfeld, R. C.; Wang, B.; Wong, A. De Novo Fragment Design: A Medicinal Chemistry Approach to Fragment-Based Lead Generation. *J. Med. Chem.* **2013**, *56*, 3115−3119.
(17) Zender, M.; Klein, T.; Henn, C.; Kirsch, B.; Maurer, C. K.; Kail, D.; Ritter, C.; Dolezal, O.; Steinbach, A.; Hartmann, R. W. Discovery and Biophysical Characterization of 2-Amino-Oxadizaoles as Novel Antagonists of PqsR, an Important Regulator of *Pseudomonas aeruginosa* Virulence. *J. Med. Chem.* **2013**, *56*, 6761−6774.
(18) Reader, J. C.; Matthews, T. P.; Klair, S.; Cheung, K-M. J.; Scanlon, J.; Proisy, N.; Addison, G.; Ellard, J.; Piton, N.; Taylor, S.; Cherry, M.; Fisher, M.; Boxall, K.; Burns, S.; Walton, M. I.; Westwood, I. M.; Hayes, A.; Eve, P.; Valenti, M.; de Haven Brandon, A.; Box, G.; van Montfort, R. L. M.; Williams, D. H.; Aherne, G. W.; Raynaud, F. I.; Eccles, S. A.; Garrett, M. D.; Collins, I. Structure-Guided Evolution of Potent and Selective CHK1 Inhibitors through Scaffold Morphing. *J. Med. Chem.* **2011**, *54*, 8328−8342.
(19) Barker, J. J.; Barker, O.; Courtney, S. M.; Gardiner, M.; Hesterkamp, T.; Ichihara, O.; Mather, O.; Montabelti, C. A. G. N.; Müller, A.; Varasi, M.; Whittaker, M.; Yarnold, C. J. Discovery of a Novel HSP90 Inhibitor by Fragment Linking. *ChemMedChem* **2010**, *5*, 1697−1700.
(20) Frank, A. O.; Feldkamp, M. D.; Kennedy, J. P.; Waterson, A. G.; Pelz, N. F.; Patrone, J. D.; Vangamudi, B.; Camper, D. V.; Rosanese, O. W.; Chazin, W. J.; Fesik, S. W. Discovery of a Potent Inhibitor of Replication Protein A Protein−Protein Interactions Using a Fragment-Linking Approach. *J. Med. Chem.* **2013**, *56*, 9242−9250.
(21) Hung, A. W.; Silvestre, H. L.; Wen, S.; Ciulli, A.; Blundell, T. L.; Abell, C. Application of Fragment Growing and Fragment Linking to the Discovery of Inhibitors of *Mycobacterium tuberculosis* Pantothenate Synthetase. *Angew. Chem., Int. Ed.* **2009**, *48*, 8452−8456.
(22) Saxty, G.; Woodhead, S. J.; Berdini, V.; Davies, T. G.; Verdonk, M. L.; Wyatt, P. G.; Boyle, R. G.; Barford, D.; Downham, R.; Garrett, M. D.; Carr, R. A. Identification of Inhibitors of Protein Kinase B Using Fragment-Based Lead Discovery. *J. Med. Chem.* **2007**, *50*, 2293−2296.
(23) Sun, H.; Tawa, G.; Wallqvist, A. Classification of Scaffold-Hopping Approaches. *Drug Discovery Today* **2012**, *17*, 310−324.
(24) Langdon, S. R.; Ertl, P.; Brown, N. Bioisosteric Replacement and Scaffold Hopping in Lead Generation and Optimization. *Mol. Inf.* **2010**, *29*, 366−385.
(25) Ertl, P.; Lewis, R. IADE: A System for Intelligent Automatic Design of Bioisosteric Analogs. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1207−1215.
(26) Evers, A.; Hessler, G.; Wang, L-h.; Werrel, S.; Monecke, P.; Matter, H. CROSS: An Efficient Workflow for Reaction-Driven Rescaffolding and Side-Chain Optimization Using Robust Chemical Reactions and Available Reagents. *J. Med. Chem.* **2013**, *56*, 4656−4670.
