JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

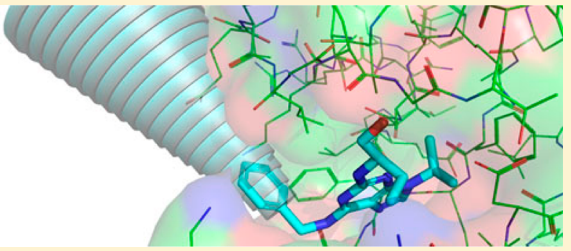
CSBB-ConeExclusion, Adapting Structure Based Solution Virtual Screening to Libraries on Solid Support

Steven Shave* and Manfred Auer*

School of Biological Sciences and School of Biomedical Sciences, University of Edinburgh, The King's Buildings, CH Waddington Building, Mayfield Road, Edinburgh, Scotland EH9 3JD, U.K.

Supporting Information

ABSTRACT: Combinatorial chemical libraries produced on solid support offer fast and cost-effective access to a large number of unique compounds. If such libraries are screened directly on-bead, the speed at which chemical space can be explored by chemists is much greater than that addressable using solution based synthesis and screening methods. Solution based screening has a large supporting body of software such as structure-based virtual screening tools which enable the prediction of protein−ligand complexes. Use of these techniques to predict the protein bound complexes of compounds synthesized on solid support neglects to take into account the conjugation site on the small molecule ligand. This may invalidate predicted binding modes, the linker may be clashing with protein atoms. We present CSBB-ConeExclusion, a methodology and computer program which provides a measure of the applicability of solution dockings to solid support. Output is given in the form of statistics for each docking pose, a unique 2D visualization method which can be used to determine applicability at a glance, and automatically generated PyMol scripts allowing visualization of protein atom incursion into a defined exclusion volume. CSBB-ConeExclusion is then exemplarically used to determine the optimum attachment point for a purine library targeting cyclin-dependent kinase 2 CDK2.



INTRODUCTION

Virtual screening efforts undertaken in academic drug discovery laboratories are hindered greatly by restrictions on making chemical space experimentally accessible based on *in silico* results. Starting from SAR (Structural Activity Relationship), industry routinely places large teams of chemists on the experimental exploration around virtual hit compound series. By the very nature of funding schemes, academic groups have much less capacity and fewer members who are often specializing in distinct areas of a discipline central to the group focus. This results in massive restrictions on target specific chemical output and places academic groups at the mercy of, often commercial, screening compound suppliers for hit diversification and SAR exploration. Once a hit is identified, the two routes to explore SAR are most of the time based on commercially available similars or limited chemical derivatization projects. The Tagged One-Bead One Compound (TOBOC)[1,2] technology in place at the University of Edinburgh attempts to address this disparity using a miniaturized high throughput synthesis and screening platform to access large areas of chemical space through both diversity and combinatorial scale in a fully quantitative and quality controlled process. Using combinatorial library synthesis methods such as split and mix[3,4] on solid support, large libraries around core scaffolds can be rapidly synthesized with the inclusion of useful chemical moieties or handles such as a variety of tagging sites and spacers. Inclusion of these elements to present a chemical entity enables use of the bead as both a synthesis compartment and primary screening compartment at the single bead and single molecule level. Coupled with a cleavable linker, compounds can be removed from solid support for screening in solution, in cells, and in model organisms, for affinity determination to targets, for purity check, and for structural decoding. We refer to the critical importance of determining the three parameters, compound purity, ligand binding affinity (K_d), and chemical structure, from the amount of substance produced on one microbead, e.g. from ∼50−100 picomoles contained on one 90 μm bead as "rule of 3 of on-bead screening". Following this "rule of 3" removes the need for time, personnel, and cost consuming resynthesis efforts too early in the hit compound discovery process.

In our hands, screening on-bead typically takes the form of bead incubation with fluorescently labeled target protein, confocal nanoscanning (CONA), and hit bead retrieval. A hit bead being defined by appearing with a fluorescent ring in the confocal image by virtue of having fluorescently labeled protein bound to the outer 2 μm of the bead "shell". Isolated hit beads are processed by labeling of the compound with a fluorophore at a generic tagging site and cleavage, after which 50% of the ∼50−100 picomoles of substance load is used for a HPLC purity check, followed by affinity determination in solution via detection techniques based on confocal fluorescence analysis at single molecule resolution. A third portion is taken forward to be decoded via mass spectrometric methods in order to reveal the identity of the solution confirmed binder.

