JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

In Silico Deconstruction of ATP-Competitive Inhibitors of Glycogen Synthase Kinase-3β

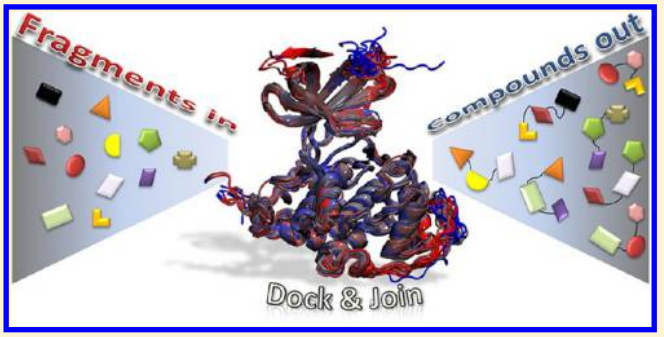
Paola Bisignano,[†] Chiara Lambruschini,[†] Manuele Bicego,[‡,§] Vittorio Murino,[§] Angelo D. Favia,*[,†,⊥,○] and Andrea Cavalli[†,∥,○]

[†](Department of) Drug Discovery and Development and [§]Pattern Analysis & Computer Vision, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy

[‡]Dipartimento di Informatica, Università degli Studi di Verona, Ca' Vignal 2, strada Le Grazie 15, 37134 Verona, Italy

[∥]Dipartimento di Scienze Farmaceutiche, Università di Bologna, via Belmeloro 6, 40126 Bologna, Italy

ABSTRACT: Fragment-based methods have emerged in the last two decades as alternatives to traditional high throughput screenings for the identification of chemical starting points in drug discovery. One arguable yet popular assumption about fragment-based design is that the fragment binding mode remains conserved upon chemical expansion. For instance, the question of the binding conservation upon fragmentation of a molecule is still unclear. A number of papers have challenged this hypothesis by means of experimental techniques, with controversial results, "underlining" the idea that a simple generalization, maybe, is not possible. From a computational standpoint, the issue has been rarely addressed and mostly to test novel protocols on limited data sets. To fill this gap, we here report on a computational retrospective study concerned with the in silico deconstruction of leadlike compounds, active on the pharmaceutically relevant enzyme glycogen synthase kinase-3β.



INTRODUCTION

Fragment-based (FB) methods[1] have recently emerged as alternatives to traditional high throughput screenings (HTS) for the identification of chemical starting points in drug discovery programs.[2] The success of this approach is underlined both by the number of compounds, coming from FB drug design (FBDD) efforts, that are currently in phase 1 and 2 clinical trials: ABT-263[3,4] from Abbott, directed against cancer target Bcl-Xl; AT7519[5−7] from Astex, against CDK2; LY-517717[8,9] from Lilly, against factor Xa of coagulation; Indeglitazar[10] from Plexxikon, against PPARγ, involved in type 2 diabetes mellitus; and Tideglusib[11] from Noscira, a glycogen sinthase kinase-3β (GSK-3β) inhibitor for the treatment of Alzheimer's disease and progressive supranuclear palsy. In addition, there is a plethora of compounds currently in preclinical development.[12−18]

FB methods, officially introduced by Abbott in 1996,[19] rely on the identification of small-sized binders of the target of interest with affinities typically in the micromolar range. Screening collections of small compounds (MW < 300 Da) has the advantage to cover larger chemical spaces with respect to leadlike molecules.[20] Compounds with up to 12 heavy atoms make up to about 10^7 diverse molecules, while this number rises to 10^{60} for those containing up to 30 heavy atoms. All those features result in reduced costs for biological assays, quicker experimental responses, and enhanced probability of finding hits.[21−23] A further appealing advantage of FBDD strategies is that, even in crowded intellectual property spaces, new areas of worth-exploring chemical landscape can be unveiled.

Once a hit is found, it has to be grown into a bigger molecule by stepwise addition of functional groups. When possible, direct joining of two or more fragments might be attempted,[22] relying on the hypothesis that fragment binding modes will be retained upon expansion.[24−26]

On the downside, because of their low binding affinities for macromolecules (>10 μM), the hit detection via experimental screening can be done only with assays possessing a very high sensitivity, able to detect affinities in the micromolar to millimolar range. Among others, biophysical techniques, such as nuclear magnetic resonance (NMR),[14,17,19,25,27,28] surface plasmon resonance (SPR),[15,17,28,29] or isothermal titration calorimetry (ITC),[28,30] have played a major role in the determination of protein−fragment binding affinities. Such methods can be used in conjunction with X-ray crystallography[12,13,26,31] to gain detailed structural information useful for the subsequent fragment growth. However, X-ray crystallography is costly, might not be suitable for some fragment−protein pairs, and can hardly be used routinely.[27]

In this scenario, computational approaches,[32] such as molecular docking, could effectively complement experimental techniques and aid the fragment hit identification[28,33,34] and hit to lead[35−38] phases. However, a systematic reliability of computational methods tuned for bigger molecules has not been proven yet in FBDD. Doubts might arise on the adequacy of current scoring functions to discriminating energetically close binding poses and to ranking

Received: July 28, 2012
Published: November 30, 2012

dx.doi.org/10.1021/ci300355p | J. Chem. Inf. Model. 2012, 52, 3233−3244

Table 1. Data Set of the 29 GSK-3β Structures Analyzed in This Study, by PDB Codes (in Alphabetical Order)

PDB ID	ligand ID	ligand name	resolution (Å)
1GNG	none	NA	2.60
1H8F	none	NA	2.80
1I09	none	NA	2.70
1J1B	ANP	adenylylimidodiphosphate	1.80
1J1C	ADP	adenosine diphosphate	2.10
1PYX	ANP	adenylylimidodiphosphate	2.40
1Q3D	STU	staurosporine	2.20
1Q3W	ATU	alsterpaullone	2.30
1Q41	IXM	indirubin-3′-monoxime	2.10
1Q4L	679	I-5,3-anilino-4-arylmaleymide	2.77
1Q5K	TMU	*N*-(4-methoxybenzyl)-*n*′-(5-nitro-1,3-thiazol-2-yl) urea (AR-A01448)	1.94
1R0E	DFN	3-indolyl-4-aryl maleimide	2.25
1UV5	BRW	6-bromoindirubin-3′-oxime	2.80
2JLD	AG1	ruthenium pyridocarbazole	2.35
2OW3	BIM	bis-(indole) maleimide pyridinophane	2.80
2O5K	HBM	2-(2,4-dichloro-phenyl)-7-hydroxy-1*h*-benzoimidazole-4-carboxylic acid [2]-amide	3.20
3DU8	553	(7*s*)-2-(2-aminopyrimidin-4-yl)-7-(2-fluoroethyl)-1,5,6,7-tetrahydro-4*h*-pyrrolo[3]pyridin-4-one	2.20
3F7Z	34O	2-(1,3-benzodioxol-5-yl)-5-[(3-fluoro-4-methoxybenzyl) sulfanyl]-1,3,4-oxadiazole	2.40
3F88	3HT, 2HT	5-[1-(4-methoxyphenyl)-1*h*-benzimidazol-6-yl]-1,3,4-oxadiazole-2(3*h*)-thione and 3-methylbenzonitrile (the cocrystal structure with compound 20x could not be fully characterized, due to cleavage of the S−C bond in the X-ray beam to form the debenzylated compound)	2.60
3GB2	G3B	2-methyl-5-(3-{4-[(*s*)-methylsulfinyl]phenyl}-1- benzofuran-5-yl)-1,3,4-oxadiazole	2.40
3I4B	Z48	*N*-[(1*s*)-2-hydroxy-1-phenylethyl]-4-[5-methyl-2-(phenylamino)pyrimidin-4-yl]-1*h*-pyrrole-2-carboxamide	2.30
3L1S	Z92	4*z*)-4-[(4-chlorophenyl)hydrazono]-5-(3,4- dimethoxyphenyl)-2*h*-pyrazol-3-one	2.90
3M1S	DW1	ruthenium pyridocarbazole	3.13
3PUP	OS1	ruthenium octasporine	2.99
3Q3B	55E	4-(4-hydroxy-3-methylphenyl)-6-phenylpyrimidin-2(5*H*)-one	2.70
3ZRL	ZRL	7-bromo-2-pyridin-4-yl-5*H*-thieno[3,2-C]pyridin-4-one	2.48
3ZRK	ZRK	2-pyridin-4-ylfuro[3,2-*c*]pyridin-4(5*H*)-one	2.37
3ZRM	ZRM	7-(4-hydroxyphenyl)-2-pyridin-4-ylthieno[3,2-*c*]pyridin-4(5*H*)-one	2.49
3SD0	TSK	3-(5-fluoranyl-6-iodanyl-1-methyl-indol-3-yl)-4-(7-methoxy-1-benzofuran-3-yl)pyrrole-2,5-dione	2.70

fragment data sets.[39,40] An accurate evaluation of the fragment interactions is crucial in order to pinpoint true binders and to develop sound protocols apt to evolve hits into promising leads.[21]

One popular assumption about FB design is that the fragment binding mode remains conserved upon chemical expansion. For instance, the question of the binding conservation upon fragmentation of a molecule is still unclear. Do the fragments still display detectable affinity for the target? And if so, do they keep their initial binding site on the macromolecule? A number of papers have challenged this assumption by means of experimental techniques, with controversial results, "underlining" the idea that a simple generalization is unlikely.[25,26,41,42] From a computational standpoint, the issue has been rarely addressed and mostly to test novel protocols on limited data sets.[43] To fill this gap, in this paper, we report on a computational retrospective study aimed at deconstructing, in silico, leadlike compounds active on the enzyme GSK-3β.