(27) Thompson, D. C.; Denny, R. A.; Nilakantan, R.; Humblet, C.; Joseph-McCarthy, D.; Feyfant, E. CONFIRM: Connecting Fragments Found in Receptor Molecules. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 761−772.
(28) Dey, F.; Caflisch, A. Fragment-Based de Novo Ligand Design by Multiobjective Evolutionary Optimization. *J. Chem. Inf. Model.* **2008**, *48*, 679−690.
(29) Ichihara, O.; Barker, J.; Law, R. J.; Whittaker, M. Compound Design by Fragment Linking. *Mol. Inf.* **2011**, *30*, 298−306.
(30) Pirard, B. The Quest for Novel Chemical Matter and the Contribution of Computer-Aided de Novo Design. *Expert Opin. Drug Discovery* **2011**, *6*, 225−231.
(31) Schneider, G.; Fechner, U. Computer-Based de Novo Design of Drug-like Molecules. *Nat. Rev. Drug Discovery* **2005**, *4*, 649−663.
(32) Böhm, H.-J. On the Use of LUDI To Search the Fine Chemicals Directory for Ligands of Proteins of Known Three-Dimensional Structure. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 623−632.
(33) Böhm, H.-J. LUDI: Rule-Based Automatic Design of New Substituents for Enzyme Inhibitor Leads. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 593−606.
(34) Böhm, H.-J. The Computer Program LUDI: A New Method for the de Novo Design of Enzyme Inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61−78.
(35) Degen, J.; Rarey, M. FlexNovo: Substructure-Based Searching in Large Fragment Spaces. *ChemMedChem* **2006**, *1*, 854−868.
(36) Teodoro, M.; Muegge, I. BIBuilder: Exhaustive Searching for de Novo Ligands. *Mol. Inf.* **2011**, *30*, 63−75.
(37) Beccari, A. R.; Cavazzoni, C.; Beato, C.; Costantino, G. LiGen: A High Performance Workflow for Chemistry Driven de Novo Design. *J. Chem. Inf. Model* **2013**, *53*, 1518−1527.
(38) Hoffer, L.; Renaud, J.-P.; Horvath, D. In Silico Fragment-Based Drug Discovery: Setup and Validation of a Fragment-to-Lead Computational Protocol Using S4MPLE. *J. Chem. Inf. Model.* **2013**, *53*, 836−851.
(39) Goldberg, D. R.; Hao, M.-H.; Qian, K. C.; Swinamer, A. D.; Gao, D. A.; Xiong, Z.; Sarko, C.; Berry, A.; Lord, J.; Magolda, R. L.; Fadra, T.; Kroe, R. R.; Kukulka, A.; Madwed, J. B.; Martin, L.; Pargellis, C.; Skow, D.; Song, J. J.; Tan, Z.; Torcellini, C. A.; Zimmitti, C. S.; Yee, N. K.; Moss, N. Discovery and Optimization of p38 Inhibitors via Computer-Assisted Drug Design. *J. Med. Chem.* **2007**, *50*, 4016−4026.
(40) Artis, D. R.; Lin, J. J.; Zhang, C.; Wang, W.; Mehra, U.; Perreault, M.; Erbe, D.; Krupka, H. E.; England, B. P.; Arnold, J.; Plotnikov, A. N.; Marimuthu, A.; Nguyen, H.; Wilb, S.; Signaevski, M.; Kral, J.; Cantwell, J.; Settachatgull, C.; Yan, D. S.; Fong, D.; Oh, A.; Shi, S.; Womack, P.; Powell, B.; Habets, G.; West, B. L.; Zhang, K. Y. J.; Milburn, M. V.; Vlasuk, G. P.; Hirth, K. P.; Nolop, K.; Bollag, G.; Ibrahim, P. N.; Tobin, J. F. Scaffold-Based Discovery of Indeglitazar, a PPAR Pan-Active Anti-Diabetic Agent. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 262−267.
