A Shape-Based 3-D Scaffold Hopping Method and Its Application to a Bacterial Protein−Protein Interaction

Thomas S. Rush III,*,† J. Andrew Grant,‡ Lidia Mosyak,† and Anthony Nicholls§

Department of Chemical & Screening Sciences, Wyeth Research, 87 Cambridge Park Drive, Cambridge, Massachusetts 02140, AstraZeneca Pharmaceuticals, EST(Chem) 26F17, Mereside, Macclesfield, Cheshire SK10 4TG, U.K., and Openeye Scientific Software, 3600 Cerrillos Road, Suite 1107, Santa Fe, NM 87507

Received September 8, 2004

In this paper, we describe the first prospective application of the shape-comparison program ROCS (Rapid Overlay of Chemical Structures) to find new scaffolds for small molecule inhibitors of the ZipA−FtsZ protein−protein interaction, a proposed antibacterial target. The shape comparisons are made relative to the crystallographically determined, bioactive conformation of a high-throughput screening (HTS) hit. The use of ROCS led to the identification of a set of novel, weakly binding inhibitors with scaffolds presenting synthetic opportunities to further optimize biological affinity and lacking development issues associated with the HTS lead. These ROCS-identified scaffolds would have been missed using other structural similarity approaches such as ISIS 2D fingerprints. X-ray crystallographic analysis of one of the new inhibitors bound to ZipA reveals that the shape comparison approach very accurately predicted the binding mode. These experimental results validate this use of ROCS for chemotype switching or "lead hopping" and suggest that it is of general interest for lead identification in drug discovery endeavors.

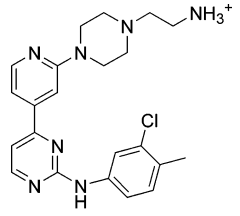
Introduction

It is now well-known that bacterial resistance to first-line antibiotics poses a serious threat to public health. For this reason many pharmaceutical and biotechnology companies are aggressively pursuing novel ways to kill bacteria. Toward this end, Wyeth Research began investigating a new bacterial target with the aim of developing a small molecule antibiotic with activity against resistant strains. This target is the ZipA−FtsZ protein−protein interaction.

The bacterial protein ZipA appears to play an essential role in the formation and dynamics of the septal ring − a membrane-associated organelle that drives constriction and formation of new cell walls during cell division.[1] It is believed that ZipA stabilizes and anchors the septal ring to the cell membrane by directly interacting with a primary component, FtsZ. In fact, it has been shown that the two proteins interact via their C-terminal domains, and that this interaction appears to be critical to promote cell division; cells depleted of ZipA grow, but cannot divide.[2]

ZipA is a 36.4 kDa integral membrane protein comprising of three domains; a short N-terminal membrane-anchoring domain, a central domain rich in proline and glutamine, and a large cytoplasmic C-terminal domain (residues 185−328). Recently, a high-resolution crystal structure for the C-terminal domain of ZipA was solved, along with a structure of a complex between ZipA and a 17 residue C-terminal fragment of FtsZ.[3] It was found that the C-terminal domain of ZipA is composed of a six stranded β-sheet structure packed against three α-helices. On the uncovered face of the β-sheet resides a large and shallow hydrophobic cavity exposed to solvent. It is within this cavity that the FtsZ 17-mer binds (Figure 1). The K_D for the ZipA−FtsZ 17mer interaction was measured to be $\sim$7 μM.[4]

To find small molecule inhibitors of this critical protein−protein interaction, Wyeth performed a fluorescence polarization-based high-throughput screen of $\sim$250,000 corporate compounds. Approximately 30 reproducible and interesting hits were identified from this screen, of which, compound **1** was the most potent (K_i



1

of 12 μM). This lead was cocrystallized with ZipA and the X-ray structure was subsequently used to guide further development of the series (see Figure 2).[4] Computational MM-PBSA-based calculations,[5] determined that for the peptides and small molecules cocrystallized with ZipA, van der Waals interactions and the hydrophobic effect were the major contributors to the free energy of binding.

Unfortunately, it was deemed very early on that **1** had potential issues with toxicity as it showed rather strong, nonspecific activity in both bacterial and yeast-based (*Candida albicans*) assays. In addition, there were

* To whom correspondence should be addressed. Telephone: (617) 665−8500. Fax: (617) 665−5682. E-mail: trush@wyeth.com.
† Wyeth Research.
‡ AstraZeneca Pharmaceuticals.
§ Openeye Scientific Software.