The work is concerned with the following: If an inhibitor is virtually fragmented, do the constituting parts recapitulate their native binding poses upon docking? How thorough must the sampling be in order to reach the native pose? GSK-3β, a protein kinase (transfers the γ-phosphate group of ATP to the target substrates), belonging to the serine/threonine family was selected as a case study,[44] based on the following considerations: (i) GSK-3β is a therapeutically relevant target that has attracted considerable attention in the past decade because of its involvement in neurodegenerative diseases and several other ailments;[45−49] (ii) the richness of structural data on this enzyme. In particular, there are to date 30 X-ray structures of this enzyme available at the Worldwide Protein Data Bank (wwPDB),[50] mostly in complex with inhibitors, each of which contains the complete kinasic domain. Thanks to those features, GSK-3β represented an ideal benchmark to study the potential of in silico fragment-based joining techniques.

RESULTS AND DISCUSSION

Data Sets. *Proteins.* The wwPDB was parsed to retrieve all the available crystal structures of wild-type human GSK-3β. Twenty-nine structures were retrieved and analyzed. A comprehensive list with some structural details of the whole set is reported in Table 1. The average resolution was reasonably good (2.5 Å), and most of the X-ray models were holo structures (i.e., 26 out of 29) in complex with diverse ATP-mimetic inhibitors.

As it can be inferred from Figure 1, where the whole data set is spatially superimposed, most of the structural variations could be observed in the glycine-rich loop and in the activation loop. Conversely, the structural fluctuations in the ATP binding site were very limited.

These findings could also be anticipated from the experimental B-factors of the entries, taken singularly. In particular, the glycin-rich loop functions as a molecular adapter that adjust according to the features of the molecule bound at the hinge-region. Therefore, given the structural diversity of the molecules found in complex with GSK-3β, structural variations in glycin-rich loop could be anticipated.

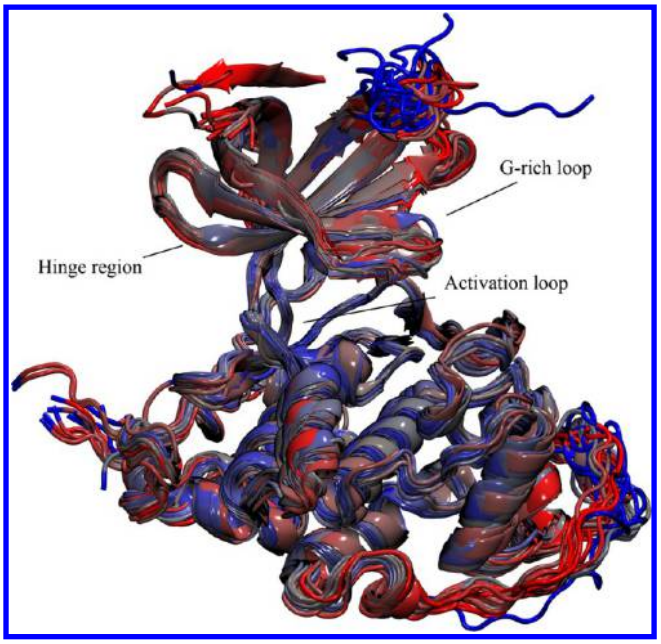


Figure 1. Complete collection of wwPDB X-ray structures of GSK-3β, as available at the wwPDB. Proteins are superimposed, represented as ribbons, and colored according to the experimental B-factors. The glycine rich loop, activation loop (where the activating phosphorylation occurs), and the site where the ATP binds (hinge region) are highlighted. The picture was realized using VMD-1.9.1.[51]

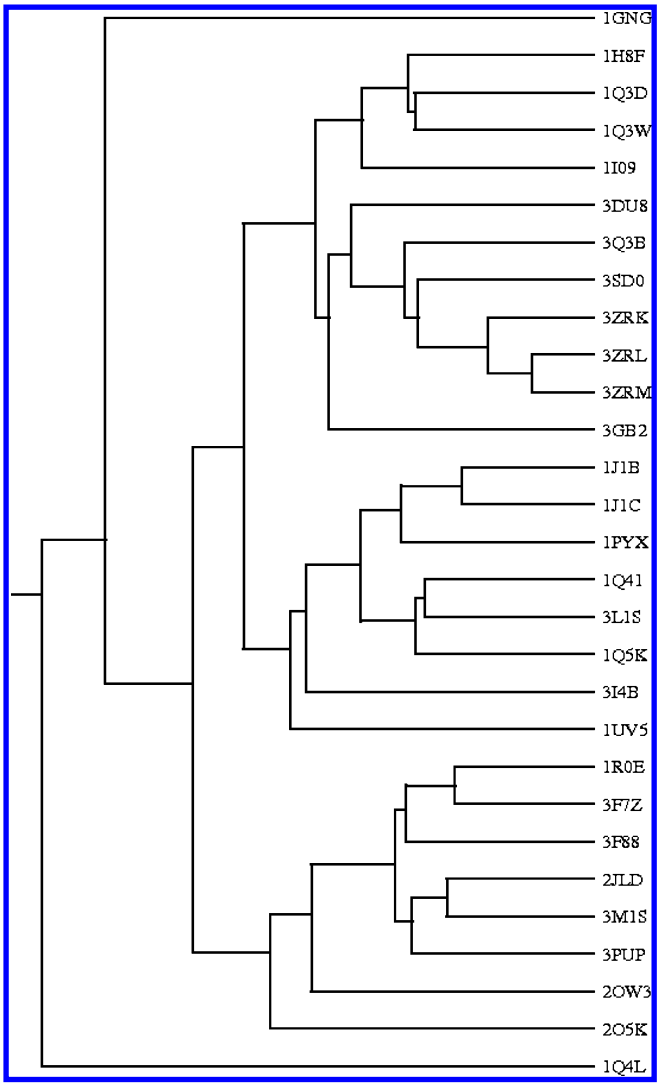


Figure 2. UPGMA-generated tree of the 29 wild-type protein structures of GSK-3β as available at the wwPDB. The distances were obtained through the comparison of LJ potentials calculated at the ATP-binding sites.

To further investigate the structural diversity of the models deposited at the wwPDB, we performed a pairwise comparison of the 29 wild type structures, using a method recently developed by us[52] that relies on the calculation of Lennard-Jones (LJ) potentials at regularly spaced grid points within a site of interest. A tree, based on the distances between ATP-binding regions calculated with this method, is reported in Figure 2.

The distance between different entries reflects directly the conformational differences of the X-ray models caused by the chemical nature of the ligands. In particular, proteins 1GNG and 1I09 stand as singletons, since they were both apo structures. Interestingly, the other apo structure present in the initial set, protein 1H8F, seems to be somehow related to 3Q3B, a protein structure in complex with a phenylpyrimidinone derivative. Proteins 1J1B, 1J1C, and 1PYX belong to the same cluster, and this is most likely due to the fact that their complexed ligands share some chemical similarities, being all very close ATP analogues. Despite some structural difference (e.g., side chain of F67), the pdb structure 3DU8 was located close to the cluster composed of 1J1B, 1J1C, and 1PYX. A fact worth highlighting is that even ligands that are not apparently closely related may induce the protein to adopt similar conformations, such as for the case of the pairs 1Q5K and 3I4B or 3F88 and 3PUP. Such analogies and dissimilarities would be partially visible from a coarser comparison made via rmsd, since no interaction fields (the one experienced by the interacting ligand) would be taken into account.

Eight structures had to be discarded (see Methods for details related to the structures selection), thus yielding a grand total of 21 structures to work with.

Fragmentations of Leadlike Compounds. Each of the selected structures had bound an ATP-mimetic inhibitor of GSK-3β. In Figure 3, the 21 cocrystallized inhibitors and their constituting fragments are shown. The circular tree reflects the Tanimoto distances between leadlike compounds calculated on a molecular description based on the Daylight fingerprinting algorithm.[54]

Grouping at a distance of 0.3 resulted in 16 clusters, thus highlighting the presence of fairly different chemotypes. The only trivially similar leadlike compounds present were non-hydrolyzable types of ATP and ADP (entries 1PYX, 1J1B, and 1J1C), and two indirubin derivatives (entries 1Q41 and 1UV5). Not including the ATP derivatives, whose importance as GSK-3β blockers in drug discovery is extremely limited, the majority of the extracted compounds were nanomolar inhibitors. Some of them had also been tested for cross-selectivity toward other protein kinases, in order to anticipate the possible appearance of undesired effects upon in vivo administration. Being all ATP-mimetics, the cocrystallized molecules were all found at the protein hinge region, where the typical kinase H-bond (HB) donor/acceptor alternation is present. Depending on the analyzed cases, other types of contacts could be observed (e.g., cation$-\pi$, hydrophobic, etc.); however, the enthalpy-driven interaction with the hinge region was possibly the most common main driving interaction.