Received: June 27, 2013
Published: November 17, 2013

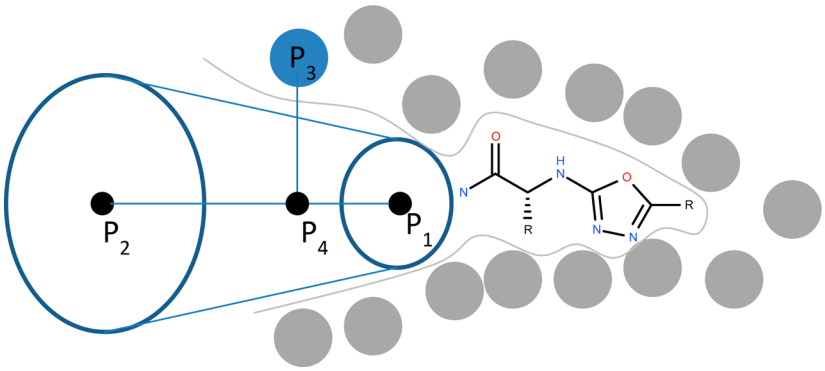


Figure 1. Cone exclusion volume defined by attachment point P_1 and P_2, specified by P_1, the length of the cone and the position origination from the connected atoms of the ligand. The closest point within the vector defining the center of the cone to the protein atom P_3 is defined as P_4.

Focusing on libraries containing core scaffolds deemed "privileged"[5−7] against a target greatly increases the chances of finding small molecule binders when compared with random or blind screening of libraries against a target. Virtual screening is a wide field broadly divisible by the type of data utilized by the technique in an attempt to achieve an enrichment factor greater than that of random screening or chance alone. For this reason, augmentation of combinatorial planning and library design through any method which offers an enrichment rate, the ratio of confirmed actives to screened compounds, over random chance proves extremely desirable. Virtual screening can be split into structure-based[8−12] and ligand-based[13,14] techniques. Virtual screening classically refers to structure-based virtual screening where knowledge of the receptor guides a docking/posing and scoring algorithm, creating a model of the predicted protein−ligand binding event being simulated and returning a score rating the predicted interaction energy. In contrast, ligand-based techniques rely on the molecular similarity principle, a quantified idea[15−18] which simply states that two "similar" molecules should be capable of making the same or similar interactions and therefore bind to common binding sites. While a plethora of virtual screening tools enabling the simulation of small molecules in solution binding to a protein are available, virtual screening tools, both structure and ligand-based, are severely lacking when we consider compounds on solid support. CSBB-ConeExclusion has been developed as a postdocking filter applicable to every structure-based virtual screening code that generates a protein−ligand complex among its output. The use of CSBB-ConeExclusion adapts these screening tools from solution to solid support. To achieve this, we ask a very simple question − "If the compound were attached to solid support via one of the possible chemical conjugation sites, e.g. position 2, 6, 8, or 9 on a purine scaffold, would the docking be valid?". Second, we address the question of the certainty with which we can say that the chemical linker does not interfere with binding. In developing an automated protocol for rapid and repeated answering of these questions we define the geometrical and dynamic characteristics of a linker attached between a bead and a compound bound to protein in order to construct a descriptive mathematical model. The frame of reference within which we work is defined by the complex. With the protein relatively static and compound docked, flexibility and variability is only present in the linker portion of the compound. As we move away from the fixed protein−ligand binding site, along the spacer attachment point the position of atoms becomes evermore uncertain. This distribution can be thought of as a cone of uncertainty.

METHODS

To illustrate the practical steps required in cone exclusion volume filtering, we consider the situation where substructure searching has been applied to a docked molecule to identify a common linker i.e. solid support attachment point which is then used to define an orientation for the cone exclusion volume. Coupled with an attachment atom, we have an orientation and displacement within the 3D space defined by the docking result with which to work. We now define the dimensions of the cone. As previously mentioned, this cone is in effect a cone of uncertainty (relating to linker atom positions as we move away from the binding site). These dimensions can be defined arbitrarily, through trial and error, or by using empirical methods. For algorithm demonstration purposes, we define the cone to be of the following dimensions $l = 10$, $d_1 = 2$, $d_2 = 10$ where l is the length of the cone, and d_1 is the diameter of the base of the cone: this is the thin end, closest to the attachment atom of the scaffold and is arbitrarily assigned a value of 2 Å. A logical assignment for this value would be the radius of a nitrogen (1.55 Å) or carbon atom (1.7 Å); however, the algorithm is rather insensitive to this value when of the order of single digit angstroms, and therefore a simple value of 2 Å was chosen. Further parametrization choices were taken from the results of a parameter sweeping exercise documented later. d_2 is the diameter at the end of the cone. This is the thickest end, furthest away from the attachment atom. We then define two points, P_1 and P_2. P_1 takes the coordinates of the linker and support attachment atom, while P_2 takes the coordinates of P_1 scaled with the unit vector defining the direction and a constant denoting the desired length of cone l. This vector is determined by examination of the core scaffold to which the linker is attached. In the case of attachment through an amide bond, the directional vector can be defined as the vector from the carbonyl carbon atom to the nitrogen atom. A subtly different case is shown in Figure 3, where we define the simulation of an attachment point at the para-position of a ring. In this case, it was necessary to define P_1 as position 1 of the ring, and the vector as the atom connected to the ring position 1 atom, to the ring position 1 atom itself. This is all achieved in an automated manner with substructure searching. With this vector and P_1, we can calculate P_2, that is $P_2 = P_1 l\hat{v}$. We then step through the following process: for each protein atom, P_3, find the point, designated as P_4, along the vector defined by $\overrightarrow{P_1P_2}$ which minimizes the distance between P_3 and P_4 (see Figure 1).