10.1021/jm040163o CCC: $30.25
Published on Web 02/10/2005

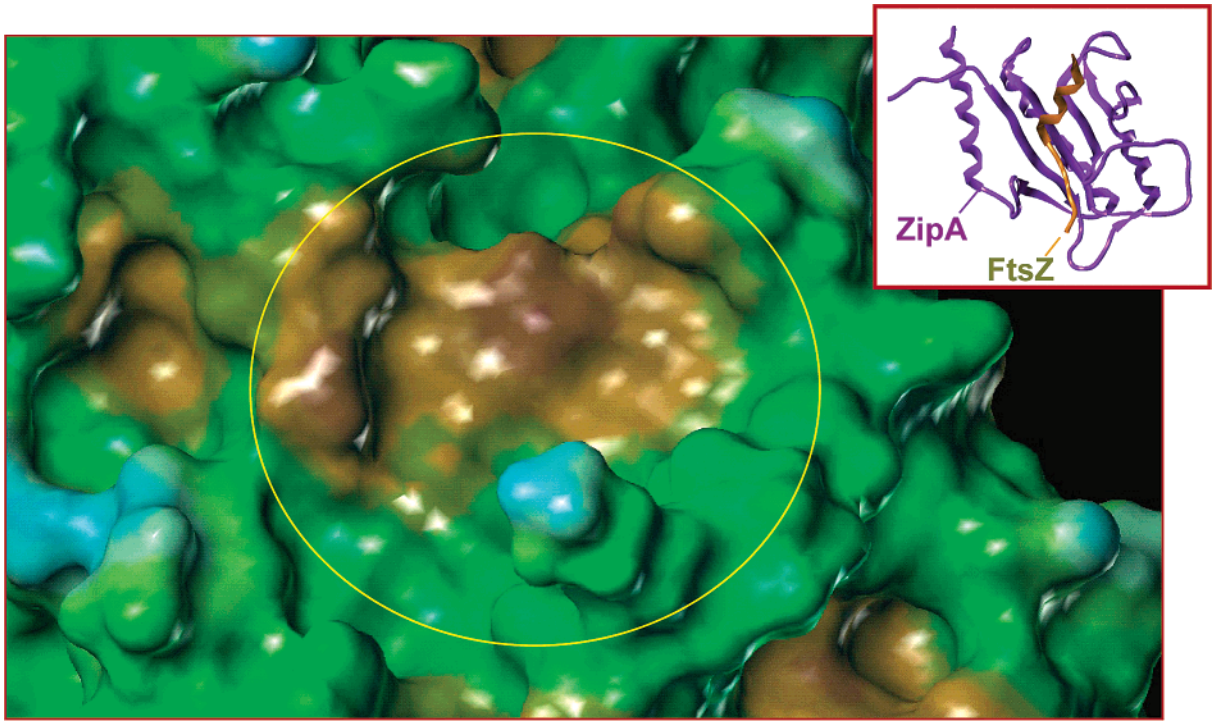


Figure 1. The solvent-accessible surface of the FtsZ binding site on ZipA, colored by lipophilic potential. Created by the MOLCAD module within Sybyl 6.8 (Tripos, Inc.); brown = hydrophobic, green = neutral, blue = hydrophilic. The inset is a ribbon diagram of the entirety of the ZipA−FtsZ structure. The yellow circle represents the primary interaction site.

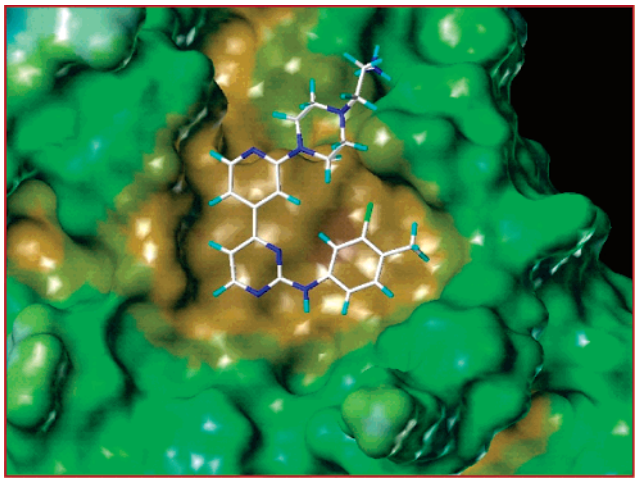


Figure 2. The structure of **1** bound to ZipA. Surface details are the same as in Figure 1.

potential patent issues as amine substituted pyridylpyrimidines are well represented in the literature in the context of kinase inhibition with both antitumor and antiviral applications.[6−10] This prevalence in the literature is also suggestive that the scaffold could show other toxic characteristics later on in development. So rather than forfeit the information contained in both the molecule and the nature of its interaction with the protein, a new shape-based computational procedure was used to identify other molecules within Wyeth's corporate database that could fill the same region of the binding pocket and make comparable interactions with the protein. This paper describes the computational procedure employed, and its successful application to finding new inhibitor scaffolds for the ZipA−FtsZ interaction.

Computational and Experimental Procedures

Shape Comparisons and ROCS. The method implemented in this work to lead or scaffold hop is known as ROCS − Rapid Overlay of Chemical Structures.[11] ROCS is used to perceive similarity between molecules based on their three-dimensional shape. The algorithm is based on earlier implementations of molecular shape comparison.[12] The general objective of these methods is to find and quantify the maximal overlap of the volume of two molecules. This rigorously defines a strong notion of shape that is either lacking in other approaches or used weakly. By the latter we mean the application of the word "shape" to concepts that lack connection to the everyday notion of the shape of objects. For instance, comparing the shapes of two coffee cups does not require definition of pharmacophores or topological ring indices, whereas ROCS could easily be applied to such a task.