dx.doi.org/10.1021/ci300355p | J. Chem. Inf. Model. 2012, 52, 3233−3244

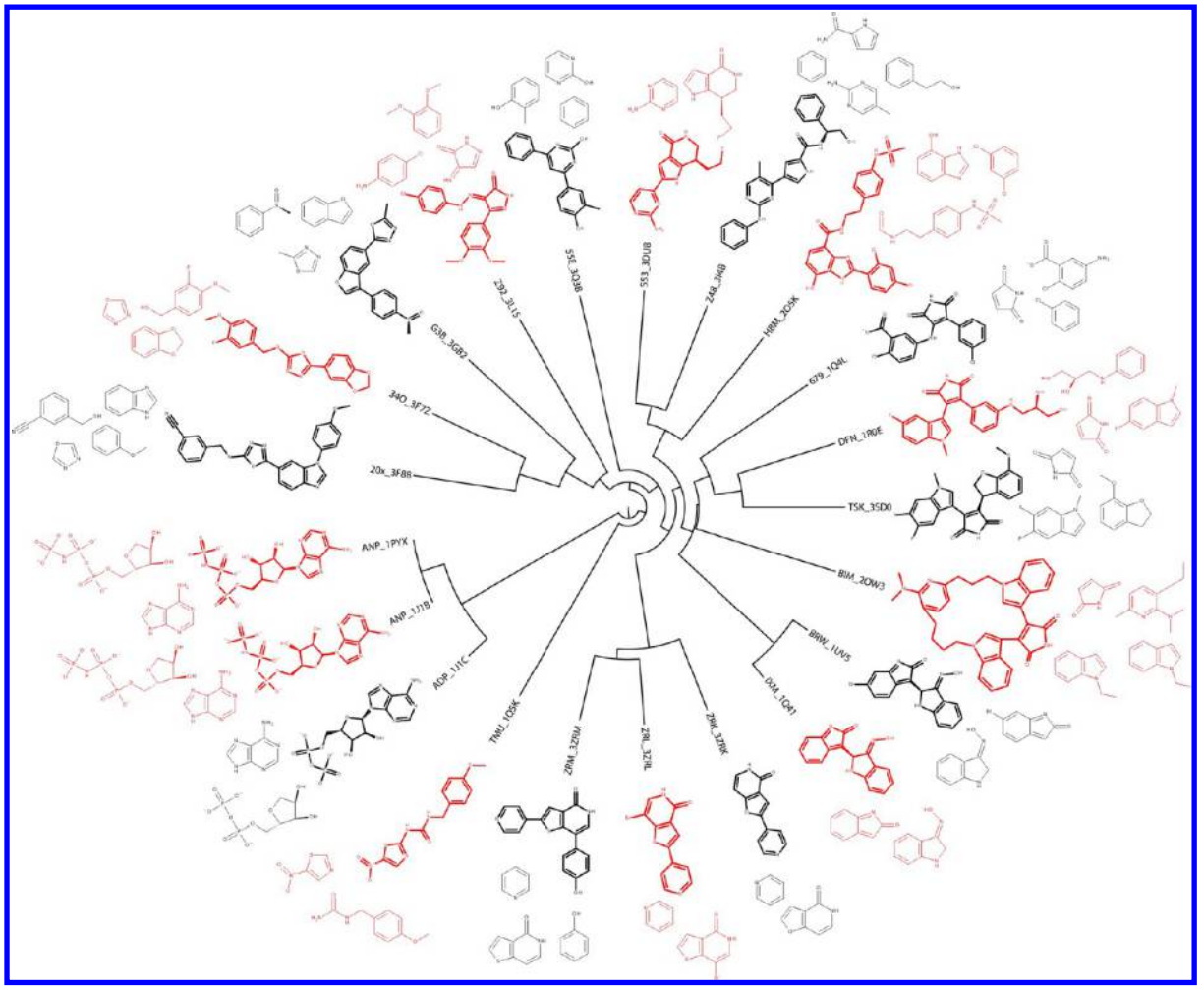


Figure 3. UPGMA-generated circular tree of the 21 leadlike molecules and their corresponding 57 constituting fragments employed in this study. The distances between leadlike compounds were calculated using the Tanimoto metrics, based on the Daylight fingerprinting algorithm. The labels show the molecule's 3-letter code and the PDB code. To help interpretation, bonds of the leadlike molecules are bolder than those of the fragments.

The deconstruction of the compounds originated 57 fragments. Eliminating redundancy (see Methods for details), the final fragment set consisted of 45 diverse small molecules. In Figure 4, eight descriptors are employed to illustrate the average physicochemical features of both leadlike and fragment sets.

The average molecular weight of the cocrystallized molecules was close to 400 Da. This value dropped to 130 Da for the fragment set. The great variability observed for the number of rotatable bonds, HB propensity, and polar surface area was significantly triggered by the presence in the set of ATP-derivatives. Such molecules, in fact, differ greatly from the rest of the set, since they were not meant to be developed as drugs but rather to be used as chemical tools. One clear trend that could be observed was that both fragments and leadlike molecules possessed more HB acceptors than donors. A great variability could also be observed for the logP, as a consequence of the fact that GSK-3β is involved in a number of diseases, some of which require inhibitors able to cross the blood−brain barrier (BBB).

Docking of Fragments. In order to recreate the starting lead, each of its constituting fragments had to dock into the proper position so to enable the joining algorithm to form the right connections. Since docking aims to find energy minima of a given molecular pair, one might investigate whether the isolated fragments, as directly extracted from the X-ray, lie in an energy minimum or, at least, in its close proximity. In Figure 5, a summary of this analysis is reported.

Almost 73% of the times, the fragments were close either to an absolute or to a relative energy minimum. In fact, in these cases, a simple local minimization displaced the fragment of less than 1 Å rmsd. In almost 25% of the cases analyzed, the minimization induced a drift comprised between 1 and 2 Å rmsd. In a very small percentage of cases (1.7%), the movement was more pronounced (between 2 and 3 Å rmsd). Considering that 1.5−2 Å rmsd could be used to classify as successful a docking run, the three different above-reported scenarios were increasingly challenging due to the fact that the higher the drift, the lower the chances to recapitulate the pose within acceptable energy windows.

The fragments were then blindly docked into the protein structures. For the sake of completeness, we report in Figure 6 the results of the docking simulations using as a reference both the starting X-ray and the fragments after the local minimization.

Approximately 70% of the time, we were able to find a pose within 1 Å rmsd from the reference X-ray. This result was remarkably in agreement with the 75% of X-ray fragments already in an energetically stable pose. In addition, although rmsd could be an arguable metric to assess how close in space two configurations are, this was an encouraging result, as far as the lead regeneration process was concerned.[55] Conversely, the chances to recapitulate the starting lead were lower when none

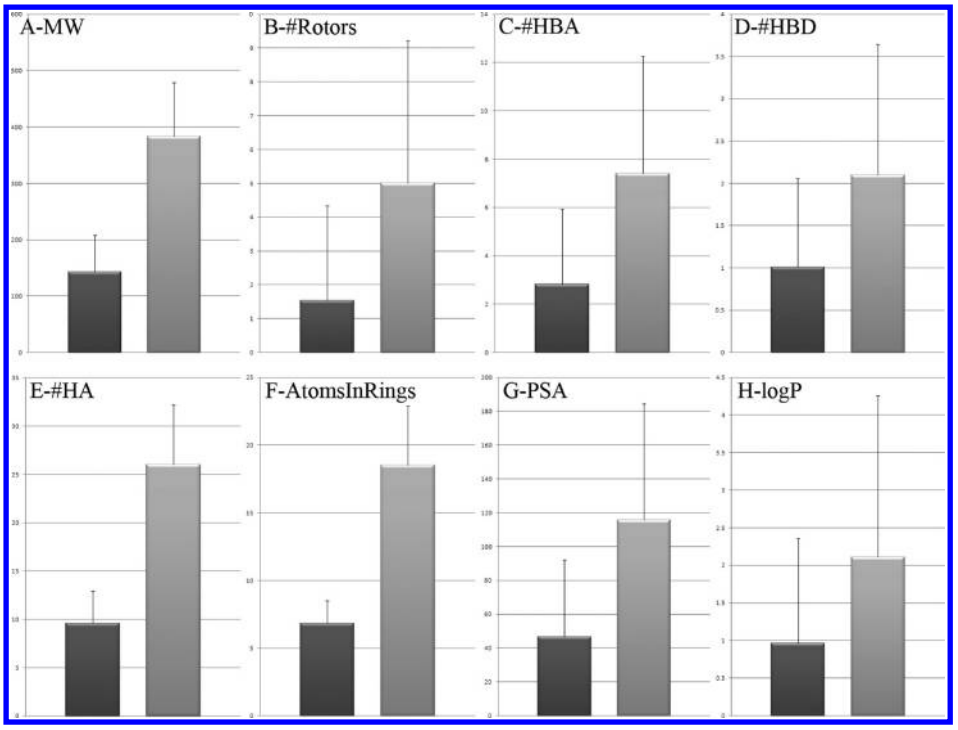


Figure 4. Eight average molecular properties of the leadlike and fragment sets (depicted in gray and black, respectively): (A) molecular weight, (B) number of rotors, (C) HB acceptors count, (D) HB donors count, (E) heavy atoms count, (F) number of atoms in rings, (G) polar surface area, and (H) logP.

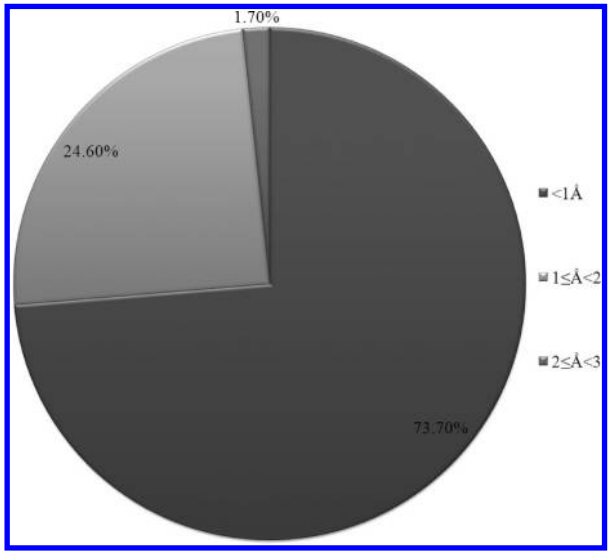


Figure 5. Pie chart summarizing, in the rmsd space, how close to an energy minimum are the fragments obtained from the deconstruction of the leadlike compounds found in the X-ray structures.

of the constituting fragments could approach the crystallographic pose, such as the cases where the closest docking solution lay at more than 2−3 Å rmsd from the reference. Expectedly, results on the isolated fragments were better if the fragments minimized in the proteins frame were taken as a reference instead. In this case, the percentage of poses found within 1 Å rmsd from the reference rose to almost 85%, while the percentage of totally irreproducible configurations was limited to 3.5% (i.e., rmsd > 3 Å). We should however comment that the proximity to the minimized pose of the fragment does not speak to the ease of recapitulating the starting lead but rather to the energetic difference between the fragment closest energy minimum and the other minima found by the docking algorithm.

In summary, as shown, in a certain number of the cases we were able to reproduce the configuration of the fragments originating from the deconstruction of a bigger lead. A relevant aspect to take into account while analyzing these data is related to the computational cost. How many poses do we have to generate to dock in close proximity of the true solution? This relates to the energetic difference between the target pose and the global minimum. Typically, in a docking effort, only the uppermost fraction of the energetically ranked distribution is taken into account for further validation. This is particularly true when molecules of high molecular weight are considered. In the case of fragments, the analysis might differ. Because of their intrinsic simplicity, fragments could adopt multiple iso-energetic configurations.[55,56] Hence, the number of poses one must consider should be higher, so not to miss important solutions. In this retrospective work, the number of generated poses per fragment was exceptionally set to 150, to allow us to draw some conclusion on the topic.

