

Targeting the BH3 Domain Mediated Protein–Protein Interaction of Bcl-xL through Virtual Screening

Prasenjit Mukherjee,[†] Prashant Desai,[‡] Yu-Dong Zhou, and Mitchell Avery*

Department of Medicinal Chemistry and Department of Pharmacognosy, School of Pharmacy, University of Mississippi, University, Mississippi 38677

Received January 25, 2010

Apoptosis, or programmed cell death, forms an important part of the cellular regulation machinery. The Bcl-2 protein family, comprising of proapoptotic and antiapoptotic members, forms an important part of the cells internal apoptotic pathway. Overexpression of the antiapoptotic members of the family in a number of cancer cell lines renders them immune to apoptosis and the ability to survive under conditions of cellular stress. Inhibition of the antiapoptotic members of the Bcl-2 family are, therefore, an interesting target for the development of anticancer therapy. An innovative structure-based virtual screening strategy was developed to identify inhibitors of Bcl-xL, an antiapoptotic member of the Bcl-2 family. Various innovative filters, such as receptor-based pharmacophore, cascade docking approach, cross-docking, and composite scoring with docking pose based descriptors were designed through exhaustive validation studies and implemented in the screening funnel. The 1.8 million ‘big-n-greasy’ subset from ZINC was screened using the protocol, and 45 compounds were finally selected for biological evaluation against Bcl-xL. The evaluation led to the identification of one low-micromolar and two weaker inhibitors belonging to novel scaffolds. Further evaluation of structure–activity relationships around these scaffolds could help in the development of anticancer leads against Bcl-xL.

INTRODUCTION

Apoptosis¹ is a specific form of programmed cell death observed in multicellular organisms and is characterized by specific morphological changes in a cell, culminating in its death. Deregulation of apoptotic pathways results in the loss of cellular homeostasis, leading to potentially harmful phenotypes, including cancer.^{2–4} Inactivating mutations of the proapoptotic proteins or overexpression of antiapoptotic proteins lead to a shutdown of the apoptosis pathways, allowing the cancer cell to survive in the presence of harmful mutations, DNA damage, and other cellular stress signals which would trigger apoptosis and cell death in a normally functioning cell. Moreover, cells which possess anomalies in the cell death signaling proteins are highly resistant to chemotherapeutic agents which function primarily by activating the apoptosis pathways. Restoration of the apoptotic pathways in a cancer cell can, therefore, be utilized as an effective anticancer therapy.^{1,5–9}

The apoptosis signaling pathways can be primarily divided into two groups,¹⁰ which are mediated through the mitochondria (intrinsic) and the death receptor (extrinsic) signaling pathway. A number of protein–protein interactions form critical components of the cellular apoptotic pathways and have been identified as therapeutic targets. In the extrinsic apoptosis pathway, the tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL)^{6,11–13} interacts with its

corresponding death receptor (CD4 or CD5), leading to its activation and to the corresponding induction of apoptosis. TRAIL, which mimics binding to the death receptor, can be used similarly to activate the external apoptotic pathway. In the intrinsic apoptotic pathway, the b-cell lymphoma 2 (Bcl-2) protein family,^{4,14,15} which acts as regulators of apoptosis, are therapeutically viable targets. The antiapoptotic Bcl-2 family members bind to the proapoptotic members, thereby preventing them from performing their function. In cancer cells, the antiapoptotic Bcl-2 family members are overexpressed. Inhibitors which bind to the antiapoptotic Bcl-2 family members would release the proapoptotic members, allowing them to perform their natural role in the induction of apoptosis. Another protein–protein interaction of therapeutic relevance is the interaction between the inhibitor-of-apoptosis (IAP) protein¹⁶ and its negative regulator SMAC¹⁷ (second mitochondria-derived activator of apoptosis)/DIABLO¹⁸ (direct IAP-binding protein with low pI). IAP proteins are overexpressed in cancer cells, and they prevent the functioning of the caspase cascade by complexing with these enzymes. Designed SMAC mimetics^{19–21} can bind to IAP and prevent it from interacting with the caspases, thereby allowing them to carry out proteolytic degradation and cell death. The p53–MDM2^{22–24} interaction has also been widely targeted for the development of apoptosis-based anticancer therapies. The proapoptotic p53 is sequestered in cancer cells by its negative regulator MDM2, which is overexpressed in cancer cells. Mimics of p53, which bind to MDM2, would disrupt the p53–MDM2 interaction^{25,26} leading to the release of free p53, resulting in the induction of apoptosis.

Protein–Protein Interactions as Targets. Protein–protein interaction interfaces are usually shallow, solvent exposed, and relatively featureless, making it difficult to find

* Corresponding author phone: 662-915-5879; fax: 662-915-5638; e-mail: mavery@olemiss.edu.

[†] Current address: Novartis Institutes for Biomedical Research, 4560 Horton Street, Emeryville, CA 94608.