Finding point P_4 becomes trivial given that the shortest distance between a point in space and a point residing on a line

dx.doi.org/10.1021/ci400371q | J. Chem. Inf. Model. 2013, 53, 3156−3162

can be defined by another vector perpendicular to the first. See the Supporting Information.

A value is then obtained to define a distance cutoff outside of which the separation between P_3 and P_4 must lie for the atom to be considered outside of the exclusion volume defined by the cone. See the Supporting Information for formulas. This value is used to scale the diameter of the cone at the small end smoothly to the diameter of the cone at the larger end and obtain a distance, beyond which a point is deemed to no longer be within the cone volume. We define this value as k and use it to define a cutoff distance within which a protein atom would lie within the exclusion cone volume (note that we take half the cone diameters in order to work with cone radii): $d_{cutoff,k} = (d_1)/2 + ((d_2)/2 - (d_1)/2)k$. We then calculate P_4; $P_4 = k(P_2 - P_1)$. Knowing P_4, we determine the length of $\overrightarrow{P_4P_3}$ and test to see if the atom lies within the cone exclusion volume ($\overrightarrow{P_4P_3} \leq d_{cutoff,k}$).

Using the testing scheme and mathematical model outlined above to define the exclusion volume, the ends of the cone take on a rounded shape (rather than a flat surface or cap in a traditional cone) with a radius equal to half the diameter specified for each end of the cone. A flat ended cone may be achieved by clamping the calculated value of k to lie between 0 and 1.

This technique has been taken to its logical conclusion with a practical implementation creating a program named CSBB-ConeExclusion, written in C++ and making use of the OpenBabel API[19,20] for substructure searching, reading, and manipulating molecular information, along with the CImg[21] library for image writing. CSBB-ConeExclusion provides output of protein atom incursion into defined cone exclusion volumes in three ways, producing the following: 1) A comma separated file containing statistics on cone occupancy for a docking pose. 2) A series of PNG images giving a 2D overview of protein atom encroachment into the exclusion volume. 3) A PyMol[22] (Open source) script is generated, allowing visualization of the cone exclusion volume along with protein and docked poses.

2D Exclusion Volume Plots. To aid in rapid visual inspection of a docking run and to identify the solution poses amenable to solid support, 2D exclusion volume plots are generated for each docking result tested with the algorithm. With the cone exclusion volume defined in the model detailed above, protein atoms encroaching within the volume are noted and kept in a list with their associated k values as a measure of their position along the length of the cone. We deem protein atoms located at the thin end of the cone next to the ligand-linker attachment point to be more important in determining solid support suitability of a specific scaffold linkage, compared to atoms located further along the cone. Therefore, we must heavily emphasize their importance in a visual representation of the solution poses amenability to solid support. This is achieved by scaling a circle representing the atom by the cone diameter at the associated k value. In addition, we scale the color of circles representing atoms from red, $k = 0$, protein atom close to small end of cone and ligand, to orange, $k = 1$, protein atom at large end of cone. Encroaching protein atoms are then rendered in order of increasing k values (small end of the cone to large end) onto a circle representing a view into the cone from the large end, looking down toward the ligand at the small end. The scaling of circle sizes as a function of cone cross-sectional area adds importance to critical atoms at the small end of the cone, occupying a larger proportion of cone cross-sectional area. This counterintuitive viewing style with a reversal of classical perspective correction emphasizes the impact of atoms encroaching within the thin end of the exclusion volume, rendering them as large circles and encroaching atoms at the thick end of the cone, close to the conceptual viewer's eye, as small circles. Rather than relying upon an arbitrary positioning of the viewer's orientation around the axis defined by $\overrightarrow{P_1P_2}$, we define the angle, $\theta_0 = 0$, at the position of the first encroaching atom (defined as E_1) around the cone axes, in order of increasing k values. With $\theta_0 = 0$ defined, we can translate the coordinates of E_1 along the $\overrightarrow{P_1P_2}$ axes, adapting the circular orientation of E_1 for every k value. With the translation of E_1 along the axis, the calculated P_4 center point of the cone (see Figure 1) becomes the closest point along the vector $\overrightarrow{P_1P_2}$ to E_1. P_4 is therefore located at the same k value as subsequent encroaching protein atoms, defined as E_n. The angle θ_E between E_1P_4 and E_n in any cross section of the cone can then be calculated as follows:

$$\theta_E = \cos^{-1}\left(\frac{(E_1 - P_4)\cdot(E_n - P_4)}{\| E_1 - P_4\| \times \| E_n - P_4\|}\right)$$