The pie charts in Figure 7 are concerned with the number of poses needed to be within 1 Å rmsd of the reference X-ray or of the minimized fragment (top and bottom, respectively).

If the X-ray was taken as a reference, in 44% of the cases, less than 10 poses were enough to recap the reference structure. This represents an ideal scenario where, despite the fragment propensity to bind at alternative sites, the docking solutions approximated adequately the experimental complex within an acceptable energy window. Generating 20 poses to recapitulate a reference structure might still be an acceptable threshold. Considering this upper limit, almost 50% of attempts resulted in a reliable docking prediction. The remaining cases were not accurately reproduced and more than 20 poses per fragment were needed to identify the X-ray pose. Worryingly, in 32% of the cases, 150 poses were not enough to reach the proximity of

dx.doi.org/10.1021/ci300355p | J. Chem. Inf. Model. 2012, 52, 3233−3244

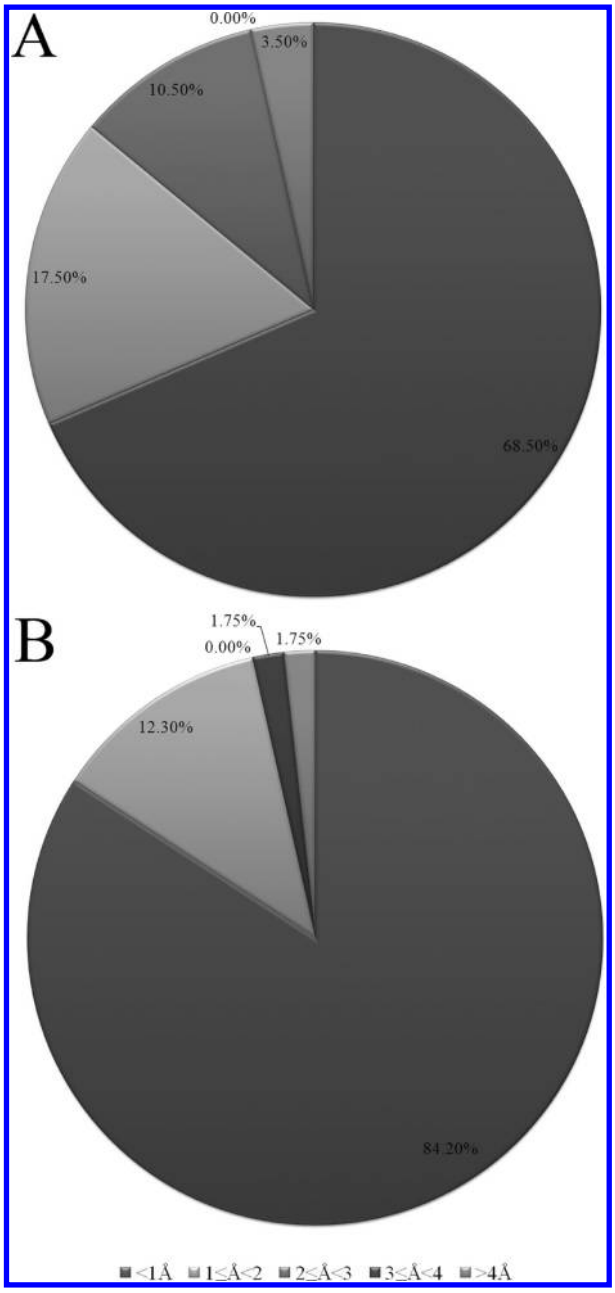


Figure 6. Pie charts summarizing how close in the rmsd space the fragment docking poses could get to the true solutions. The original X-ray (A) and the fragments after minimization (B) were used as a reference.

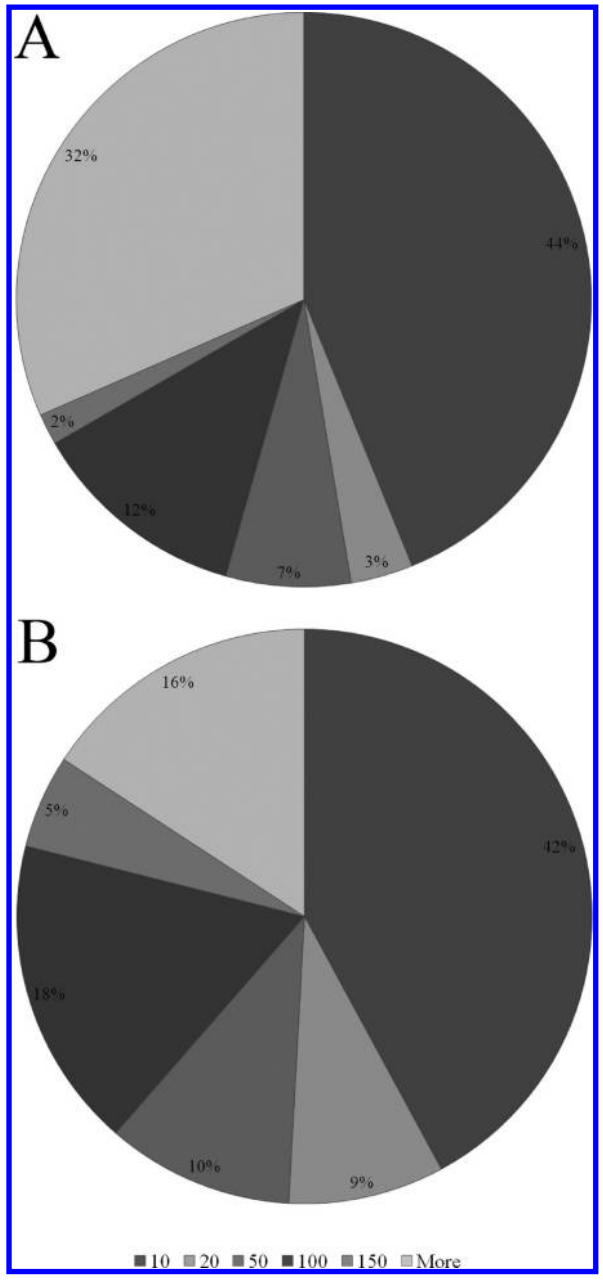


Figure 7. Pie charts summarizing the number of poses needed to be within 1 Å rmsd of the reference pose. The original X-ray (A) and the fragments after minimization (B) were used as a reference.

the reference configuration. Also in this case, figures got better if the minimized fragments were used as a reference. In this case, only 16% of the studied protein−fragment pairs were not recapitulated within 150 docking poses. In addition, we tried to increase the rmsd tolerance to a more conservative value of 2 Å and to rescore the docking poses via MM-GBSA,[57−59] but the overall figures did not change significantly (data not shown). One reason behind these failures can be found in the rules employed here to dissect the lead compounds. Especially when an heteroatom was involved in the fragmentation, the chemical properties of the resulting molecule were likely to differ substantially from the original one, due, for example, to altered H-bond properties (e.g., tertiary amines originating secondary amines).

It is worth noting here that, in each analyzed case, at least one of the fragments constituting the lead could be found within 1 Å rmsd from the reference, among the top 20 scoring poses This was most likely due to the fact that the driving interaction, i.e. hydrogen bond donor/acceptor, with the hinge region was constantly found, while the other interactions (where present) were re-enacted with alternate success. For instance, flat rings containing the H-bond donor/acceptor alternation such as maleimide (e.g., PDB IDs 2OW3, 3SD0, and 1R0E), adenine (e.g., PDB IDs 1PYX, 1J1B, and 1J1C), or pyrazolone (e.g., PDB ID 3L1S), to name a few, re-enacted regularly the configuration of the starting lead molecule. Rather than being a docking intrinsic limitation (most of those rings are, in fact, known to be pan-kinase binders), this speaks to the complexity of the system under investigation. In fact, ATP-mimetic inhibitors

dx.doi.org/10.1021/ci300355p | J. Chem. Inf. Model. 2012, 52, 3233−3244

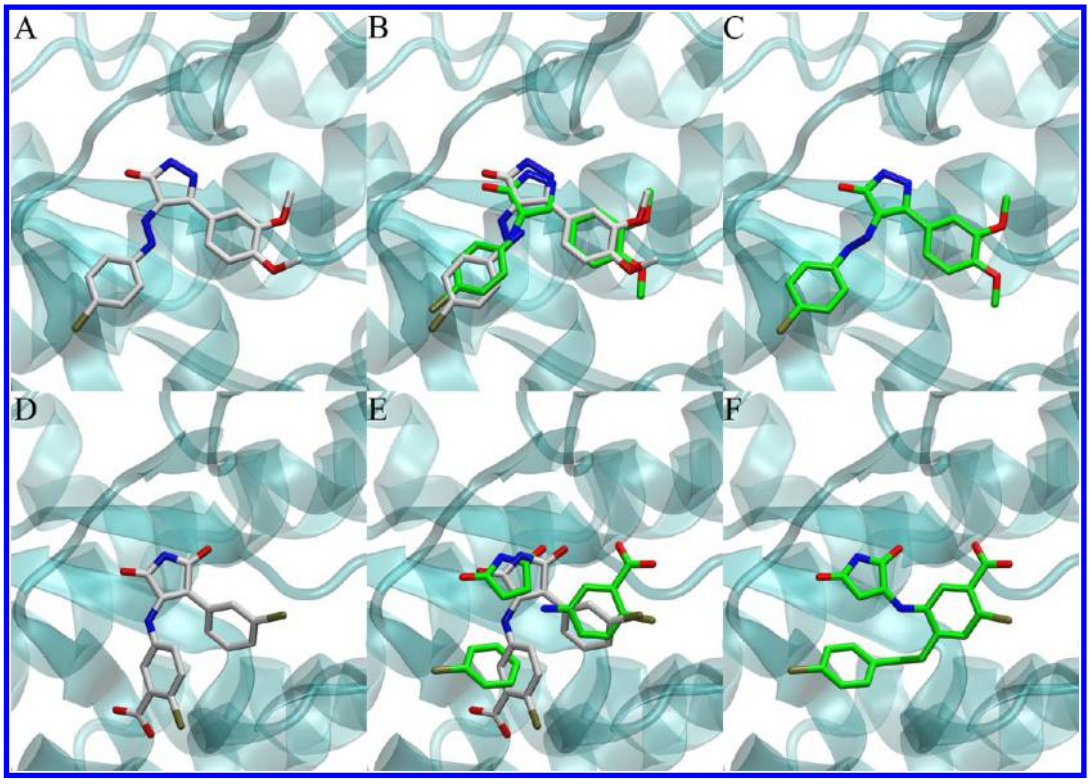


Figure 8. Dock-and-join process for PDB IDs 3L1S and 1Q4L (top and bottom panels, respectively). In A and D, the two leadlike compounds, as taken from the wwPDB, are shown. In B and E, energetically favorable docking solutions of the constituting fragments are shown along with the reference X-ray. In C and F, the resulting leads, as obtained through the joining procedure and subsequent minimization, are depicted. Proteins are highlighted in transparent cyan, and the native poses are highlighted C-colored in white, while the docked fragment poses as well as the newly built molecules are C-colored in green.