[‡] Current address: Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285.

anchor points for ligands. In certain cases, the proteins undergo pronounced induced-fit effects from their apo form, forming unique interaction interfaces in the presence of specific protein binding partners²⁷ (or small molecule inhibitors), making it difficult to make a priori prediction of these changes. Together, these factors make the discovery of a protein–protein interaction inhibitor^{28–30} through structure-based approaches a very formidable task. Another important consideration from the structural aspects of a protein–protein interaction is the location of “hot-spots”^{31–33} on the protein–protein interface, where protein residues belonging to either of the protein partners are involved in a strong, energetically favorable intersubunit interaction and may be supplanted by small molecule interactions. While it is not essential to utilize all of the potential “hot-spots”, one should utilize two important criteria, targeting “hot-spots” providing the: (i) maximum contributions to the binding affinity and (ii) those which are located at close proximity to each other to balance the seemingly opposing goals of potency and drug likeness.

Targeting the Bcl-2 Protein Family. As discussed previously, the Bcl-2 protein family act as regulators of apoptosis. So far, 25 Bcl-2 family members³⁴ have been identified, which can be functionally classified into two groups, pro and antiapoptotic. The proapoptotic Bcl-2 family members include Bak, Bax, Bid, and Bim, while the antiapoptotic family members are exemplified by Bcl-2, Bcl-xL, Bcl-w, Mcl-1, and Ced-9 from *C. elegans*. The members can be further classified based on the presence of Bcl-2 homology (BH) domains in the protein. All of the antiapoptotic and some of the proapoptotic Bcl-2 family members are characterized by the presence of sequence homology in three BH domains BH1, BH2, and BH3. Among the proapoptotic proteins some members show sequence homology only in the BH3 domain and are classified as BH3-only proteins. The BH3 domain is critical for the protein–protein interaction between a pro and antiapoptotic member. The BH3 domain of a proapoptotic protein binds into a groove formed by the BH1, BH2, and BH3 domain of the antiapoptotic member. The proapoptotic members possessing three BH domains, such as Bak and Bax, undergo relocation and oligomerization^{35–38} on the mitochondrial surface, leading to mitochondrial membrane permeation (MMP) followed by the release of cytochrome c^{39,40} from the mitochondria and the induction of the caspase cascade resulting in cell death. The BH3-only proapoptotic proteins⁴¹ lack the capability for inducing MMP and function by binding to the antiapoptotic members or inducing the oligomerization of Bak or Bax.

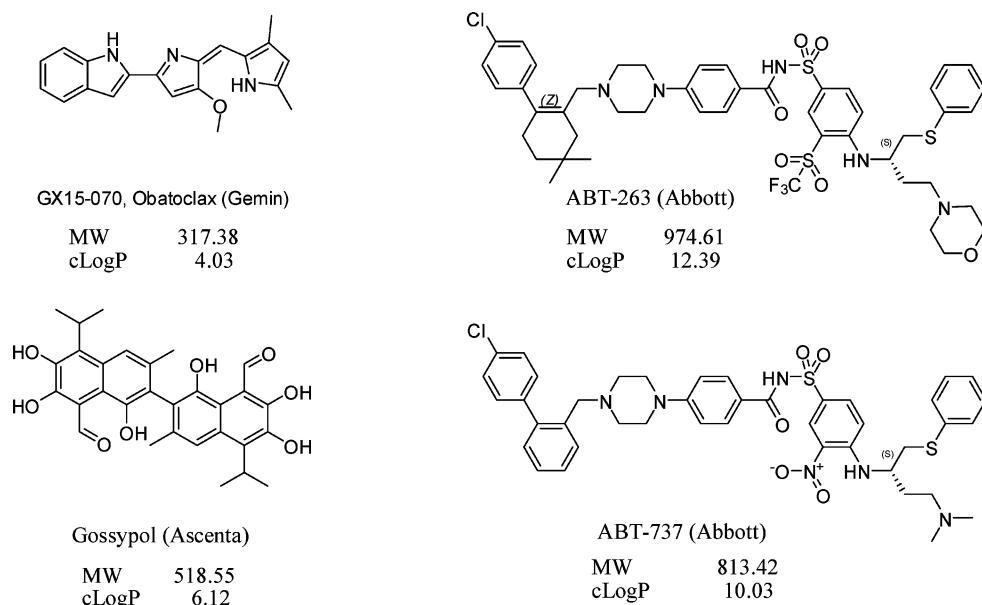
The antiapoptotic Bcl-2 family members are considered as attractive targets⁴² for the development of anticancer therapy. They have been shown to overexpress in a number of cancers including chronic lymphocytic leukemia (CLL), multiple myeloma, human follicular lymphoma, small cell lung carcinoma (SCLC), cervical, ovarian, bladder, prostate, pancreatic, gastric, colorectal, and breast cancers. In most of these cases, overexpression of Bcl-2 or Bcl-xL have been shown to correlate with the progression of the disease and with poor survival rate. Overexpression of Bcl-2 and Bcl-xL has been also associated with resistance¹⁴ development against standard chemotherapeutic agents and radiotherapy. A study⁴³ on the cancer cell resistance to a panel of 122 standard chemotherapeutic agents suggested a strong cor-

relation for the overexpression of Bcl-xL with resistance development. Reduction in levels of Bcl-2 has demonstrated higher sensitivity⁴⁴ for anticancer drugs and enhancement in survival rates in *in vivo*⁴⁵ experiments. These evidences clearly establish the antiapoptotic Bcl-2 proteins as relevant targets for anticancer drug development.