Storing each encroaching atoms position around a central point with the axis defined by $\overrightarrow{P_1P_2}$ allows us to radially plot atom positions as follows:

$$E_{x,n} = \frac{d_{cutoff,k}}{2*\| p_4 - p_3\|} h \cos((P_2 - P_1)\cdot(P_4 - P_3)) + \frac{canvassize}{2}$$

$$E_{y,n} = \frac{d_{cutoff,k}}{2*\| p_4 - p_3\|} h \sin((P_2 - P_1)\cdot(P_4 - P_3)) + \frac{canvassize}{2}$$

$$color = \langle 255|(k_n*255) - 50|0\rangle$$

where $E_{x,n}$ and $E_{y,n}$ are the respective x and y coordinates of the output image, $d_{cutoff,k}$ is the calculated cutoff value at the k value of interest, and *canvassize* is the diameter in pixels of the circular image being drawn. Color is a vector containing three values representing red, green, and blue intensities, allowing transition from red to orange with increasing k values.

3D Exclusion Volume Visualization. CSBB-ConeExclusion facilitates 3D visualization of the exclusion volume through the generation of PyMol scripts, which, when loaded into PyMol with the protein and docked poses, displays a cone volume and allows easy identification and inspection of encroaching protein atoms. While PyMol does not include a cone primitive, the cone is approximated using multiple cylindrical sections of varying diameters.

With docking results loaded into different PyMol states, poses may be examined using common "VCR" style controls, the different poses bringing with them the correct cone exclusion volume.

In order to demonstrate CSBB-ConeExclusion using a simple one protein-one ligand system, a potent[23] Cyclin Dependent Kinase 2 (CKD2) inhibitor CYC202 (R-roscovitine)[24] (see Supporting Information Figure S1) has been docked using the program FRED[25] (OpenEye Scientific Software Inc.) to an apo CDK2 structure (PDBID: 3DDQ),[26] generating 100 predicted protein−ligand complexes. For demonstration purposes we ran our cone exclusion software on all FRED outputs including energetically unfavorable poses. Defining the spacer attachment

dx.doi.org/10.1021/ci400371q | J. Chem. Inf. Model. 2013, 53, 3156−3162

point to be the para position of the aromatic ring attached at the 6 position of the purine core, CSBB-ConeExclusion finds the spacer attachment point for each of the 100 docked poses. Any chemical moiety may be specified and used as an attachment point. Along with the PDB file representing the protein, CSBB-ConeExclusion generates statistics in the form of a CSV file for use as input to automated filters, 100 images in the PNG file format for quick visual inspection and rejection of poses, and finally a Python script which may be read into PyMol allowing visualization of the cone exclusion volume and protein atom encroachment (see the Supporting Information). Figure 2 shows

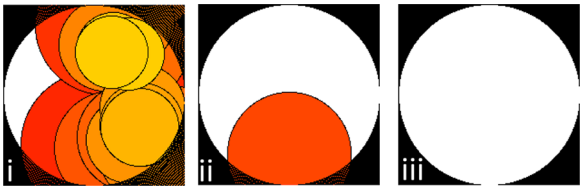


Figure 2. Exemplary 2D exclusion volume plots. i) Solid support attachment site blocked by protein atoms - solution docking pose unsuitable for solid support synthesis. ii) Encroachment of one protein atom into cone exclusion volume, investigation of structural features recommended determining if solution pose would work on solid support. iii) No encroachment of protein atoms into the exclusion volume, solution pose is amenable to solid support library production.

three exemplary 2D exclusion volume plots from the output of 100 docked conformations of CYC202. Figure 3 shows use of an