of GSK-3β bind in a defined cleft and some of their constituting parts could be solvent exposed and not directly involved in the binding. Moreover, functional groups of a lead could be present in X-ray structures for a number of different reasons, such as PK tuning or to improve relevant physicochemical features. However, from a broader drug design stand-point, having one small binder correctly predicted could be a very promising starting point toward a stepwise fragment growth.[31,60,61]

Construction of Leadlike Molecules. Fragments were joined according to criteria of proximity (see Methods for details) and the so-formed molecules were minimized within the protein frame to reach the closest energy minimum. At this stage, looking at the top-scoring reconstructed molecules, one of these three scenarios could be anticipated:

(i) The starting molecule was exactly rebuilt.
(ii) The joining procedure produced one or more molecules similar to the reference.
(iii) The joining procedure did not produce any molecule trivially resembling the starting lead.

The dock-and-join process along with two exemplary outcomes (PDB IDs 3L1S and 1Q4L) are reported in Figure 8.

In the case of protein 3L1S (top panels in Figure 8), the procedure originated the starting compound among the top scoring solutions. Conversely, in the case of protein 1Q4L, the protocol was not able to rebuilt the reference X-ray, and alternative solutions, such as the one depicted in Figure 8F, were created. In the latter case, lacking experimental evidence of inhibition or any clear similarity to the starting X-ray, any speculation on the molecule's activity would be arguable. To handle cases like these, we decided to analyze the results in two different ways, hereafter described.

A prompt and commonly used way to analyze molecular similarities is to calculate the Tanimoto[63] distance between fingerprints. Using the Daylight fingerprinting algorithm (Daylight Chemical Information Systems Inc., Laguna Niguel, CA), we compared the newly produced molecules to the reference structures. We found that approximately 85% of the time, we managed to reobtain the leadlike compounds at a Tanimoto distance lower than 0.2 from the reference. In 9 out of 21 cases, the starting compounds were exactly reproduced (PDB IDs 1J1B, 1PYX, 1Q41, 3L1S, 3ZRL, 1J1C, 1Q5K, 3DU8, and 3ZRK), whereas the protocol failed in just three cases (PDB IDs 3F88, 3F7Z, and 3I4B). In the remaining 9 cases (PDB IDs 3GB2, 1UV5, 2OW3, 1R0E, 2O5K, 3I4B, 3GB2, 3SD0, and 3Q3B), the X-ray structure was regenerated within a Tanimoto distance of 0.2.

However, a description based on fingerprints does not bear information on how the compound binds to the enzyme. Since a metrics based on rmsd could be sometimes misleading and could be exploited only in the case of perfect matches, we applied a similarity measure based on Tanimoto distances calculated between Lennard-Jones potentials (see Methods for details). In this way, we could assess not only whether reference binding modes were recapitulated but also whether molecules that apparently did not share any resemblance with the references retained the same interactions with the enzyme. Notably, this metric is similar in spirit to those proposed by others[39,64,65] in similar contexts. A detailed summary of the results is reported in Figure 9.

dx.doi.org/10.1021/ci300355p | J. Chem. Inf. Model. 2012, 52, 3233−3244

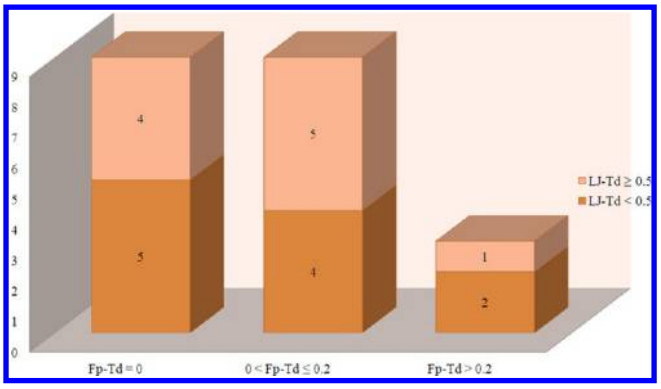


Figure 9. Summary of the joining procedure. The bars are concerned with reconstructed molecules that (i) matched exactly the starting X-ray or were at (ii) less or (iii) more than 0.2 fingerprint-Tanimoto distance from the reference (left, middle, and right, respectively). The bars color-code reflects the Lennard-Jones potentials Tanimoto distance from the reference.

Only 5 out of 9 molecules, previously considered perfect matches, recapitulated exactly the starting X-ray complexes (PDB IDs 1J1B, 1PYX, 1Q41, 3L1S, and 3ZRL). In the remaining four cases, even if the lead-like molecule was correctly rebuilt, its pose differed from the starting configuration by more than 3 Å rmsd. In the case of nonperfect matches (at varying fingerprint-Tanimoto distances from the references), the interactions with the enzyme were maintained in 6 out of 12 examined cases (PDB IDs 1Q4L, 1R0E, 1UV5, 2OW3, 3F88, and 3I4B).

The proposed computational protocol worked perfectly (i.e., the starting X-ray complexes were rightly matched) 5 out of 21 times. Nonetheless, in four cases, close analogs of the complexed inhibitors that re-enacted the key interactions observed in the experimentally determined structure were found (PDB IDs 1Q4L, 1R0E, 1UV5, and 2OW3). Three cases were particularly challenging, and no close analogs of the cocrystallized inhibitors could be generated (PDB IDs 3F7Z, 3F88, and 3I4B) even though two of them (PDB IDs 3F88 and 3I4B) retained the interaction fields of the crystal-native binder. Despite the topological dissimilarities, the protocol created molecules that overlapped nicely with the cocrystallized inhibitors, in terms of Lennard-Jones potentials as well as in terms of docking score. From a drug design standpoint, the latter represents a very interesting situation in which novel chemical entities, likely to retain the same interactions observed in potent inhibitors, can be unveiled.

Two exemplary cases of Lennard-Jones and fingerprint similarities are reported in Figure 10.

The molecule shown in panel A has worse topological overlap with the reference X-ray structure with respect to the molecules reported in panel B. However, the latter compound shows a significant decrease in the LJ overlap, mostly due to the inversion of the carboxylic group. The comparison of the two reported cases clearly depicts the usefulness of using an interaction-based protocol to evaluate the similarity between conformations of different compounds.

A final consideration concerns the estimated ligand efficiencies and binding affinity of both lead compounds and fragments employed in this study. The average docking score recorded for the lead molecules in the X-ray binding pose was −8.13 kcal/mol (SD = 1.4 kcal/mol), while the constituting fragments had an average estimated binding affinity of −5.11 kcal/mol (SD = 1.17 kcal/mol). Given the nature of the data set, this type of

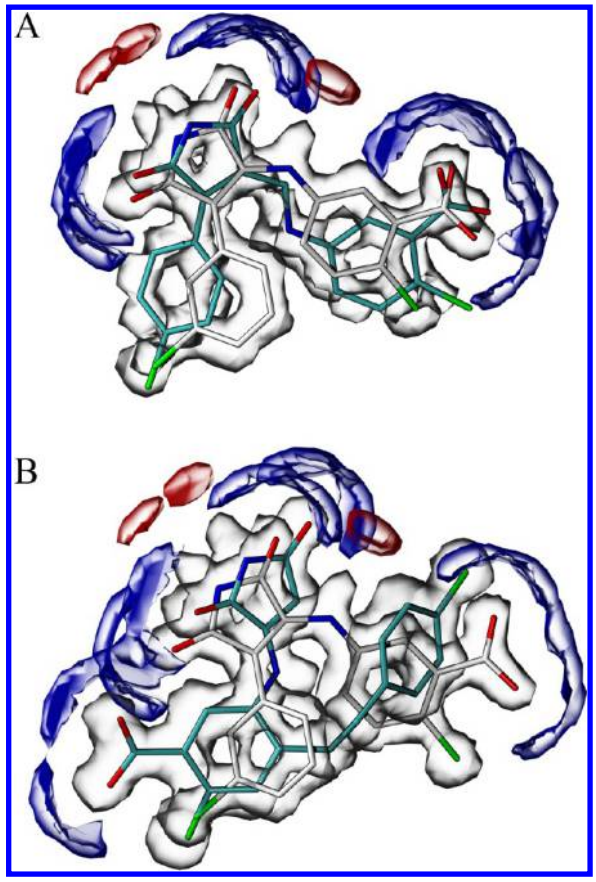


Figure 10. Two exemplary cases of molecules reconstructed compared to the reference compound (cmpID 679, PDB ID 1Q4L). In A, the newly generated compound and the target inhibitor share good topological and LJ overlaps (FP-Td 0.235, LJ-Td 0.351). In B, the novel compound, if compared with the one shown in A, has greater topological overlap with the reference (FP-Td 0.149). However, the LJ overlap is lower (LJ-Td 0.472). The C, H, and O probes are depicted as isocontour volumes gray, blue, and red colored, respectively. The reference compound, as extracted from the X-ray structure 1Q4L is C-colored in white, while the newly generated compounds are C-colored in cyan.