One of the approaches utilized for the targeting of the antiapoptotic Bcl-2 protein is to introduce antisense oligonucleotides⁴⁴ which would lead to the reduced expression levels of these proteins. A prominent example is the antisense oligonucleotide Oblimersen⁴⁶ designed against Bcl-2, which has been advanced to phase-III clinical trials for the treatment of myeloma, multiple myeloma, and CLL. Another approach⁴⁷ utilizes synthetic BH3 domain peptides complexed to active transporters, such as the antennapedia homeoprotein internalization domain for facilitating cell entry. These peptides would enter into the cell and bind to the proapoptotic Bcl-2 family members. Bid-based BH3 peptides have been shown to be effective in the human leukemia cell lines and prolonged survival in the mice leukemia model. The most ideal approach to the development of Bcl-2 inhibitors focuses on the design of small molecule inhibitors (Figure 1) which would mimic the BH3 peptides and bind into the groove of the antiapoptotic Bcl-2 family members. Several approaches, such as high-throughput screening,^{48,49} structure-based virtual screening,^{50,51} and protein NMR screening^{52,53} have been utilized in the discovery of small molecule inhibitors binding to Bcl-2, Bcl-xL, Bcl-w, and Mcl-1. Small molecules, such as Chelerythrine,⁴⁸ Antimycin,⁵⁴ Gossypol,⁵² Epicatechin Gallate,⁵⁵ YC-137,⁵⁶ TW-37,⁵⁷ HA14-1,⁵¹ Obatoclax,⁵⁸ and BH3I^{49,59} have been found to induce apoptosis in cells. Recently ABT-737^{53,60,61} and ABT-263,⁶² subnanomolar Bcl-2, Bcl-xL, and Bcl-w inhibitors, have been discovered by the scientists at Abbott and are in preclinical and early phase-I development. The molecules are capable of single-agent mechanism-based killing in lymphoma and leukemia cell lines. It shows synergistic cytotoxicity with other chemotherapeutic agents or radiation in various cell lines. These developments provide the proof of concept for use of the antiapoptotic Bcl-2 proteins as targets for anticancer drug discovery.

High-throughput screening^{10,63,64} of synthetic and natural product libraries has been the industry standard for the discovery of novel hits for therapeutic targets of relevance. However, with the increasing size of these compound libraries, a greater pressure is levied on the manual and material resources making the exercise economically unviable. *In-silico* or virtual screening^{65–69} of compound libraries provides an interesting alternative to this problem. Large compound libraries in the range of millions of compounds can be rationally screened to a small subset which is then utilized in a biological assay to identify hits. This cost-effective alternative has a high throughput and is gaining significant interest in the industry as a means to discovering novel leads for drug discovery. Based on the knowledge of the target and the desired throughput, one may conduct virtual screening using various techniques, such as similarity/substructure searching, pharmacophore searching, and protein structure-based docking or using a combination of one or more of these techniques.

Identification of hits against the Bcl-2 protein family have primarily focused on the use of methodologies, such as high-

**Figure 1.** Small molecule inhibitors of Bcl-xL.

throughput screening of synthetic/natural product libraries and protein NMR based fragment screening. The initial hits identified from these screens have ranged in potency from low micromolar up to the millimolar range. Structure-based modification on the identified hits have led to the identification of highly potent compounds. The discovery of ABT-737 is a prominent example where a fragment with $300\text{ }\mu\text{M}$ (K_d) affinity discovered through NMR screening was developed into the preclinical candidate with subnanomolar affinity using a combination of structural and medicinal chemistry based optimization schemes. However, there are only a few examples for the discovery of hits against the Bcl-2 protein family using *in silico* screening. To the best of our knowledge, only two efforts^{50,51} have been reported so far on the identification of hits against the Bcl-2 protein family using structure-based virtual screening. We, therefore, decided to conduct a dual pharmacophore and docking based virtual screening against one of the Bcl-2 family members, Bcl-xL, to identify novel hits.

METHODS

Computational Resources. Calculations were performed on an 8 processor SGI Origin 350 server equipped with R16000 chipsets and 2GB of memory. All docking calculations were carried in Gold 3.0.1 (CCDC, Cambridge, U.K.). Rescoring using multiple scoring functions was carried out in the Cscore module of Sybyl 6.9 (Tripos Inc., St. Louis, MO). Docking pose based descriptors (DPBDs) were generated using Silver1.1 (CCDC, Cambridge, U.K.). The receptor-based pharmacophore prefilter was generated using Catalyst 4.6 (Accelrys Inc., San Diego, CA). The dbselect and dbtranslate utility from the Unity module of Sybyl6.9 were utilized for file format conversions and selection of the enrichment test set decoys. The Descriptor+ module of Cerius2 (Accelrys Inc., San Diego, CA) was also utilized for the selection of the enrichment set decoys. Macromodel v9.5 suite from Schrödinger (Portland, OR) was used for the eMBRACE calculations. Scripts used for range scaling calculations and evaluation of Silver descriptor outputs were written using Python 2.4.