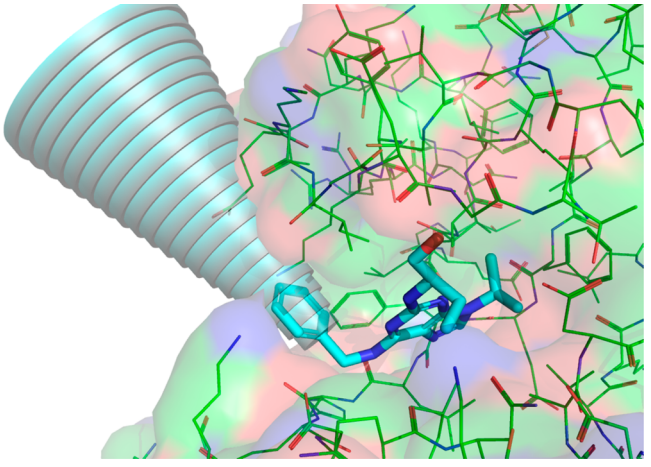


Figure 3. CYC202 docked to apo-CDK2 (PDBID: 3DDQ). Cone exclusion volume shown.

automatically generated PyMol script used to visualize the cone of uncertainty. Execution time for CSBB-ConeExclusion is negligible on the data set: sub 5 s on one processor of a modest 2010 Linux server for 100 docking poses.

Applications within Library Design. The CSBB-ConeExclusion technique has been brought to its logical conclusion through application of the technique to guide combinatorial solid supported library design. A combinatorial library based around a purine core for drug discovery efforts with the protein CDK2 offered a promising starting point. Using the previously mentioned structure 3DDQ from the PDB, two virtual libraries were constructed, each consisting of a purine core substituted in the 9 position with an isopropyl group and combinatorially enumerated with two selections of primary amines in both the 2 and 6 positions. One set of primary amines was restricted to the NH2 functionality (245 unique R-groups), while the other set of amines, in addition, contained a derivatizable acid functionality for attachment to solid support (18 unique R-groups). We define Library 1 to have been enumerated using a bead attachment point on the amine off of position 6 of the purine ring. Library 2 was designed with a bead attachment point on the amine off of the 2 position (see Figure 4). Enumerating creates 4410 unique molecular entities for each library.

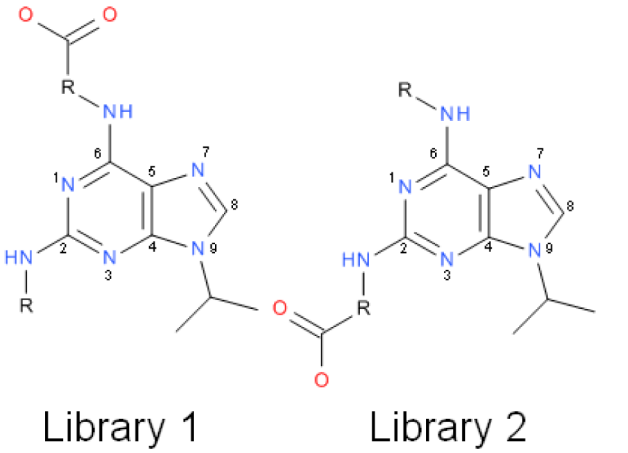


Figure 4. Two experimentally possible virtual purine libraries. Library 1 bead attachment point provided by an amine in the purine 6 position, and Library 2 bead attachment point provided by an amine in purine 2 position.

After enumeration, up to 200 of the lowest energy conformers for each unique compound are retained within the two libraries from an Omega2[27] run (OpenEye Scientific Software Inc.). Library conformers were docked to CDK2 using the molecular docking program FRED (full parameters are available in the Supporting Information). The three top scoring docked conformations for each library member were retained. A filtering procedure was then performed to remove invalid dockings where the predicted binding mode of the purine core had an RMSD of greater than 2 Å to the crystallographic purine core of CYC202 complexed to CDK2. To illustrate the robustness of our process and to parametrize the algorithm, shell scripting was then used to run the CSBB-ConeExclusion program with a range of parameters for d_2 (radius of the thick end of the cone) and l (cone length) for both Library 1 and Library 2. It is suggested that application of CSBB-ConeExclusion to library design always be coupled to a parameter sweeping exercise. This ensures that with a wide range of different R-groups and parameters, the user obtains a view of the robustness of the approach applied to a specific library design and protein to which the dockings have been made.