trend was expected being in line with what can be usually found experimentally when micromolar binders are evolved in nanomolar inhibitors. However, not every part of a nanomolar inhibitor is directly involved in interacting with the protein. In fact, fragment hits, although possessing worse absolute binding affinities, are interesting starting points, due to the fact that most of their constituting atoms are actively involved in binding. To this end, a parameter worth analyzing is the ligand efficiency (LE), calculated in this study as the estimated ΔG divided by the heavy atoms count. The LE speaks to the ability of molecules in using, at best, their atoms to displace solvent molecules from the binding site and engage the protein. Furthermore, LE is a useful metrics for normalizing data both when molecules of very different size are compared, or in the process of hit-to-lead, where the addition of small functional groups has to be evaluated. In our study, we found that the average LE of the lead-like compounds considered was −0.32 kcal/mol (SD = 0.07). On the other hand, if the most efficient fragments per case analyzed were considered, the LE was on average −0.8 kcal/mol (SD = 0.1). Due to all the limitations of scoring functions, those values are exceptionally high if compared to experimental outcomes. Nonetheless, this analysis reflects the fact that, during

dx.doi.org/10.1021/ci300355p | J. Chem. Inf. Model. 2012, 52, 3233−3244

fragment evolution, efficiency is likely to decrease in order to improve equally important physicochemical properties that ultimately translate into desirable PK profiles for sustained target engagement in vivo.

CONCLUSIONS

The study has shown that the reconstruction of nanomolar inhibitors of GSK-3β, starting from their constituting fragments, is possible. In 9 out of 21 analyzed cases, we were able to reobtain exactly the starting molecules, while in 9 other cases, close analogs were generated. Three cases were particularly challenging and resulted in the generation of molecules that did not resemble the starting inhibitors. An alternative view of the results can be gathered if, instead of looking at the topological similarities, one focuses on the interaction field similarities. In this way, one could pinpoint molecules that are chemically distant from the reference compounds and yet retain similar interactions. Regardless of their Tanimoto distance from the reference structure, 11 times we were able to generate molecules that shared worth analyzing 3D fields overlaping with the reference inhibitors. Even though the descriptors of interactions used here are clearly an oversimplification of what really goes on at a molecular level, in perspective, the present approach could allow one to overcome intellectual property issues by detecting unprecedented molecules that possess the same binding determinants of known modulators.

Because of their intrinsic simplicity, fragments are more prone to bind to proteins with low specificity.[23,56,62] This behavior is particularly appealing at the hit identification step and is among the reasons that have contributed to make FBDD so popular. However, because of this aspect, inferring the spatial arrangement of a fragment in complex with a macromolecular entity is far from a trivial task, since narrow energetic windows are likely to host a plethora of possible alternative solutions. Results from our computational docking reflect this observation and, in order to reconstruct the starting molecules, tens of alternative poses had to be generated. This was due to the fact that usually, not all the constituting parts of a binder are directly involved in interactions with the molecular target. Moreover in our case, an extra layer of complexity was given by the fact that GSK-3β binding site is a markedly solvent exposed cleft and the driving binding interactions were confined to the hinge region. As a consequence, we found a consistent correlation between success of docking and some physicochemical properties. Fragments bearing H-bond acceptor and donor groups were more likely to reproduce the X-ray crystal structures within the uppermost fraction of the top scoring solutions.

Upon lead fragmentation, hydrophobic interactions were the hardest to reproduce. Although results might not directly translate to other protein systems this could be due to the poor treatment of the entropic contribution to binding by computational algorithms. Moreover, it is interesting to note how this tendency has been already reported in some experimental studies where the specificity of enthalpy driven interaction versus the promiscuity of entropy driven interactions were analyzed.[23,55,62,66] Apart from the case-specific considerations, and analogously to some previous experimental techniques,[53,67−71] the proposed methodology might be used as an alternative cost-free way to generate complex molecules starting from simpler constituents.

Finally, the study reported here shows that, although varying degrees of success are possible, fragments preferred binding sites can be different from the site they occupy when they are included in larger molecules. Nonetheless, the potential to generate computationally novel chemical entities by addition of fragments preoptimized in a protein frame can be extremely valuable in finding empty slots even in crowded intellectual property spaces.

METHODS

Selection of PDB Structures. The wwPDB (as of February 10, 2012) was parsed to retrieve all the available crystal structures of wild-type human GSK-3β. Twenty-nine experimentally solved structures of wild-type GSK-3β were retrieved and analyzed. Apo structures were discarded (i.e., 1GNG, 1H8F, and 1I09) as well as structures complexed with organo-metallic compounds (i.e., 2JLD, 3M1S, and 3PUP) or structures bound to ligands that were unsuitable for the fragmentation and joining procedure employed in this study (i.e., 1Q3D and 1Q3W). These criteria led to the inclusion, in the final set, of 21 X-ray structures.

Workflow. The cocrystallized ligands were extracted and manually dissected in fragments. The final data set was composed of 57 commercially available fragments, representing 45 unique chemical compounds. Each protein was challenged, via docking, with the "native" fragment set. If compared with standard docking runs, the number of generated poses was purposely enhanced to 150, in order to allow us to draw some conclusions on the ease of docking in recapitulating the experimentally observed poses. The fragments were then joined, according to simple proximity criteria, thus generating bigger and more complex compounds. The resulting molecules were subjected to a local optimization in the protein frame to reach the closest local minimum.

The whole protocol was entirely based on modules available from the Schrödinger suite[72,75] of software. Once all the parameters were tuned, the workflow could be easily implemented in KNIME,[73] for easier data generation.

Protein Preparation. Each structure was subjected to the preparation steps implemented in the Protein Preparation Wizard of Maestro.[74] Only one molecular chain (chain A) was kept. Waters beyond 5 Å of the het-group were removed, bond orders were assigned, hydrogens were added, and metal atoms were treated automatically. Missing side chains were rebuilt with Prime.[75] The orientation of the amide (Asn, Gln), hydroxyl (Ser, Thr, Tyr), and thiol groups (Cys) and the protonation and tautomeric state of His residues were optimized. In the final step, all complexes were optimized by applying a stop criterion of 0.30 Å rmsd usingn the OPLS_2005 force field.[76] At the end, all waters and ions were deleted, and for each structure, a grid box of 26 Å × 26 Å × 26 Å with a default inner box (10 Å × 10 Å × 10 Å), was centered on the corresponding ligand.

Ligand Preparation. Each ligand was visually inspected and if necessary bond orders were adjusted, hydrogens were added, and protonation state was assigned with LigPrep,[77] according to physiological conditions. The fragmentation criteria were the following:

(i) Each fragment had to be constituted of, at least, 1 ring.
(ii) The terminal groups belonged to the attached rings.
(iii) The chemical bridges were assigned to one of the attaching rings.

This fragmentation scheme is similar, in spirit, to the one applied by Barelier and colleagues in one of their recent experimental works.[25]

Docking Protocol. Fragment docking calculations were performed with Glide SP,[78] applying flexible docking, by selecting expanded sampling which is very suitable for fragments. Nitrogen inversion or ring conformations sampling were excluded and only

trans conformations of amide bonds were allowed. The number of initial poses for ligand was increased to 50000, the best 700 poses were kept for energy minimization. As output options, we increased the number of docking poses to 150. The MM-GBSA[59] rescoring procedure was applied according to the flowing equation:

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} + \Delta G_{SA}$$

ΔE_{MM} is the difference in energy between the complex structure and the sum of the energies of the ligand and the protein, using the OPLS force field. ΔG_{solv} is the difference in the GBSA solvation energy of the complex and the sum of the solvation energies for the ligand and the protein. ΔG_{SA} is the difference in the surface area energy for the complex and the sum of the surface area energies for the ligand and the protein. Corrections for entropic changes were not applied.

To minimize the fragments in their original X-ray position, the *refine* option, as available in Glide SP precision mode, was employed.

Joining of Fragments and Refinement. After docking, fragment poses were joined with combine_fragments,[79] a python script from Schrödinger suite. In order to allow a robust sampling, we impose to reject only deviation of 25° for angles and 3 Å for bonds and distances between centers less than 2 Å. The connection points were hydrogen positions. All possible leadlike molecules were automatically sorted, and the 1000 best were saved.

After joining fragments, the newly born molecules were minimized, with a docking refinement in the protein frame, with Glide SP, optandscore option, increasing the number of minimization steps to 300, in order to reject molecules with steric clashes.

Lennard-Jones Potential Similarities. The methodology to devise similarity between proteins using Lennard-Jones potentials calculated at regularly spaced points within a box of interest is discussed in detail elsewhere.[52] The potential maps (spanning 14.25, 11.25, and 21 Å along the three axes, with grid spacing 0.375) on the prealigned structures were computed with Autogrid 4.2,[80] using carbon, hydrogen, and oxygen probes. The alignment was performed on the α carbon of the ATP binding site with Profit 3.1.[81] The grid was centered on the center of mass of the aligned ligands.

A similar methodology was applied to devise the ligand similarity. In this case, after the calculation of the grid points, the C, O, and H maps (spanning 30 Å in the three directions, with grid spacing 0.375) were merged in a single map files using these criteria:

(i) If the O LJ potential was lower or equal to −1 kcal/mol, a value of 1.00 was assigned to that point.
(ii) If the H LJ potential was lower or equal to −0.6 kcal/mol, a value of −1.00 was assigned to that point.
(iii) If the C LJ potential was lower or equal to 0 kcal/mol, a value of 0.00 was assigned to that point.
(iv) If none of the above conditions was fulfilled, that grid point had a "non-assigned" (ND) value.