Formally, the overlap of shape between two objects or molecules A and B, $O_{A,B}$, is defined as

$$O_{A,B}(\vec{q}^{A},\vec{q}^{B}) = \int\int\int \chi^{A}(\vec{r},\vec{q}^{A})\chi^{B}(\vec{r},\vec{q}^{B})\,d\vec{r} \qquad (1)$$

Here **r** is a position in space, q is a set of variables that determine orientation and position, and χ is the "characteristic volume" function. The volume integral appearing in eq 1 is over the whole space ($d\vec{r} = dx\ dy\ dz$). The characteristic function is typically taken as one inside and zero outside. This was the original approach of Masek et al.,[12] wherein the volume of a molecule is taken to be that defined by a set of fused spheres of appropriate atomic radii. However, the actual functional form of χ is then very complicated and discontinuous. This made the critical step of optimizing the overlap $O_{A,B}$ with respect to the orientation and position variables q very challenging. In practice, the procedure was slow and not robust in finding the global minimum. The crucial step forward was the introduction of a description of molecular shape using continuous functions constructed from atom-centered Gaussians. This was used to define a χ of very similar behavior to the hard-sphere volume-function, for instance, predicting total

volumes and areas of small molecules to within one percent.[13] The function used in the Gaussian shape model is

$$\chi(\vec{r}) = 1 - \prod_{i=1}^{i=N}(1 - g_i(\vec{r})) \quad (2)$$

where

$$g_i(\vec{r}) = p_i \, e^{-\gamma \vec{r}_i^{\,2}} \quad (3)$$

and the local coordinate system, $r_i = |\vec{r}_i| = |r - \vec{c}_i|$, is defined as a distance vector from the atomic center ($\vec{c}_i$). Equation 2 can be applied to any set of atom-based functions and accounts for the overlap between such functions, to any order. Such overlaps become very complex when χ is a hard-sphere function but trivial when a Gaussian (products of Gaussians are merely other Gaussians). An atom of radius R is modeled as a Gaussian of width γ given by

$$\gamma = \pi\left(\frac{3p}{4\pi R^3}\right)^{2/3} \quad (4)$$

This choice of γ ensures that for an atom of radius R, the Gaussian volume computed from the volume integral of eq 4 equals the equivalent hard-sphere atomic volume. A universal value of 2.7 for the Gaussian weight parameter reproduces molecular volumes, when used in eq 2.

Given such a Gaussian volume function, most of the difficulties associated with the hard-sphere approach are solved. Convergence to the global minimum is usually possible from one of four starting positions, those for which the axes of the moments of inertia are superimposed, followed by rotation of π radians about each axis. In addition, the extreme smoothness of the Gaussian functions allows for robust optimization. Further experimentation showed that the orientation found from this approach closely correlated with that from the hard-sphere functions and that the relative orientation of ligands binding to proteins could be reasonably accurately predicted based on their Gaussian volume overlay.[14] The accuracy of the method is little affected by using a reduced set of Gaussians generated from the neglect of overlap Gaussians generated by eq 2. Since there are typically 10 times more Gaussians of appreciable volume (greater than 0.01 cubic Angstroms) in the χ from eq 2 as there are atoms, this is an appreciable simplification.

Finding $O_{A,B}$ allows for the calculation of a fundamental quantity between any two molecules, the shape distance, $D_{A,B}$:

$$D_{A,B} = \sqrt{O_{A,A} + O_{B,B} - 2O_{A,B}} \quad (5)$$

This property is a true metric (obeying the triangle inequality), and predicts that shape is actually an intrinsic, not an extrinsic or relative property. One can also produce a related property that has more familiarity, a Tanimoto, $T_{A,B}$:

$$T_{A,B} = O_{A,B}/(O_{A,A} + O_{B,B} - O_{A,B}) \quad (6)$$

This shape Tanimoto has the recognizable quantity that it is 1.0 if two shapes are identical, and 0.0 if completely different. Two shapes are never completely different, i.e., have zero overlap, but shapes may be identical for different molecules. This Tanimoto is the measure used in this study to compare molecules.

This is the approach of ROCS. Matches are based only on volume overlap of optimally aligned molecules; therefore, they are virtually independent of the atom types and bonding patterns present in the query and search molecules. The emphasis and goal of this approach is to identify molecules that can adopt shapes extraordinarily similar to the query and in doing so increase the chance of "scaffold hopping" or "lead hopping".

Database Preparation. The *in silico* screening library for this study was a "lead-like" subset of Wyeth's corporate compound collection. The set included all molecules with molecular weights between 180 and 400, less than 9 rotatable bonds, no more than one chiral center, a log D(7.4) of less than 4.0, fewer than 4 ionizable groups, and believed to contain "nonreactive" functionalities. The SMILES strings[15] for the database of lead-like molecules were expanded into a set of 3D conformers using the program OMEGA[16] with torsional sampling that ranged between 5 and 60 degrees, depending on bond type. All conformations within 5 kcal/mol of the lowest energy conformer were kept. For chiral compounds in the database that lacked a stereospecific designation, both enantiomers were created and conformationally expanded.