Protein Preparation. Two NMR-derived structures of Bcl-xL, 1BXL and 1YSI (PDB codes), were utilized in the virtual screening. In case of 1BXL, Bcl-xL is complexed with a 16-membered peptide derived from the Bak BH3 domain, while in case of 1YSI, the protein is complexed with a small molecule inhibitor N3B. The complexes are minimized, average structures derived from solution phase NMR. Since the reported structures have hydrogens added, the reported protonation and tautomeric states of the residues were preserved. To prepare the proteins for docking the Bak peptide (1BXL) and the ligand N3B (1YSI) were extracted from the relevant complexes. The protein structures were saved in the .mol2 format and atom- and bond-type misrepresentations were corrected to generate the final protein for docking.

Ligand and Screening Database Preparation. The N3B ligand from the structure 1YSI was utilized in docking pose validation studies. The original ligand coordinates were extracted from the complex and saved in the .mol2 format after atom- and bond-type corrections. The ligands (see Supporting Information, Table 1) utilized in the docking score-based separation and enrichment studies were sketched in Sybyl 6.9 and minimized using a conjugate gradient to 0.01 kcal/mol-Å, applying the Gasteiger-Hückel partial-charge method and the Tripos force field. To generate the enrichment test set, a group of decoys were selected from the ZINC ‘big-n-greasy’ subset. A set of molecules were chosen based on the cutoffs of physiochemical properties $300 \leq \text{molecular weight} \leq 600$, $4 \leq \text{hydrogen-bond acceptor} \leq 10$, $1 \leq \text{hydrogen-bond donor} \leq 7$, $1 \leq \text{clogP} \leq 8$, and $5 \leq \text{rotatable bonds} \leq 13$. While the molecular weight and clogP filters were applied using the dbselect utility of Sybyl 6.9, the hydrogen-bond acceptor and donor and the rotatable bond filters were applied using the Descriptor+ module of Cerius2. Finally, 1024 molecules were selected from a randomized distribution of the filtered set.

The actual screening was conducted using the ZINC ‘big-n-greasy’ subset (1.8 million compounds) which was downloaded from <http://zinc.docking.org/> in the .sd format (38 segments). The CatDB server was utilized to generate a

multiconformer catalyst database. In addition to the structure files, the database configuration files (.bdb) serve as inputs for the actual database generation and were, therefore, generated and configured beforehand in an automated fashion utilizing a shell script. These files contained the hard drive paths for the structure files and two-dimensional (2D) and 3D indices as well as the feature dictionary. The actual database creation was carried out using the CatDB program, handled through a shell script written for this purpose. Each molecule in the database was assigned a unique “cref” number and enumerated by 100 conformers generated using the “fast” mode. The database generation was carried out in 38 segments to facilitate data handling and other operations. The prepared database was then submitted for virtual screening using the Cat search background utility. The filtered output from the Cat search program was obtained in the .sd format and was subsequently converted to the .mol2 format using the dbtranslate utility of Sybyl 6.9, for use in Gold docking. A similar catalyst database was generated for the enrichment test set and utilized in the validation of the receptor-based pharmacophore model.

Receptor-Based Pharmacophore Model. The receptor-based pharmacophore model was generated in Catalyst using the 1BXL structure. Variations in the tolerance values of the hydrophobic features and the location of the exclusion spheres were varied intuitively based on the selectivity and enrichment obtained from the enrichment study. The Cat search background utility was utilized for the enrichment study and the fast method was utilized to study the fitting of the relevant molecules to the pharmacophore. The pharmacophore selected through this study consisted of four hydrophobic occupancy features defined around the geometric centroids of the relevant hydrophobic side chains of the Bak peptide. The tolerance values of the hydrophobic features were varied to attain optimal enrichment. Twenty-three exclusion spheres were added based on the position of the residues constituting the binding site. The exclusion spheres were defined using either the C- α atoms of the relevant amino acids or the geometric centroid of the relevant side chains. Molecules possessing ≥ 3 of the hydrophobic features were considered to have passed the filter.

Docking Pose Based Descriptors. DPBDs were generated for the structures 1BXL and 1YSI using the Silver program. In each case, the descriptor set consisted of four hydrophobic occupancy descriptors, one hydrogen-bond acceptor descriptor and one buried hydrophobic surface area descriptor. The Bak peptide and the bound ligand **N3B** were utilized in the definition of the descriptors in case of 1BXL and 1YSI, respectively. The hydrophobic occupancy descriptors were defined as 2 Å spheres around the geometric centroid of the relevant ligand hydrophobic groups. The descriptor reports the percentage occupancy of the spherical volume by the hydrophobic atoms of the ligands docked pose. An occupancy of 70% or greater was considered as fulfillment of the descriptors. The hydrogen-bond acceptor feature was also defined as a sphere around the sulfone oxygen atom of **N3B** in case of 1YSI, while it was intuitively placed in case of 1BXL and coded for absence or presence of such a feature. The buried hydrophobic surface area descriptor was defined using the ligand hydrophobic atoms of **N3B** from the 1YSI structure. A value corresponding to 70% or more of the

ligand buried hydrophobic surface area of the **N3B** ligand was considered as fulfillment of this feature. The same definition was carried over for the 1BXL descriptor. The fulfillment of a particular feature was given a Boolean value of 1, while a lack of fulfillment was described by a value of 0. Thus any given pose could be assigned a descriptor-based score ranging from 0 to 6. In this calculation, the Silver output was obtained as a comma delimited text file which was then processed by a python script to generate the final score.