In doing so, a net compression of the data was possible, yet the relevant information was retained and conveniently encoded into a single, ternary vector. The threshold values above-reported were heuristically defined. The map discretization allowed us to apply a Tanimoto-like metric to evaluate similarities. The metric was adapted in this way in order to account for the presence of four possible values (1.00, 0.00, −1.00, and ND):

(i) The bit was common if the values were equal.
(ii) The number of common bits rose by 0.5 if, in a given point, the two maps had a C and O preference or a C and H preference.
(iii) The number of common bits was decreased by 0.5 if, in a given point, the two maps had a O and H preference.

These numbers were set so to softly treat situations in which the discretized maps did not perfectly match or to penalize situations in which H bond acceptor and donor propensities clearly clashed.

AUTHOR INFORMATION

Corresponding Author

*Phone: +86-21-20805557. Fax: +86-21-20805501. E-mail: favia_angelo@lilly.com.

Present Address

⊥Lilly China R&D Center, Building 8, 338 Jia Li Lue Road, Zhangjiang Hi-Tech Park, Pudong, 201203, Shanghai (CN).

Notes

The authors declare no competing financial interest.

○A.D.F. and A.C. should be regarded as joint last authors.

ACKNOWLEDGMENTS

We thank the IIT platform "Computation" for computational resources.

ABBREVIATIONS

FB, fragment-based; HB, hydrogen bond; HTS, high-throughput screening; FBDD, fragment-based drug design; NMR, nuclear magnetic resonance; SPR, surface plasmon resonance; ITC, isothermal titration calorimetry; GSK-3β, glygogen synthase kinase-3β; UPGMA, Unweighted Pair Group Method with Arithmetic Mean; wwPDB, Worldwide Protein Data Bank; LE, ligand efficiency; LJ, Lennard-Jones

REFERENCES

(1) Warr, W. A. Fragment-based drug discovery. *J. Comput.-Aided Mol. Des.* **2009**, *23* (8), 453−458.

(2) Whittaker, M.; Law, R. J.; Ichihara, O.; Hesterkamp, T.; Hallett, D. Fragments: past, present and future. *Drug Discovery Today: Technol.* **2010**, *7* (3), e163−e171.

(3) Tse, C.; Shoemaker, A. R.; Adickes, J.; Anderson, M. G.; Chen, J.; Jin, S.; Johnson, E. F.; Marsh, K. C.; Mitten, M. J.; Nimmer, P.; Roberts, L.; Tahir, S. K.; Mao, Y.; Yang, X. F.; Zhang, H. C.; Fesik, S.; Rosenberg, S. H.; Elmore, S. W. ABT-263: A potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res.* **2008**, *68* (9), 3421−3428.

(4) Ackler, S.; Mitten, M.; Foster, K.; Oleksijew, A.; Refici, M.; Tahir, S.; Xiao, Y.; Tse, C.; Frost, D.; Fesik, S.; Rosenberg, S.; Elmore, S.; Shoemaker, A. The Bcl-2 inhibitor ABT-263 enhances the response of multiple chemotherapeutic regimens in hematologic tumors in vivo. *Cancer Chemother. Pharm.* **2010**, *66* (5), 869−880.

(5) Santo, L.; Vallet, S.; Hideshima, T.; Cirstea, D.; Ikeda, H.; Pozzi, S.; Patel, K.; Okawa, Y.; Gorgun, G.; Perrone, G.; Calabrese, E.; Yule, M.; Squires, M.; Ladetto, M.; Boccadoro, M.; Richardson, P. G.; Munshi, N. C.; Anderson, K. C.; Raje, N. AT7519, A novel small molecule multi-cyclin-dependent kinase inhibitor, induces apoptosis in multiple myeloma via GSK-3 beta activation and RNA polymerase II inhibition. *Oncogene* **2010**, *29* (16), 2325−2336.

(6) Santo, L.; Vallet, S.; Hideshima, T.; Cirstea, D.; Pozzi, S.; Vaghela, N.; Ikeda, H.; Okawa, Y.; Gorgun, G.; Perrone, G.; Calabrese, E.; Squires, M. S.; Ladetto, M.; Munshi, N. C.; Boccadoro, M.; Anderson, K. C.; Raje, N. AT7519, a Novel Small Molecule Multi-Cyclin Dependent Kinase Inhibitor, Induces Apoptosis in Multiple Myeloma VIA GSK3 beta. *Blood* **2008**, *112* (11), 99−100.

(7) Squires, M. S.; Cooke, L.; Lock, V.; Qi, W. Q.; Lewis, E. J.; Thompson, N. T.; Lyons, J. F.; Mahadevan, D. AT7519, a Cyclin-Dependent Kinase Inhibitor, Exerts Its Effects by Transcriptional Inhibition in Leukemia Cell Lines and Patient Samples. *Mol. Cancer Ther.* **2010**, *9* (4), 920−928.
(8) Agnelli, G.; Haas, S.; Ginsberg, J. S.; Krueger, K. A.; Dmitrienko, A.; Brandt, J. T. A phase II study of the oral factor Xa inhibitor LY517717 for the prevention of venous thromboembolism after hip or knee replacement. *J. Thromb. Haemostasis* **2007**, *5* (4), 746−753.
(9) Talele, T. T.; Khedkar, S. A.; Rigby, A. C. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. *Curr. Top. Med. Chem.* **2010**, *10* (1), 127−141.
(10) Artis, D. R.; Lin, J. J.; Zhang, C.; Wang, W.; Mehra, U.; Perreault, M.; Erbe, D.; Krupka, H. I.; England, B. P.; Arnold, J.; Plotnikov, A. N.; Marimuthu, A.; Nguyen, H.; Will, S.; Signaevsky, 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.; 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* (1), 262−267.
(11) Dominguez, J. M.; Fuertes, A.; Orozco, L.; del Monte-Millan, M.; Delgado, E.; Medina, M. Evidence for irreversible inhibition of glycogen synthase kinase-3beta by tideglusib. *J. Biol. Chem.* **2012**, *287* (2), 893−904.
(12) 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. Fragment-based drug discovery applied to Hsp90. Discovery of two lead series with high ligand efficiency. *J. Med. Chem.* **2010**, *53* (16), 5942−5955.
(13) Murray, C. W.; Callaghan, O.; Chessari, G.; Cleasby, A.; Congreve, M.; Frederickson, M.; Hartshorn, M. J.; McMenamin, R.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to beta-secretase. *J. Med. Chem.* **2007**, *50* (6), 1116−1123.
(14) Jacobsen, J. A.; Fullagar, J. L.; Miller, M. T.; Cohen, S. M. Identifying chelators for metalloprotein inhibitors using a fragment-based approach. *J. Med. Chem.* **2011**, *54* (2), 591−602.
(15) Geitmann, M.; Elinder, M.; Seeger, C.; Brandt, P.; de Esch, I. J. P.; Danielson, U. H. Identification of a Novel Scaffold for Allosteric Inhibition of Wild Type and Drug Resistant HIV-1 Reverse Transcriptase by Fragment Library Screening. *J. Med. Chem.* **2011**, *54* (3), 699−708.
(16) 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* (16), 5836−5857.
(17) 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* (7), 3285−3306.
(18) Bastien, D.; Ebert, M. C.; Forge, D.; Toulouse, J.; Kadnikova, N.; Perron, F.; Mayence, A.; Huang, T. L.; Vanden Eynde, J. J.; Pelletier, J. N. Fragment-based design of symmetrical bis-benzimidazoles as selective inhibitors of the trimethoprim-resistant, type II R67 dihydrofolate reductase. *J. Med. Chem.* **2012**, *55* (7), 3182−3192.
(19) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274* (5292), 1531−1534.
(20) Roughley, S. D.; Hubbard, R. E. How Well Can Fragments Explore Accessed Chemical Space? A Case Study from Heat Shock Protein 90. *J. Med. Chem.* **2011**, *54* (12), 3989−4005.
(21) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent developments in fragment-based drug discovery. *J. Med. Chem.* **2008**, *51* (13), 3661−3680.
(22) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug. Discovery* **2007**, *6* (3), 211−219.
(23) Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41* (3), 856−864.
(24) Hajduk, P. J. Fragment-based drug design: how big is too big? *J. Med. Chem.* **2006**, *49* (24), 6972−6976.
(25) Barelier, S.; Pons, J.; Marcillat, O.; Lancelin, J. M.; Krimm, I. Fragment-Based Deconstruction of Bcl-x(L) Inhibitors. *J. Med. Chem.* **2010**, *53* (6), 2577−2588.
(26) Babaoglu, K.; Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2006**, *2* (12), 720−723.
(27) Carr, R. A.; Congreve, M.; Murray, C. W.; Rees, D. C. Fragment-based lead discovery: leads by design. *Drug Discovery Today* **2005**, *10* (14), 987−992.
(28) Gozalbes, R.; Carbajo, R. J.; Pineda-Lucena, A. Contributions of computational chemistry and biophysical techniques to fragment-based drug discovery. *Curr. Med. Chem.* **2010**, *17* (17), 1769−1794.
(29) Retra, K.; Irth, H.; van Muijlwijk-Koezen, J. E. Surface Plasmon Resonance biosensor analysis as a useful tool in FBDD. *Drug Discovery Today: Technol.* **2010**, *7* (3), e181−e187.
(30) Edink, E.; Jansen, C.; Leurs, R.; de Esch, I. J. P. The heat is on: thermodynamic analysis in fragment-based drug discovery. *Drug Discovery Today: Technol.* **2010**, *7* (3), e189−e201.
(31) 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.* **2011**, *55* (2), 576−586.
(32) Pellecchia, M. Fragment-based drug discovery takes a virtual turn. *Nat. Chem. Biol.* **2009**, *5* (5), 274−275.
(33) Zoete, V.; Grosdidier, A.; Michielin, O. Docking, virtual high throughput screening and in silico fragment-based drug design. *J. Cell. Mol. Med.* **2009**, *13* (2), 238−248.
(34) Chen, Y.; Pohlhaus, D. T. In silico docking and scoring of fragments. *Drug Discovery Today: Technol.* **2010**, *7* (3), e149−e156.
(35) Yuan, H.; Lu, T.; Ran, T.; Liu, H.; Lu, S.; Tai, W.; Leng, Y.; Zhang, W.; Wang, J.; Chen, Y. Novel Strategy for Three-Dimensional Fragment-Based Lead Discovery. *J. Chem. Inf. Model.* **2011**, *51* (4), 959−974.
(36) Li, H.; Liu, A.; Zhao, Z.; Xu, Y.; Lin, J.; Jou, D.; Li, C. Fragment-Based Drug Design and Drug Repositioning Using Multiple Ligand Simultaneous Docking (MLSD): Identifying Celecoxib and Template Compounds as Novel Inhibitors of Signal Transducer and Activator of Transcription 3 (STAT3). *J. Med. Chem.* **2011**, *54* (15), 5592−5596.
(37) Lin, F. Y.; Tseng, Y. J. Structure-based fragment hopping for lead optimization using predocked fragment database. *J. Chem. Inf. Model.* **2011**, *51* (7), 1703−1715.
(38) Bollini, M.; Domaoal, R. A.; Thakur, V. V.; Gallardo-Macias, R.; Spasov, K. A.; Anderson, K. S.; Jorgensen, W. L. Computationally-Guided Optimization of a Docking Hit to Yield Catechol Diethers as Potent Anti-HIV Agents. *J. Med. Chem.* **2011**, *54* (24), 8582−8591.
(39) Marcou, G.; Rognan, D. Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195−207.
(40) Verdonk, M. L.; Giangreco, I.; Hall, R. J.; Korb, O.; Mortenson, P. N.; Murray, C. W. Docking Performance of Fragments and Druglike Compounds. *J. Med. Chem.* **2011**, *54* (15), 5422−5431.
(41) Andersen, O. A.; Nathubhai, A.; Dixon, M. J.; Eggleston, I. M.; van Aalten, D. M. Structure-based dissection of the natural product cyclopentapeptide Chitinase inhibitor argifin. *Chem. Biol.* **2008**, *15* (3), 295−301.