As can be seen from Figure 6, this test case shows that with only one exception, where both $d_2 \geq 12$ Å and $l \leq 6$ Å, Library 1 has produced more poses amenable to solid support. With low l values, the cone takes the form of a disc; in the case where $d_2 \geq 12$ Å, the insertion of this disc-shaped exclusion volume into the space of the attachment atom creates meaningless results. Only when the cone reaches a sensible length and moves from a disc shape to a cone shape, as described in the requirements for such a filter, beginning around 5 Å can we gather and draw conclusions from the data.

dx.doi.org/10.1021/ci400371q | J. Chem. Inf. Model. 2013, 53, 3156−3162

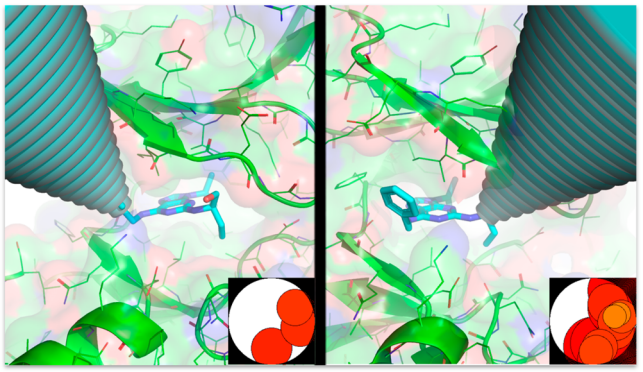


Figure 5. Comparison of CYC202 attachment points. With the attachment point at the para position of the ring located off of the 6 position of the purine core (left image), 3 protein atoms are predicted to encroach within the cone exclusion volume (exclusion volume defined as $d_1 = 2$, $d_2 = 20$, $l = 20$). Visual inspection of the crystal structure suggests this is a viable strategy for attachment to solid support. An alternative attachment strategy (right) where the alcohol of CYC202 is transformed to an acid to provide an attachment point can be seen to produce many clashes with protein atoms encroaching into the exclusion volume.

Building, docking, and filtering of the virtual purine libraries has shown a strong preference for a solid support/cross-linking attachment point at the 6 position of the purine ring. Literature support of this recommendation is present in two forms. Bach et al. attach CYC202 to agarose in an affinity chromatography approach at the 6 position of the purine ring and state that no significant affinity changes were observed with other well-known CYC202 targets: CDK1, CDK5, ERK1, ERK2, and DYRK1A.[28] Knockaert[29] et al. attach purvalanol, a close similar of CYC202 which can be observed in crystal structures to adopt an almost identical binding mode in complex with CDK2, to agarose beads via the 6 position of the purine ring via a PEG spacer and are able to identify known interacting proteins from various cell types and tissues via affinity chromatography. Figure 5 shows a run of CSBB-ConeExclusion on the crystal structure of CYC202, defining attachment points at both the 2 and 6 positions. Position 6 is a clearly favored attachment point.

CONCLUSION

We have demonstrated an automated algorithm to determine the suitability and adaptability of solution docking poses to solid support. Output can be used in three ways: (a) automated gating of poses based on statistics for each pose, (b) 2D visual inspection using uniquely generated 2D exclusion volume plots, and (c) rigorous 3D visualization using augmented PyMol views created through automatically generated PyMol scripts. We have also applied the filtering process to combinatorial library design, tailoring the choice of attachment point to an exemplary target, CDK2. Previous to the realization of this technique, the choice of attachment point relied heavily upon a modelers intuition garnished from available structural information. We are now able