Docking Protocols for Validation and Screening. The docking protocols for both the structures 1BXL and 1YSI were generated using the Gold docking program. Score-based active/inactive separation and docking score vs activity correlation studies were conducted for the 1BXL structure, wherein the variations in the definition of the binding site, radius of cavity detection, and scoring functions utilized in the calculation of the docking poses were evaluated. Additionally, various combinations of hydrophobic constraints were also utilized in the docking calculations and were defined as spheres of 2 Å radius around geometric centroid of side chain heavy atoms of the Bak peptide pose. Seven congeneric series of Bcl-xL actives (Supporting Information) with a range of biological activity were selected for this study. In case of the 1YSI, a pose validation study was conducted using the bound pose of the **N3B** ligand. In this case as well, variations in the definition of hydrophobic constraints, radius of cavity detection, binding site definition, and scoring function for docking pose calculations were conducted.

$$EF_n = \frac{\frac{Hits_n}{Hits_{tot}}}{\frac{Lig_n}{Lig_{tot}}} \quad (1)$$

$$\text{Selectivity}_n = \frac{Lig_n}{Lig_{tot}} \quad (2)$$

where,

$Hits_n$ = Hits recovered at threshold n .

Lig_n = Total ligands recovered at threshold n .

$Hits_{tot}$ = Total number of hits.

Lig_{tot} = Total number of ligands.

The formalism above was used to calculate the enrichment factor (EF) and the selectivity in the validation studies. The enrichment studies were conducted for both 1BXL and 1YSI in triplicate using the same enrichment test set utilized in the receptor-based pharmacophore enrichment study. Rescoring of the top-ranked poses obtained through docking was conducted using Gold Score (GL) as well as the five scoring functions Fscore (F), Gscore (G), PMF score (PMF), Dscore (D), and Chemscore (C) available from the Cscore module of Sybyl 6.9. Combinations of scores obtained from different scoring functions were conducted after range scaling using the equation described below. In the cross-docking phase of enrichment studies, the scores obtained from the docking pose based descriptors were also combined with scores obtained from the scoring function(s) after range scaling.

$$S_{n(\text{Range scaled})} = \sum \frac{S_n - S_{\min}}{S_{\max} - S_{\min}} \quad (3)$$

For eMBRACE calculations in the difference mode, the following equation was used

$$\Delta E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}} \quad (4)$$

where,

E_{complex} = Energy of the protein–ligand complex using a generalized Born surface area (GBSA) implicit solvent treatment.

E_{ligand} = Energy of the free ligand using a GBSA implicit solvent treatment.

E_{protein} = Energy of the free protein using a GBSA implicit solvent treatment.

Biological Evaluation. The C-terminal truncated His tagged Bcl-xL protein (residues 1–212, purity > 95% by SDS-PAGE) was obtained from RnDsystems (Minneapolis, MN). The fluorescein-labeled Bad peptide, F-Bad (Ac-NLWAAQRYGRELRRMSDK (fluorescein)FVD-OH, purity ≥ 95% by HPLC) was also obtained from Calbiochem (San Diego, CA). A 20 mM phosphate buffer [0.015 M Na₂HPO₄ and 0.004 M Na₂HPO₄ in ddH₂O] (pH 7.4), 1 mM ethylenediaminetetraacetic acid (EDTA), 0.05 M NaCl, and 0.05% v/v pluronic-F68 (Invitrogen, Carlsbad, CA) in ddH₂O] was prepared and used in the assay. All assays were carried out in 96-well, black, flat-bottom plates with a nonbinding surface (NBS) available from Corning (Acton, MA). A 16-residue peptide from the Bak BH3 domain (GQVGRQLAIIGDDINR, Res 72–87) (RnD systems, Minneapolis, MN) was used as a positive control. The F-Bad peptide was reconstituted as a 1 mM solution in dimethyl sulfoxide (DMSO), aliquoted, and stored at –20 °C. The Bak peptide was similarly reconstituted as a 500 μM solution in DMSO and stored at –20 °C. Graph Pad Prism 4.0 was utilized for all calculations including K_d and IC₅₀ determinations. All readings were obtained in a BMG polarstar 96-well plate reader (BMG Labtech, Durham, NC). The calculated K_i values were generated using an online K_i calculator^{70,71} for fluorescence-based competitive binding assay.