(42) Brandt, P.; Geitmann, M.; Danielson, U. H. Deconstruction of Non-Nucleoside Reverse Transcriptase Inhibitors of Human Immunodeficiency Virus Type 1 for Exploration of the Optimization Landscape of Fragments. *J. Med. Chem.* **2011**, *54* (3), 109−118.
(43) 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* (10), 761−772.
(44) Ali, A.; Hoeflich, K. P.; Woodgett, J. R. Glycogen synthase kinase-3: properties, functions, and regulation. *Chem. Rev.* **2001**, *101* (8), 2527−2540.
(45) Hernandez, F.; Avila, J. The role of glycogen synthase kinase 3 in the early stages of Alzheimers' disease. *FEBS Lett.* **2008**, *582* (28), 3848−3854.
(46) Hernandez, F.; de Barreda, E. G.; Fuster-Matanzo, A.; Goni-Oliver, P.; Lucas, J. J.; Avila, J. The role of GSK3 in Alzheimer disease. *Brain Res. Bull.* **2009**, *80* (4−5), 248−250.
(47) Ryu, Y. K.; Lee, Y. S.; Lee, G. H.; Song, K. S.; Kim, Y. S.; Moon, E. Y. Regulation of glycogen synthase kinase-3 by thymosin beta-4 is associated with gastric cancer cell migration. *Int. J. Cancer* **2012**, *131* (9), 2067−2077.
(48) Tai, W.; Shi, E.; Yan, L.; Jiang, X.; Ma, H.; Ai, C. Diabetes abolishes the cardioprotection induced by sevoflurane postconditioning in the rat heart in vivo: Roles of glycogen synthase kinase-3 beta and its upstream pathways. *J. Surg. Res.* **2012**, *178*, 96−104.
(49) Bilim, V.; Ougolkov, A.; Yuuki, K.; Naito, S.; Kawazoe, H.; Muto, A.; Oya, M.; Billadeau, D.; Motoyama, T.; Tomita, Y. Glycogen synthase kinase-3: a new therapeutic target in renal cell carcinoma. *Br. J. Cancer* **2009**, *101* (12), 2005−2014.
(50) Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* **2007**, *35* (Database issue), D301−D303.
(51) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33−38 ; see also pages 27 and 28.
(52) Bicego, M.; Favia, A. D.; Bisignano, P.; Cavalli, A.; Murino, V. An innovative protocol for comparing protein binding sites via atomic grid maps. *Proc. Int. Conf. Knowl. Discov. Inf. Retriev.* **2011**, 412−422.
(53) Núñez, L. E.; Nybo, S. E.; Gonzalez-Sabin, J.; Pérez, M.; Menéndez, N.; Braña, A. F.; He, M.; Morís, F.; Salas, J. A.; Rohr, J.; Méndez, C. A novel mithramycin analogue with high antitumor activity and less toxicity generated by combinatorial biosynthesis. *J. Med. Chem.* **2012**, *55* (12), 5813−5825.
(54) *Daylight*, Chemical Information Systems Inc., Laguna Niguel, CA, 2008.
(55) Favia, A. D.; Bottegoni, G.; Nobeli, I.; Bisignano, P.; Cavalli, A. SERAPhiC: A Benchmark for in Silico Fragment-Based Drug Design. *J. Chem. Inf. Model.* **2011**, *51* (11), 2882−2896.
(56) Bottegoni, G.; Favia, A. D.; Recanatini, M.; Cavalli, A. The role of fragment-based and computational methods in polypharmacology. *Drug Discovery Today* **2012**, *17* (1−2), 23−34.
(57) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J. Comput. Chem.* **2011**, *32* (5), 866−877.
(58) Genheden, S.; Ryde, U. Comparison of the Efficiency of the LIE and MM/GBSA Methods to Calculate Ligand-Binding Energies. *J. Chem. Theory Comput.* **2011**, *7* (11), 3768−3778.
(59) Haider, M. K.; Bertrand, H. O.; Hubbard, R. E. Predicting Fragment Binding Poses Using a Combined MCSS MM-GBSA Approach. *J. Chem. Inf. Model.* **2011**, *51* (5), 1092−1105.
(60) Buckley, D. L.; Van Molle, I.; Gareiss, P. C.; Tae, H. S.; Michel, J.; Noblin, D. J.; Jorgensen, W. L.; Ciulli, A.; Crews, C. M. Targeting the von Hippel-Lindau E3 Ubiquitin Ligase Using Small Molecules To Disrupt the VHL/HIF-1 alpha Interaction. *J. Am. Chem. Soc.* **2012**, *134* (10), 4465−4468.
(61) Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, M. D.; Chung, C. W. Fragment-Based Discovery of Bromodomain Inhibitors Part 2: Optimization of Phenylisoxazole Sulfonamides. *J. Med. Chem.* **2012**, *55* (2), 587−596.
(62) Nobeli, I.; Favia, A. D.; Thornton, J. M. Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol.* **2009**, *27* (2), 157−167.
(63) Tanimoto, T. T. *IBM Internal Report*, IBM: Armonk, NY, Nov 17, 1957.
(64) Deng, Z.; Chuaqui, C.; Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J. Med. Chem.* **2004**, *47* (2), 337−344.
(65) Brewerton, S. C. The use of protein-ligand interaction fingerprints in docking. *Curr. Opin. Drug Discovery Dev.* **2008**, *11* (3), 356−364.
(66) Morphy, R.; Rankovic, Z. Fragments, network biology and designing multiple ligands. *Drug Discovery Today* **2007**, *12* (3−4), 156−160.
(67) Kapkovà, P.; Heller, E.; Unger, M.; Folkers, G.; Holzgrabe, U. Random Chemistry as a New Tool for the Generation of Small Compound Libraries: Development of a New Acetylcholinesterase Inhibitor. *J. Med. Chem.* **2005**, *48* (23), 7496−7499.
(68) Dolle, R. E.; Bourdonnec, B. L.; Worm, K.; Morales, G. A.; Thomas, C. J.; Zhang, W. Comprehensive survey of chemical libraries for drug discovery and chemical biology: 2009. *J. Comb. Chem.* **2010**, *12* (6), 765−806.
(69) Dolle, R. E.; Le Bourdonnec, B.; Goodman, A. J.; Morales, G. A.; Thomas, C. J.; Zhang, W. Comprehensive survey of chemical libraries for drug discovery and chemical biology: 2008. *J. Comb. Chem.* **2009**, *11* (5), 739−90.
(70) Khan, M.; Wate, P.; Krupadam, R. Combinatorial screening of polymer precursors for preparation of benzo[α] pyrene imprinted polymer: an ab initio computational approach. *J. Mol. Model.* **2012**, *18* (5), 1969−1981.
(71) Minovski, N.; Perdih, A.; Solmajer, T. Combinatorially-generated library of 6-fluoroquinolone analogs as potential novel antitubercular agents: a chemometric and molecular modeling assessment. *J. Mol. Model.* **2012**, *18* (5), 1735−1753.
(72) (a) *Schrödinger Suite 2010 Protein Preparation Wizard, Epik*, version 2.1; Schrödinger, LLC: New York, 2010. (b) *Impact*, version 5.6; Schrödinger, LLC: New York, 2010.
(73) *KNIME*, version 2.1.2; The Chair for Bioinformatics and Information Mining, University of Konstanz, Germany, 2010.
(74) *Maestro*, version 9.1; Schrödinger, LLC, New York, 2010.
(75) *Prime*, version 2.2; Schrödinger, LLC, New York, 2010.
(76) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B.* **2001**, *105* (28), 6474−6487.
(77) *LigPrep*, version 2.4; Schrödinger, LLC, New York, 2010.
(78) *Glide*, version 5.6; Schrödinger, LLC, New York, 2010.
(79) *combine_fragment*; Schrödinger, L.L.C.: New York, 2010.
(80) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785−2791.
(81) (a) Martin, A. C. R. and Porter, C. T. *ProFit*, version 3.1; SciTech Software, University College London, London, 2009. (b) McLachlan, A. D. Rapid comparison of protein structures. *Acta Crystallogr., Sect. A* **1982**, *38*, 871−873.