Scoring Protein−Ligand Interactions. The detailed analyses of protein−ligand interactions were accomplished through the use of a scoring function that includes molecular mechanics and surface area-based energy terms, where the effects of solvation were captured using a Poisson-based, continuum solvent approach. This type of scoring function (most widely known as MM-PBSA) has been well validated by Kollman and co-workers (e.g. ref 5). The form of the scoring function used in this work is[17]

$$\Delta G_{bind}^{solv} \approx [\Delta E_{elec}^{gas} + \Delta E_{vdW}^{gas}] + \Delta\Delta G_{PB} + \Delta\Delta G_{SA} \quad (7)$$

where

$$\Delta\Delta G_{PB} = \Delta G_{PB}^{PL} - (\Delta G_{PB}^{P} + \Delta G_{PB}^{L}) \quad (7a)$$

and

$$\Delta\Delta G_{SA} = \Delta G_{SA}^{PL} - (\Delta G_{SA}^{P} + \Delta G_{SA}^{L}) \quad (7b)$$

The individual components were all evaluated with OpenEye software.

Experimental Assays. Molecules selected using ROCS were screened for activity against the ZipA−FtsZ interaction by a fluorescence polarization assay described in detail in a recent publication.[4] Briefly, the assay utilizes a fluorophore-modified FtsZ peptide. Percent inhibitions at fixed concentrations are determined by comparing the fluorescence polarization of the inhibitor−ZipA−peptide solutions to that observed for the ZipA−peptide solution. IC50's and K_D's were determined only for molecules that exhibited a dose-dependent response at 500, 200 and 50 μg/mL.

Crystallography. Two of the most promising hits identified using our ROCS-based search procedure were cocrystallized with ZipA to verify their mode of binding, and to provide structural insights for further medicinal chemistry efforts. Cocrystals were obtained in hanging drops using PEG6K, 0.1M MES, pH 6.0 as precipitating solutions. Diffraction data were collected at the Advanced Light Source (Berkeley, CA), then integrated and scaled with DENZO/SCALEPACK. Crystals of ZipA with compound **1** belong to space group $P2_12_12_1$ with unit cell parameters a = 44.97 Å, b = 50.61 Å, c = 84.39 Å, and with one molecule in the asymmetric unit. ZipA with compound **3** crystallized in the space group $P2_1$ with unit cell parameters a = 52.69 Å, b = 39.52 Å, c = 70.27 Å, β = 104.9°, and with two molecules in the asymmetric unit. The crystal structures were solved using molecular replacement (AmoRe[18]) with the ZipA monomer as a search model. Refinement was performed with CNS,[19] and the process of the refinement and rebuilding was repeated until most of the water molecules fit the density. The refined models (with no inhibitors added) were then used to calculate the $3F_o - 2F_c$ and $F_o - F_c$ maps, which showed clear electron density for the bound compounds. The final models of complexes **1** and **3**, respectively, have R-factor

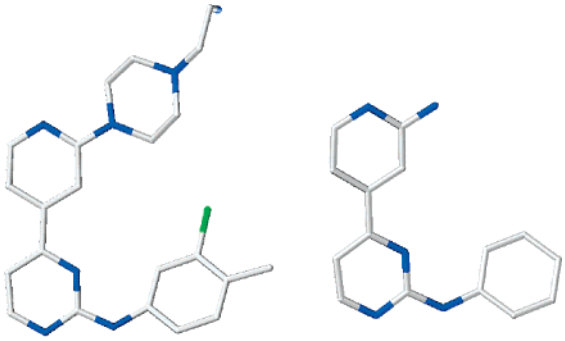


Figure 3. The two structures used in the ROCS query. To the left is the "FULL" query where all heavy atoms of **1** were used, and to the right is the "PART" query where the heavy atoms of the (2-aminoethyl)piperazine side chain were removed.

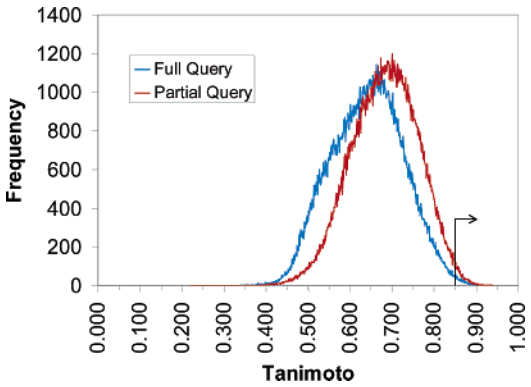


Figure 4. A plot showing the distribution of the best Tanimoto scores found for each molecule in the in silico-screening database. The line depicts the cutoff chosen for further study.