To determine the linear range for the F-Bad peptide, a set of dilutions were made in assay buffer. 125 μL of each dilution was added in triplicate to a 96 well plate and readings were taken at an excitation = 485nM and an emission = 530 nM. The optimal linear range for the parallel and perpendicular fluorescence intensities ranged from 10 nM to 30 μM (Supporting Information, Figure 1). However, the optimal linear range for the polarization signal where the values did not show much deviation extended from 10 nM to 10 μM (Supporting Information, Figure 1).

The next critical criterion of evaluation was the amount of protein and fluorescent substrate that was present in the assay conditions. A low concentration of the substrate would lead to a poor signal, while a high concentration would make it difficult to identify weaker inhibitors. Similarly, a low concentration of the protein would bias the assay against the detection of weaker inhibitors. On the other hand, a high concentration of the protein would put constraints on the economic viability of the assay. Since 10 nM is the lowest concentration the fluorescent substrate shows a linear signal,

two concentrations 10 and 20 nM were chosen for conducting the experiment. To determine the K_d of F-Bad interactions with Bcl-xL, two 2.5X F-Bad dilutions (10 and 20 nM final concentration) and seven 5X Bcl-xL dilutions (1, 3, 10, 30, 100, 300, and 600 nM final concentration) were prepared in assay buffer. To each well of a 96-well plate, 50 μL of F-Bad, 25 μL of Bcl-xL, and 50 μL of assay buffer (125 μL final volume) were added, and readings were taken in triplicate after a 30 min of incubation. Saturation binding curves (Supporting Information, Figure 2) were plotted to evaluate the dynamics range of the binding as well as the K_d for the F-Bad/Bcl-xL interaction. The dynamic range for the binding response was ~133 mP units. The K_d value for the binding was found to be 3.08 and 4.19 nM for 10 and 20 nM of the F-Bad peptide, respectively. On the basis of this experiment, a 20 nM concentration of Bcl-xL and a 10 nM concentration of the F-Bad peptide were selected for the assay. At this concentration, approximately 80% of the dynamic range could be used for detecting the inhibitory profile of the tested compounds.

For the Bak peptide (positive control) competition/inhibition assay 25 μL of 100 nM (20 nM final concentration) Bcl-xL was added to each well of a 96-well plate. Bak peptide dilutions (2.5 × final concentration) were prepared in assay buffer from a 10 mM DMSO stock. In triplicate, 50 μL of peptide dilutions was added, and the plate incubated for a period of 30 min. After incubation, 50 μL of 25 nM F-Bad (10 nM final concentration) was added to each well to make the final volume of 125 μL. After an incubation of 30 min, the plates were read using excitation and emission filters of 485 and 530 nM, respectively. The percentage inhibition for a given compound was calculated as

$$\% \text{Inhibition} = \left[1 - \frac{(mP - mP_f)}{(mP_b - mP_f)} \right] \times 100 \quad (5)$$

where,

mP = Polarization value from the compound well.

mP_f = Polarization value for the free peptide.

mP_b = Polarization value for the bound peptide.

Polarization was calculated as

$$\text{Polarization}[mP] = \frac{(F_{\parallel} - K \times F_{\perp})}{(F_{\parallel} + K \times F_{\perp})} \quad (6)$$

where,

F_{\parallel} = Parallel fluorescence intensity.

F_{\perp} = Perpendicular fluorescence intensity.

K = K factor.

F-Bad showed an IC₅₀ value of 1.1 μM (Supporting Information, Figure 3) against the F-Bad/Bcl-xL inhibition. The calculated K_i value for this peptide was 0.16 μM, which is close to the reported value of 0.34 μM.⁷² Compounds were evaluated at three single-point concentrations of 10, 100, and 1000 μM using 20 nM of Bcl-xL and 10 nM of F-Bad peptides. Compounds showing dose-related inhibitory activity were carried forward to a detailed dose response study using quarter log dilutions of the compound.

RESULTS AND DISCUSSION

Binding Site Interactions of Bcl-xL. The protein structural topology of the Bcl-xL protein can be described as two

hydrophobic helices surrounded by five amphiphilic helices. The binding groove of the Bcl-xL protein is a shallow, solvent exposed, and primarily hydrophobic cavity (Figure 2 a and c). It is formed by the side chains originating from the BH (Bcl-2 homology) 1, BH2, and the BH3 domains of the protein. In the apo form, the cavity is in a collapsed state, and, in the event of a heterodimerization with a proapoptotic partner, such as Bak or Bad, the protein undergoes pronounced induced-fit effects in order to form the cavity. The BH3 domain of the proapoptotic protein is thought to be important for the induction of apoptosis. Coincidentally, the BH3 domain of the proapoptotic member inserts into the groove formed by the antiapoptotic member and is critical to the heterodimerization process. Relatively small peptides derived from the BH3 domain of the proapoptotic members are almost equally functional as the whole protein in heterodimerizing with the antiapoptotic members.