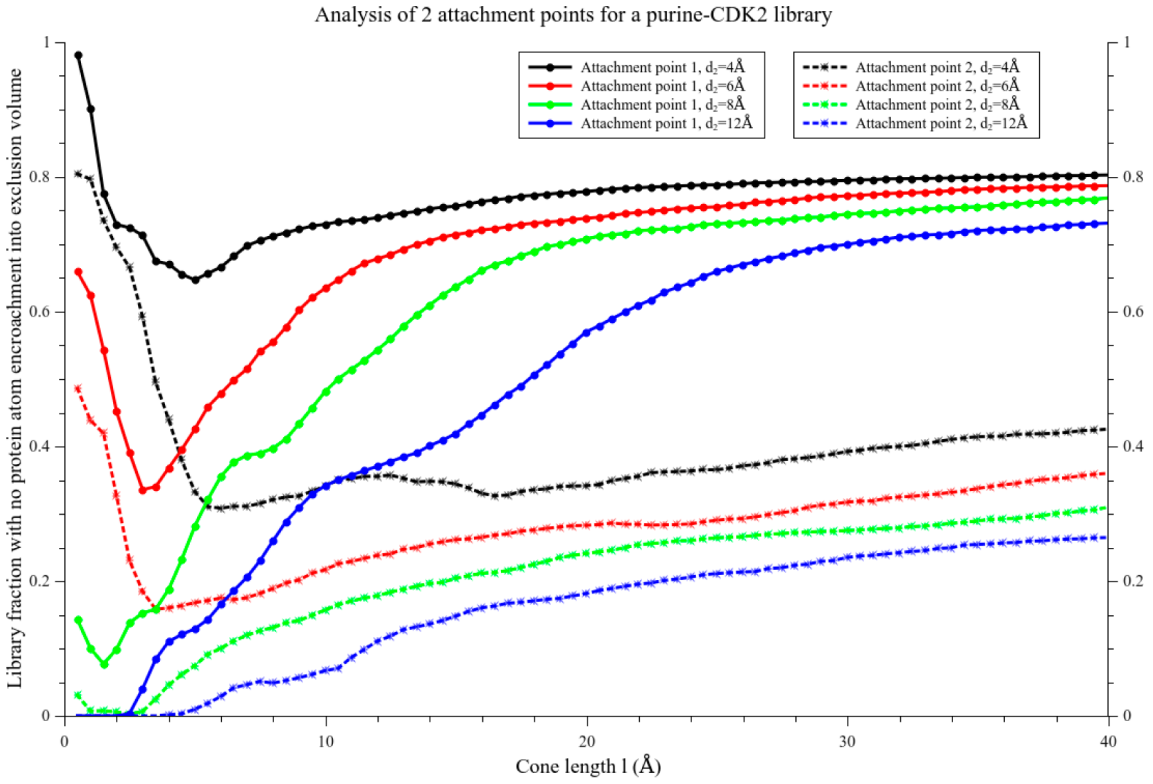


Figure 6. Plot showing measured fraction of library docked with no observed exclusion volume encroachment. Varying d_2 and l distances for the two libraries linked to bead via attachment position 6 and 2 on the purine ring, respectively. The plot shows the clear difference in library populations not interfering with different cone angles. This indicates that no atom of the target protein lies within a necessary exclusion space to allow a library to fit into a binding site. Library 1 with attachment point at position 6 of the purine ring shown with solid line. Library 2 with attachment point at position 2 of the purine ring shown with a dashed line. The chemist would conclude that only Library 1 should be produced for the exemplary CDK ATP/hinge region binding site, suggesting attachment at position 6 of the purine ring when targeting CDK2 with a purine library to be a very good option for generation of a successful targeted library.

dx.doi.org/10.1021/ci400371q | J. Chem. Inf. Model. 2013, 53, 3156−3162

to quantitatively compare possible attachment points within a library context. This reduces the number of distinct libraries that need to be made for a target, freeing up chemical capacity, reducing costs, and maximizing the likelihood of a library delivering high quality hits for follow up hit and lead compound-discovery efforts.

■ ASSOCIATED CONTENT

Supporting Information

Combinatorial libraries 1 and 2 are available in the SD file format, in both 2D and docked 3D conformations. Accompanying each library is a script file allowing PyMol visualization of cone angle encroachment of predicted complex (top 3 poses for each small molecule) and compressed (tar.gz) files containing 2D exclusion volume plots for each SDF entry. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Corresponding author address: School of Biological Sciences, University of Edinburgh, The King's Buildings, CH Waddington Building, Mayfield Road, Edinburgh, EH9 3JD, UK. E-mail: manfred.auer@ed.ac.uk (M.A.).

*Corresponding author address: School of Biological Sciences, University of Edinburgh, The King's Buildings, CH Waddington Building, Mayfield Road, Edinburgh, EH9 3JD, UK. E-mail: s.shave@ed.ac.uk (S.S.).

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to thank the PRIMES consortium, Scottish Universities Life Sciences Alliance (SULSA) and the Medical Research Council (MRC). This work is funded by the Scottish Universities Life Sciences Alliance (SULSA), Medical Research Council (MRC), and by the European Community's 7th Framework Program (FP7/2007−2013) under grant agreement no 278568 "PRIMES".

■ ABBREVIATIONS

CDK1, cyclin-dependent kinase 1; CDK2, cyclin-dependent kinase 2; CDK5, cyclin-dependent kinase 5; CONA, confocal nanoscanning; DYRK1A, dual specificity tyrosine-phosphorylation-regulated kinase 1A; ERK1, extracellular signal-regulated kinase 1; ERK2, extracellular signal-regulated kinase 2; PEG, polyethylene glycol; TOBOC, tagged one-bead one-compound

■ REFERENCES

(1) Hintersteiner, M.; Buehler, C.; Uhl, V.; Schmied, M.; Müller, J.; Kottig, K.; Auer, M. Confocal nanoscanning, bead picking (CONA): PickoScreen microscopes for automated and quantitative screening of one-bead one-compound libraries. *J. Comb. Chem.* **2009**, *11* (5), 886−894.

(2) Hintersteiner, M.; Kimmerlin, T.; Kalthoff, F.; Stoeckli, M.; Garavel, G.; Seifert, J. M.; Meisner, N. C.; Uhl, V.; Buehler, C.; Weidemann, T. Single bead labeling method for combining confocal fluorescence on-bead screening and solution validation of tagged one-bead one-compound libraries. *Chem. Biol.* **2009**, *16* (7), 724−735.