values of 22.6% (R_{free} of 22.8%) for the resolution range 45−2.0 Å and of 20.7% (R_{free} of 24.2%) for the resolution range 24−1.9 Å. The atomic coordinates (codes 1Y2F and 1Y2G) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Results

The crystal structure of compound **1** and a substructure not including the (2-aminoethyl)piperazine substituent, were used as templates for the ROCS searches. These two query molecules are denoted as ROCS_FULL and ROCS_PART, respectively (Figure 3). The distribution of Shape Tanimoto values obtained from searching the lead-like database using these queries is shown in Figure 4. The mean Shape Tanimoto for ROCS_FULL is 0.57 and 49 molecules have Shape Tanimoto values greater than 0.85, whereas the smaller query, ROCS_PART, has a mean Shape Tanimoto of 0.68 and 2203 molecules are found with a shape similarity that exceeds 0.85. Figure 4 shows that the distribution is roughly Gaussian with the median shifted toward higher Tanimotos for the smaller shape. In a separate publication (Haigh and Grant, in preparation) it is shown that this difference is predominantly a volume effect; molecules with volumes similar to the median have higher average Tanimotos than smaller or larger molecules. A general rule-of-thumb is that a Shape Tanimoto of greater than 0.75 provides visual shape similarity.

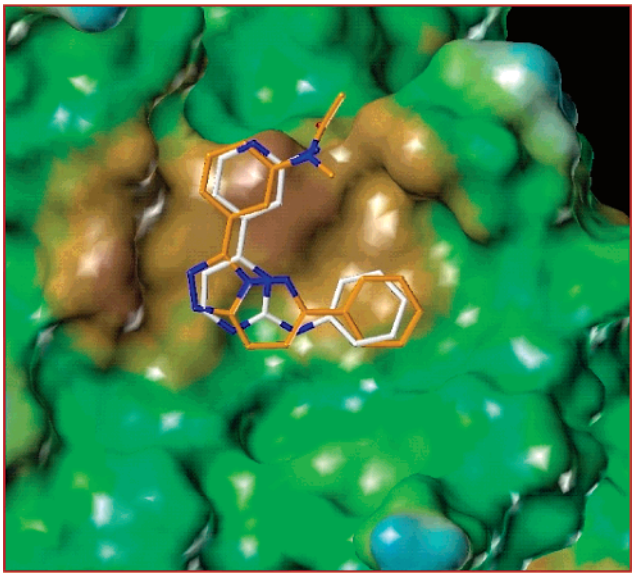


Figure 5. An example ROCS hit (ROCS_PART-18) overlaid with **1**, the query molecule.

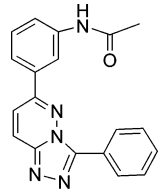
Since ROCS does not directly use protein structural information, molecules identified by the procedure can potentially have large volume overlap with regions of the ZipA binding site, and thus would be unlikely to bind to ZipA. To detect such clashes, each ROCS hit was overlaid with the query molecule (**1**) in the ZipA binding site, then the intermolecular van der Waals energy was computed between the ROCS hit and the protein. Molecules that resulted in positive values of this interaction energy were discarded from consideration. This screen was applied to all of the 49 molecules obtained from the search using the ROCS_FULL query and the 300 molecules with the highest Shape Tanimoto values obtained from the ROCS_PART query. Subsequent removal of molecules with the same scaffold as the query, and ensuring that molecules satisfied certain criteria connected with purity and availability, resulted in 29 molecules being selected for experimental screening in the ZipA−FtsZ assay. An example of a molecule selected using this overall approach is shown in Figure 5, which also shows the shape-overlay onto the underlying ROCS_PART template used in the search. The overlay itself is shown in the context of the binding site obtained for the parent compound **1.** The experimental activities of all of the molecules tested are given in Table 1, which reports separately the data for those obtained from the ROCS_FULL query (3 molecules), and the ROCS_PART query (26 molecules).

The three most interesting and active molecules from the query are ROCS_PART-11, ROCS_PART-18 and ROCS_PART-19 (compounds **2**−**4**), with K_D's determined to be 73.9, 83.1 and 85.1 μM, respectively. Although weaker than the original HTS hit, these molecules were smaller, did not show the same cytotoxic effects, and had less intellectual property concerns. For

Table 1. Fluorescence Polarization Assay Data and Tanimoto Scores for the Search Hits