(41) Damewood, J. R., Jr.; Lerman, C. L.; Masek, B. B. NovoFLAP: A Ligand-Based de Novo Design Approach for the Generation of Medicinally Relevant Ideas. *J. Chem. Inf. Model.* **2010**, *50*, 1296−1303.
(42) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindylalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(43) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(44) Haider, K.; Huggins, D. J. Combining Solvent Thermodynamic Profiles with Functionality Maps of the HSP90 Binding Site To Predict the Displacement of Water Molecules. *J. Chem. Inf. Model.* **2013**, *53*, 2571−2586.
(45) *SiteMap*, version 2.4; Schrödinger, LLC: New York, 2010.
(46) Ertl, P.; Rohde, B.; Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and its Application to the Prediction of Drug Transport Properties. *J. Med. Chem.* **2000**, *43*, 3714−3717.
(47) Ertl, P. Simple Quantum-Chemical Parameters as an Alternative to the Hammett Sigma Constants in QSAR studies. *Quant. Struct.-Act. Relat.* **1997**, *16*, 377−382.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193

(48) Walters, W. P.; Murcko, M. A. Prediction of "Drug-Likeness". *Adv. Drug Delivery Rev.* **2002**, *54*, 255−271.
(49) Ertl, P.; Schuffenhauer, A. Estimation of Synthetic Accessibility Score of Drug-like Molecules Based on Molecular Complexity and Fragment Contributions. *J. Cheminf.* **2009**, *1*, 8.
(50) Cheeseright, T.; Mackey, M.; Rose, S.; Vinter, A. Molecular Field Extrema as Descriptors of Biological Activity: Definition and Validation. *J. Chem. Inf. Model.* **2006**, *46*, 665−676.
(51) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. A.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(52) *Molecular Operating Environment (MOE)*, version 2013.08; Chemical Computing Group: Montreal, QC, Canada, 2014.
(53) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A. The VSGB 2.0 Model: A Next Generation Energy Model for High Resolution Protein Structure Modeling. *Proteins* **2011**, *79*, 2794−2812.
(54) Amir, M.; Somakala, K.; Ali, S. P38 MAP Kinase as Anti-Inflammatory Agents. *Mini-Rev. Med. Chem.* **2013**, *13*, 2082−2096.
(55) Cirillo, P. F.; Hickey, E. R.; Moss, N.; Breitfelder, S.; Betageri, R.; Fadra, T.; Gaenzler, F.; Gilmore, T.; Goldberg, D. R.; Kamhi, V.; Kirrane, T.; Kroe, R. R.; Madwed, J.; Moriak, M.; Netherton, M.; Pargellis, C. A.; Patel, U. R.; Qian, K. C.; Sharma, R.; Sun, S.; Swinamer, A.; Torcellini, C.; Takahashi, H.; Tsang, M.; Xiong, Z. Discovery and Characterization of the *N*-Phenyl-*N*′-Naphthylurea Class of p38 Kinase Inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *9*, 2386−2391.
(56) Liu, Y.; Gray, N. S. Rational Design of Inhibitors that Bind to Inactive Kinase Conformations. *Nat. Chem. Biol.* **2006**, *7*, 358−364.
(57) Regan, J.; Breitfelder, S.; Cirillo, P.; Gilmore, T.; Graham, A. G.; Hickey, E.; Klaus, B.; Madwed, J.; Moriak, M.; Moss, N.; Pargellis, C.; Pay, S.; Proto, A.; Swinamer, A.; Tong, L.; Torcellini, C. Pyrazole Urea-Based Inhibitors of p38 MAP Kinase: From Lead Compound to Clinical Candidate. *J. Med. Chem.* **2002**, *45*, 2994−3008.
(58) Xing, L.; Rai, B.; Lunney, E. A. Scaffold Mining of Kinase Hinge Binders in Crystal Structure Database. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 13−23.