(3) Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **1991**, *37* (6), 487−493.

(4) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **1991**, *354* (6348), 82−84.

(5) Welsch, M. E.; Snyder, S. A.; Stockwell, B. R. Privileged scaffolds for library design and drug discovery. *Curr. Opin. Chem. Biol.* **2010**, *14* (3), 347−361.

(6) DeSimone, R. W.; Currie, K. S.; Mitchell, S. A.; Darrow, J. W.; Pippin, D. A. Privileged structures: applications in drug discovery. *Comb. Chem. High Throughput Screening* **2004**, *7* (5), 473−493.

(7) Spandl, R. J.; Bender, A.; Spring, D. R. Diversity-oriented synthesis; a spectrum of approaches and results. *Org. Biomol. Chem.* **2008**, *6* (7), 1149−1158.

(8) Lyne, P. D. Structure-based virtual screening: an overview. *Drug Discovery Today* **2002**, *7* (20), 1047−1055.

(9) Cavasotto, C. N.; Orry, W.; Andrew, J. Ligand docking and structure-based virtual screening in drug discovery. *Curr. Top. Med. Chem.* **2007**, *7* (10), 1006−1014.

(10) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **1996**, *10* (4), 293−304.

(11) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455−461.

(12) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43* (25), 4759−4767.

(13) Willett, P. Similarity-based virtual screening using 2D fingerprints. *Drug Discovery Today* **2006**, *11* (23), 1046−1053.

(14) Ballester, P. J.; Richards, W. G. Ultrafast shape recognition to search compound databases for similar molecular shapes. *J. Comput. Chem.* **2007**, *28* (10), 1711−1723.

(15) Sheridan, R. P. The most common chemical replacements in drug-like compounds. *J. Chem. Inf. Comput. Sci.* **2002**, *42* (1), 103−108.

(16) Sheridan, R. P. Finding multiactivity substructures by mining databases of drug-like compounds. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (3), 1037−1050.

(17) Patterson, D. E.; Cramer, R. D.; Ferguson, A. M.; Clark, R. D.; Weinberger, L. E. Neighborhood behavior: a useful concept for validation of molecular diversity descriptors. *J. Med. Chem.* **1996**, *39* (16), 3049−3059.

(18) Brown, R. D.; Martin, Y. C. Use of structure-activity data to compare structure-based clustering methods and descriptors for use in compound selection. *J. Chem. Inf. Comput. Sci.* **1996**, *36* (3), 572−584.

(19) O'Boyle, N. M.; Morley, C.; Hutchison, G. R. Pybel: a Python wrapper for the OpenBabel cheminformatics toolkit. *Chem. Cent. J.* **2008**, *2* (5).

(20) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminf.* [Online] 2011, 3, Article 33. http://www.jcheminf.com/content/3/1/33 (accessed Nov 19, 2011).

(21) Tschumperlé, D. *The CImg Library: The C++ Template Image Processing Library 1.5.6.* http://cimg.sourceforge.net/ (accessed November 19, 2013).

(22) DeLano, W. L. The PyMOL Molecular Graphics System; DeLano Scientific: San Carlos, CA, 2002.

(23) Azevedo, W. F.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S. H. Inhibition of cyclin-dependent kinases by purine analogues. *Eur. J. Biochem.* **1997**, *243* (1−2), 518−526.

(24) Meijer, L.; Borgne, A.; Mulner, O.; Chong, J. P. J.; Blow, J. J.; Inagaki, N.; Inagaki, M.; Delcros, J. G.; Moulinoux, J. P. Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin dependent kinases cdc2, cdk2 and cdk5. *Eur. J. Biochem.* **1997**, *243* (1−2), 527−536.

(25) McGann, M. FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2011**, *51* (3), 578−596.

(26) Bettayeb, K.; Oumata, N.; Echalier, A.; Ferandin, Y.; Endicott, J.; Galons, H.; Meijer, L. CR8, a potent and selective, roscovitine-derived inhibitor of cyclin-dependent kinases. *Oncogene* **2008**, *27* (44), 5797−5807.
(27) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50* (4), 572−584.
(28) Bach, S.; Knockaert, M.; Reinhardt, J.; Lozach, O.; Schmitt, S.; Baratte, B.; Koken, M.; Coburn, S. P.; Tang, L.; Jiang, T. Roscovitine targets, protein kinases and pyridoxal kinase. *J. Biol. Chem.* **2005**, *280* (35), 31208−31219.
(29) Knockaert, M.; Meijer, L. Identifying *in vivo* targets of cyclin-dependent kinase inhibitors by affinity chromatography. *Biochem. Pharmacol.* **2002**, *64* (5), 819−825.