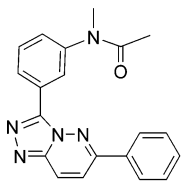
molecule	fluorescence polarization assay: % inhibition at 500 μg/mL	% inhibition at 200 μg/mL	% inhibition at 50 μg/mL	ROCS Tanimoto	ISIS 2-D similarity (×0.01)
		Partial Structure			
ROCS_PART-01	54.9	14.7	0.5	0.896	0.316
ROCS_PART-02	8.4	3.3	3.3	0.909	0.552
ROCS_PART-03	10.8	8.1	3.7	0.877	0.367
ROCS_PART-04	6.4	13.6	5.8	0.897	0.309
ROCS_PART-05	7.9	4.5	3.9	0.895	0.464
ROCS_PART-06	21.5	8.9	0.4	0.896	0.554
ROCS_PART-07	13.3	4.5	2.8	0.896	0.575
ROCS_PART-08	2.3	0.0	7.2	0.879	0.434
ROCS_PART-09	0.0	0.0	0.0	0.910	0.284
ROCS_PART-10	19.3	20.6	15.1	0.926	0.357
ROCS_PART-11	55.0	45.4	32.8	0.901	0.363
ROCS_PART-12	15.5	10.7	11.2	0.898	0.367
ROCS_PART-13	31.5	44.5	35.5	0.905	0.364
ROCS_PART-14	21.9	13.8	15.4	0.907	0.331
ROCS_PART-15	4.1	17.5	12.1	0.892	0.319
ROCS_PART-16	20.0	15.3	7.1	0.877	0.330
ROCS_PART-17	7.1	26.5	13.3	0.934	0.358
ROCS_PART-18	76.6	58.0	23.9	0.907	0.348
ROCS_PART-19	76.3	55.8	23.5	0.876	0.364
ROCS_PART-20	11.6	2.3	0.8	0.867	0.328
ROCS_PART-21	1.7	6.9	5.5	0.871	0.356
ROCS_PART-22	26.4	20.4	6.9	0.894	0.361
ROCS_PART-23	7.5	6.9	0.9	0.893	0.350
ROCS_PART-24	28.3	29.5	19.8	0.908	0.710
ROCS_PART-25	13.4	9.6	3.4	0.878	0.539
ROCS_PART-26	8.6	2.4	0.4	0.859	0.508
		Full Structure			
ROCS_FULL-01	17.6	5.5	1.6	0.855	0.360
ROCS_FULL-02	4.6	2.7	6.2	0.854	0.429
ROCS_FULL-03	13.0	10.1	3.6	0.865	0.595

reference, molecule **2** was not in the original HTS

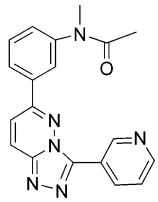


2

screening library, while molecules **3** and **4** were; how-



3



4

ever, the low activity values of **3** and **4** in the HTS setting excluded them from further consideration by the team.

A crystal structure of one of these molecules, ROCS_PART-18, in complex with ZipA at 1.9 Å resolution was obtained. The orientation predicted by ROCS (without any minimization in the binding site) is strikingly similar to that determined experimentally, as can be seen in Figure 6 (RMSD of 1.25 Å).

Discussion

The interaction between the C-terminal domains of FtsZ and ZipA primarily involves the burial of two hydrophobic faces.[3] The solvent accessible area buried upon binding is ~1350 Å^2. The hydrophobic cavity on ZipA (highlighted in Figure 1) is the binding site for the α-helical segment of the FtsZ 17-mer. The most extensively buried side chains are of the hydrophobic residues Leu372, Ile374, Phe377 and Leu378 that fit tightly into the ZipA pocket. Polar interactions made adjacent to the tight hydrophobic binding pocket include a Phe-Arg π-type interaction, and two hydrogen bonds between main chain atoms of ZipA and FtsZ ($FtsZ_{370}$/$ZipA_{250}$ & $FtsZ_{372}$/$ZipA_{248}$). A more complete description of the

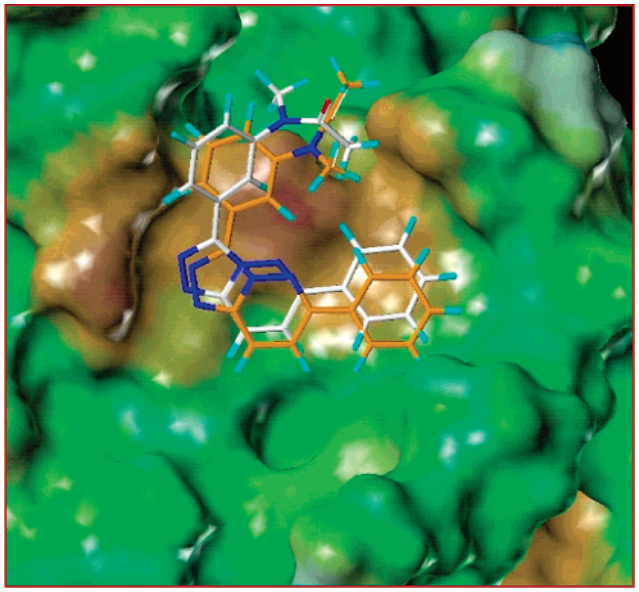


Figure 6. The X-ray structure of **3** overlaid with its ROCS determined pose.

structure of ZipA and its interaction with the FtsZ peptide can be found elsewhere.[3]

The crystal structure of **1** (Figure 2) shows that the molecule we identified from the HTS fits quite tightly into ZipA with an affinity (K_D 12 μM) similar to that of the FtsZ peptide (K_D 7 μM). However, the solvent accessible surface area buried upon binding of molecule **1** is 740 Å^2, appreciably smaller that that of the FtsZpeptide. Expressing the equilibrium binding data in terms of the binding interface area gives values of 9.2 and 5.2 cal mol^{-1} Å^{-2} for molecule **1** and the FtsZ peptide, respectively. This indicates that the relative efficiency of the small molecule inhibitor in utilizing surface area to promote binding to ZipA is nearly double that of the peptide. MM-PBSA[5] calculations suggest that the favorable contributions to the ΔG of binding of **1** arise from hydrophobic interactions, its conformational rigidity, van der Waals interactions, a π-stacking interaction with Phe269 (which forms part of the floor of the ZipA pocket), and minimal desolvation of polar groups.