The NMR-derived structure⁷² of a 16 amino acid peptide derived from the Bak BH3 domain complexed with Bcl-xL (1BXL, PDB code) (Figure 2 a and b) provides critical insights into the binding site interactions. The Bak-derived peptide forms an amphiphilic helix which binds into the hydrophobic groove of Bcl-xL. An alanine scan study⁷² (Figure 2d) of the 16 residue peptide helped in the identification of amino acids, which upon mutation led to a loss in binding affinity of the peptide. Four hydrophobic residues Val74, Leu78, Ile81, and Ile85 which are located on one side of the amphiphilic helix and point into the hydrophobic groove were found to cause maximum loss in binding affinity. Among these residues, the Leu78Ala mutation produced ~3 log order loss in binding affinity, while the others produced ~2 order loss of affinity. The hydrophobic binding groove in the 1BXL structure can be widely subdivided into the lower and upper regions (Figure 2e). The Val74 residue is located in the lower groove, while the Ile81 and Ile85 residues are located in the upper hydrophobic groove. The Leu78 residue is located at the interface of the upper and lower hydrophobic grooves. The upper hydrophobic groove is the wider of the two and is made up of the residues Leu90, Ala93, Gly94, Val141, and Tyr195. The lower hydrophobic groove is comparatively narrower and is made up of the residues Phe97, Tyr101, Phe105, Leu108, Val126, Leu130, Ala142, and Phe146. Apart from the four hydrophobic Bak residues, the Asp83Ala mutation also shows a loss in binding affinity. This residue is seen to form a moderate electrostatic interaction with Arg139 of Bcl-xL. Although the binding groove of Bcl-xL is primarily hydrophobic, there are polar residues forming the walls of the binding pocket. While most of the residues point into the solvent, two residues Arg139 and Glu96 are accessible for interaction with the binding ligand.

The bound pose of **N3B** in the ligand bound structure of Bcl-xL (1YSI, PDB code)⁶¹ (Figure 2f) shows a similar pattern of nonbonded interactions, as seen in the 1BXL structure. The prominent hydrophobic features of this ligand include four phenyl rings which surrogate for the four hydrophobic side chains of the Bak peptide. The biphenyl system lies over the lower hydrophobic groove and surrogates for the Leu78 and the Val74 interactions. The nitro-substituted phenyl group substitutes for the Ile85 interaction, while the terminal phenyl ring substitutes for the Ile81 interaction. The two phenyl rings in the upper portion of

the ligand form $\pi-\pi$ stacking interactions with Tyr195 and Tyr101, respectively, as well as an internal $\pi-\pi$ stacking interaction within themselves. The sulfone oxygen is in close proximity of the hydroxyl group of Tyr195 and the Arg139 side chain. It may surrogate for the Asp83 interaction of the Bak peptide. The carbonyl oxygen may form a hydrogen bond with the hydroxyl group of Tyr101.

The Bcl-xL protein in the *apo* form does not present a binding groove and undergoes pronounced structural changes in the presence of an appropriate ligand to form this groove and accommodate for ligand binding. Comparison of the two ligand bound structures suggests that the binding of the two structurally different ligands produce varying induced-fit effects (Figure 2g and h) in the formation of the binding groove. In case of the 1BXL structure, the binding of a larger peptidic ligand induces a larger opening of the hydrophobic groove, which is wider than the groove present in 1YSI. In case of 1YSI, the upper and the lower portions of the groove are clearly demarcated by the presence of the Phe97 residue, and the individual cavities are well-defined. On the other hand, in case of 1YSI, the demarcation between the upper and the lower regions is less prominent, and each of the cavities is shallower. The 1BXL groove spans up to 12 Å at its widest, while the 1YSI groove only spans up to a maximum of 6 Å. Two regions of the binding site show significant deviations of the protein backbone. The C-terminal region of the Bcl-xL protein forms a portion of the upper binding groove. In case of 1BXL, the C-terminal region comes closer to the binding pocket, and the residue Glu92 is not accessible from the binding site. On the other hand, the same residue is accessible in case of 1YSI where the C-terminal region is located further away from the binding site. The residue Tyr195, which is located at the beginning of this terminal region, has different locations of the C α atoms and, consequently, their side chains between the two structures. The other prominent backbone movement occurs in the α -helical region of the BH3 domain. In case of 1BXL, the region is made up of three turns of a helix, while in case of 1YSI the lowest turn and part of the second lowest turn are unstructured. Two important hydrophobic residues located on this turn assume significantly different positions in the two structures. The Tyr101 residue is located on the second lowest turn and, in case of 1BXL, points toward the lower hydrophobic groove forming the floor of the groove. In case of 1YSI, the residue shows significant C α deviation and points upward forming the side walls of the binding groove. The Tyr101 residue forms a side chain hydrogen-bonding interaction with the ligand in case of 1YSI, which is not possible in case of 1BXL due to a completely different orientation. The Phe105 residue is located on the lowest turn in case of 1BXL and showed significant deviation in the C α position and the side chain orientation from 1YSI, due to the unwinding of the helical region in the case of the 1YSI structure. Apart from these three residues where the differences are most noticeable, movements of lesser degree are noticed in the case of the residues Phe97, Leu112, and Arg139. The binding groove of Bcl-xL shows a breathing action as a response to ligand binding, and there might be many more energetically accessible binding site configurations than what is seen from the experimentally determined structures so far. The factor is further complicated in the presence of two competing ligands, such as in a competitive