(59) Pearl, L. H.; Prodromu, C. Structure and Mechanism of the HSP90 Molecular Chaperone Machinery. *Annu. Rev. Biochem.* **2006**, *75*, 271−294.
(60) Janin, Y. L. ATPase Inhibitors of Heat-Shock Protein 90, Second Season. *Drug Discovery Today* **2010**, *15*, 342−353.
(61) Dutta, R.; Inouye, M. GHKL, an Emergent ATPase/Kinase Superfamily. *Trends Biochem. Sci.* **2000**, *25*, 24−28.
(62) Biamonte, M. A.; Van de Water, R.; Arndt, J. W.; Scannevin, R. H.; Perret, D.; Lee, W. C. Heat Shock Protein 90: Inhibitors in Clinical Trials. *J. Med. Chem.* **2010**, *53*, 3−17.
(63) Murray, C. W.; Carr, M. G.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S.; Coyle, J. E.; Downham, R.; Figueroa, E.; Frederickson, M.; Graham, B.; McMenamin, R.; O'Brien, M. A.; Patel, S.; Phillips, T. R.; Williams, G.; Woodhead, A. J.; Woolford, A. J.-A. Fragment-Based Drug Discovery Applied to HSP90. Discovery of Two Lead Series with High Ligand Efficiency. *J. Med. Chem.* **2010**, *53*, 5942−5955.
(64) Huth, J. E.; Park, C.; Petros, A. M.; Kunzer, A. R.; Wendt, M. D.; Wang, X.; Lynch, C. L.; Mack, J. C.; Swift, K. M.; Judge, R. A.; Chen, J.; Richardson, P. L.; Jin, S.; Tahir, S. K.; Matayoshi, E. D.; Dorwin, S. A.; Ladror, U. S.; Severin, J. M.; Walter, K. A.; Bartley, D. M.; Fesik, S. W.; Elmore, S. W.; Hajduk, P. J. Discovery and Design of Novel HSP90 Inhibitors Using Multiple Fragment-Based Design Strategies. *Chem. Biol. Drug Des.* **2007**, *70*, 1−12.
(65) Brough, P. A.; Barril, X.; Borgognoni, J.; Chene, P.; Davies, N. G. M.; Davis, B.; Drysdale, M. J.; Dymock, B.; Eccles, S. A.; Garcia-Echeverria, C.; Fromont, C.; Hayes, A.; Hubbard, R.; Parsons, R.; Radimerski, T.; Raynaud, F. I.; Robertson, A.; Roughley, S. D.; Schoepfer, J.; Simmonite, H.; Sharp, S. Y.; Surgenor, A.; Valenti, M.; Walls, S.; Webb, P.; Wood, M.; Workman, P.; Wright, L. M. Combining Hit Identification Strategies: Fragment-Based and in Silico Approaches to Orally Active 2-Aminothieno[2,3-*d*]pyrimidine Inhibitors of the HSP90 Molecular Chaperone. *J. Med. Chem.* **2009**, *52*, 4794−4809.
(66) Wilson, T. M.; Brown, P. J.; Sternbach, D. D.; Henke, B. R. The PPARs: From Orphan Receptors to Drug Discovery. *J. Med. Chem.* **2000**, *43*, 527−550.
(67) Ertl, P. Intuitive Ordering of Scaffolds and Scaffold Similarity Searching Using Scaffold Keys. *J. Chem. Inf. Model.* **2014**, *54*, 1617−1622.
(68) Viklund, J.; Kolmodin, K.; Nordvall, G.; Swah, B.-M.; Svensson, M.; Gravenfors, Y.; Rahm, F. Creation of Novel Cores for *β*-Secretase (BACE-1) Inhibitors: A Multiparameter Lead Generation Strategy. *ACS Med. Chem. Lett.* **2014**, *5*, 440−445.

DOI: 10.1021/ci5006355
J. Chem. Inf. Model. 2015, 55, 180−193