Despite cytotoxicity issues, molecule **1** was a valuable discovery for the program, given that small molecule antagonists of protein−protein interactions are notoriously difficult to find.[20−22] Nevertheless, this molecule supports the hypothesis that small scaffolds with multiple aromatic rings connected by rigid linkers make good inhibitors of intramolecular hydrophobic collapse.[21] Further support for this hypothesis can be found in two of our recent publications on other families of ZipA−FtsZ antagonists.[23,24] Given these examples, it is tempting to speculate that screening libraries enriched with such compounds will be more successful in finding inhibitors of therapeutically interesting protein−protein interactions.[22]

The ROCS searches we performed (utilizing the experimentally determined binding mode of molecule **1**) provided the team with two new and novel scaffolds to work with. The algorithm used the shape of **1** to "scaffold hop" to molecules that are similar in their three-dimensional shape, but possess nonintuitive similarities according to their bond frameworks. By contrast, numerous 2-D substructure/similarity searches only identified more pyridylpyrimidine-like molecules that also tested positive for cytotoxicity. To show how different in 2-D similarity the ROCS hits are, we also calculated (Table 1) the 2-D MDL similarity index implemented in ISIS (MDL Information Systems).[25] As observed in Table 1, the highest 2-D similarity (x 0.01) for one of the ROCS chosen molecules was that for ROCS_PART-24 at a value of 0.710 − identical 2-D frameworks would have a value of 1.000. Most importantly, the molecules from the ROCS screen that were experimentally verified to inhibit the ZipA−FtsZ interaction in a manner consistent with compound **1** had 2-D similarities in the range of 0.348−0.364. As a point of reference, there are over 35,000 molecules in the Wyeth database with 2-D similarities values greater than this.

Figure 5 illustrates the ability of ROCS to overlay molecules based on shape. These molecules also share the low MW feature of **1**, as well as being conformationally rigid. The crystal structure of **3** bound to ZipA (Figure 6) validates the binding mode of this inhibitor, and suggests that the major contributors to binding are similar to that of **1**; namely favorable enthalpy from van der Waals interactions, the π-stacking interaction with Phe269 and minimal desolvation of polar groups, and favorable entropy gain from the hydrophobic effect. Although **3** is a weaker inhibitor than **1**, it has a lower MW and a near identical free energy of binding per "heavy" atom (i.e. non-hydrogen atoms) of 0.2 kcal mol^{-1}.[26] Ultimately, molecules **2**−**4** were all considered good "lead-like" starting points for developing antagonists of this protein−protein interaction. Further details of the opportunities present and the subsequent optimization of these scaffolds will be provided in future communications.

The approach taken here differs significantly from other 3-D search methods. In particular, it is almost orthogonal to the so-called 3-D pharmacophore approach. The active compounds here were found without *any* reference to the details of the actual chemistry of the known active. The approach relied only upon an intuitive fragmentation based upon physicochemical reasoning of the known active to generate appropriate templates for ROCS searches. This is in contrast to pharmacophore approaches, which require complex representations of molecules in terms of distances, angles, planes, centroids and very detailed chemical features. ROCS requires no preprocessing other than the generation of a set of energetically reasonable conformations of a corporate/vendor/virtual database. It currently processes conformers at a rate approaching one thousand per second on a single processor; orders of magnitudes faster than other 3-D search methods or docking programs, and is easily parallelizable. While it does not give energies of interactions as methods such as free energy perturbation or thermodynamic integration attempt to do, it does give results that are intuitive and easy to understand. In our hands it is capable of jumps in chemical structure that other methods cannot approach.

Conclusion

The new 3-D similarity algorithm described in this report rapidly identified two new and significantly different inhibitor scaffolds for Wyeth's ZipA program. The algorithm first uses ROCS to create a 3-D Gaussian volume for the query molecule and to rapidly search a multi-conformer database for molecules with similar shapes. Then, a simple evaluation of van der Waals interactions of the ROCS hits with the protein is performed to eliminate molecules with obvious steric clashes with the protein. This ultimately produced a final list of molecules similar in shape and protein binding site coverage as the query molecule. By following this procedure, the molecules ranked best need not be structurally similar (in a 2-D sense) to the query; in fact, in this example, we intentionally focused on hits that were not similar, thus allowing us to "scaffold hop" to molecules without the same toxicity and potential intellectual property issues as the query. In practice, we found that a rigorous definition of shape alone is more powerful than we, or others, had expected.

Crucial support for the viability of this approach was provided by a crystal structure for one of the ROCS hits. This structure confirmed that our new scaffold binds to ZipA at the same site as the FtsZ peptide, and the

query molecule **1**. Most importantly, the new molecules share the same tight fit within the site, similar hydrophobic interactions and a π-stacking interaction with a binding site phenylalanine. With lower molecular weights and superior ratios of binding energy to surface area to that of the FtsZ peptide, we believe that the ROCS-identified molecules provided us with ideal, lead-like scaffolds for further development of antagonists to the ZipA−FtsZ protein−protein interaction.

Acknowledgment. The authors would like to thank Alan Sutherland, Steve Haney, Alexey Ruzin, Desiree Tsao, and Cynthia Kenny of Wyeth Research for their guidance and experimental assistance in following up the results of the study.

References

(1) Lutkenhaus, J.; Addinall, S. G. Bacterial cell division and the Z ring. *Annu. Rev. Biochem.* **1997**, *66*, 93−116.
(2) Hale, C. A.; de Boer, P. A. Direct binding of FtsZ to ZipA, an essential component of the septal ring structure that mediates cell division in E. coli. *Cell* **1997**, *88*, 175−185.
(3) Mosyak, L.; Zhang, Y.; Glasfeld, E.; Haney, S.; Stahl, M. et al. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J.* **2000**, *19*, 3179−3191.
(4) Kenny, C. H.; Ding, W.; Kelleher, K.; Benard, S.; Dushin, E. G., et al. Development of a fluorescence polarization assay to screen for inhibitors of the FtsZ/ZipA interaction. *Anal. Biochem.* **2003**, *323*, 224−233.
(5) Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J. Am. Chem. Soc.* **2001**, *123*, 5221−5230.
(6) Zimmermann, J. PCT Int. Appl. WO 9509851, 1995.
(7) Furet, P.; Zimmermann, J.; Capraro, H. G.; Meyer, T.; Imbach, P. Structure-based design of potent CDK1 inhibitors derived from olomoucine. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 403−409.
(8) Albrecht, T.; Meijer, L.; Schaffer, P.; Schang PCT Int. Appl. WO 0230410, 2002.
(9) Eberle, M.; Stierli, D.; Pillonel, C.; Ziegler, H. PCT Int. Appl. WO 0193682, 2001.
(10) Eberle, M.; Ziegler, H.; Cederbaum, F.; Ackermann, P. PCT Int. Appl. WO 0253560, 2002.
(11) *ROCS*; OpenEye Scientific Software: Santa Fe, NM.
(12) Masek, B. B.; Merchant, A.; Matthew, J. B. Molecular shape comparison of angiotensin II receptor antagonists. *J. Med. Chem.* **1993**, *36*, 1230−1238.
(13) Grant, J. A.; Pickup, B. T. A Gaussian Description of Molecular Shape. *J. Phys. Chem.* **1995**, *99*, 3503−3510.
(14) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A fast method of molecular shape comaprison: a simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* **1996**, *17*, 1653−1666.
(15) James, C. A.; Weininger, D.; Delany, J. *Daylight Theory Manual*; Daylight Chemical Information Systems, Inc.: Mission Viejo, CA, 2004.
(16) *OMEGA*; OpenEye Scientific Software: Santa Fe, NM.
(17) Rush, T. S., III; Manas, E.; Tawa, G.; Alvarez, J. Solvation-Based Scoring for High Throughput Docking. *Virtual Screening in Drug Discovery*; CRC Press: Boca Raton, FL, expected 2005.
(18) Navaza, J. AMoRe − an automated package for molecular replacement. *Acta Crystallogr.* **1994**, *A50*, 157−163.
(19) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P. et al. Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Crystallogr.* **1998**, *D54*, 905−921.
(20) Toogood, P. L. Inhibition of protein−protein association by small molecules: approaches and progress. *J. Med. Chem.* **2002**, *45*, 1543−1558.
(21) Cochran, A. G. Antagonists of protein−protein interactions. *Chem. Biol.* **2000**, *7*, R85−94.
(22) Spencer, R. W. High-throughput screening of historic collections: observations on file size, biological targets, and file diversity. *Biotechnol. Bioeng.* **1998**, *61*, 61−67.
(23) Jennings, L. D.; Foreman, K. W.; Rush, T. S., 3rd; Tsao, D. H.; Mosyak, L. et al. Design and synthesis of indolo[2,3-a]quinolizin-7-one inhibitors of the ZipA-FtsZ interaction. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1427−1431.
(24) Sutherland, A. G.; Alvarez, J.; Ding, W.; Foreman, K. W.; Kenny, C. H. et al. Structure-based design of carboxybiphenylindole inhibitors of the ZipA-FtsZ interaction. *Org. Biomol. Chem.* **2003**, *1*, 4138−4140.
(25) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273−1280.
(26) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430−431.

JM040163O