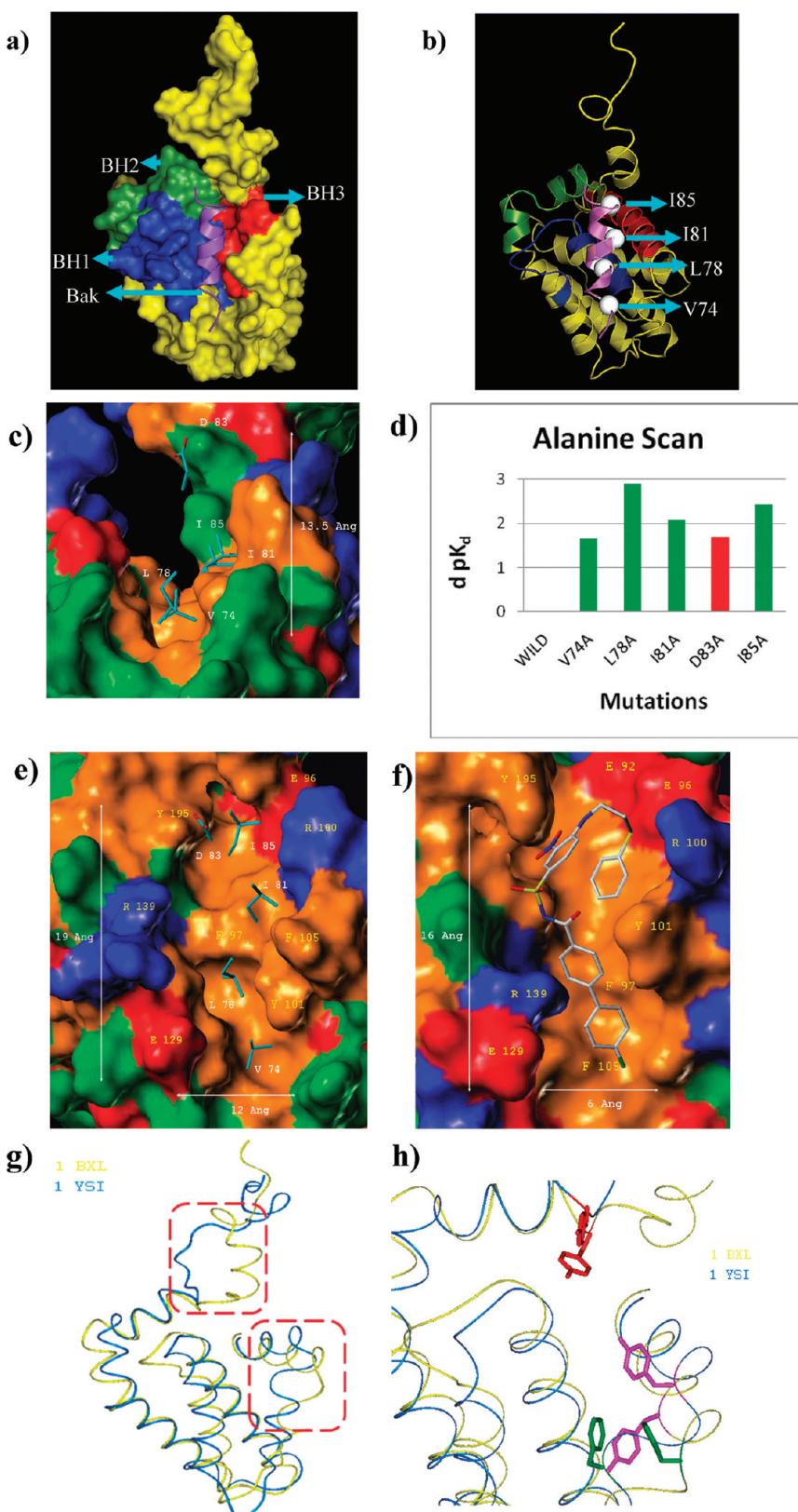


Figure 2. (a) A surface representation of the 16-residue Bak peptide (magenta, ribbon representation) bound to the binding groove of Bcl-xL (1BXL, PDB code) formed by the BH1 (blue), BH2 (green), and BH3 (red) domains of Bcl-xL. (b) A ribbon representation of the Bak/Bcl-xL complex showing the hydrophobic residues (white CPK) which point into the groove and form binding interactions. (c) A side view of the Bcl-xL protein showing the solvent exposed binding groove. (d) Results of the alanine scan study showing the loss in K_d relative to the wild-type peptide. The bars in green represent buried hydrophobic residues, while those in red represent solvent exposed polar residues. Comparison of the (e) 1BXL and (f) 1YSI structures showing the similarity of the binding interactions for the Bak peptide residues (cyan) and the acylsulphonamide ligand as well as the key protein residues (green circles), which are oriented differently in the two structures. (g) A representation of the pronounced induced-fit effects present between the two structures (yellow: 1BXL, cyan: 1YSI). The red dotted rectangles represent the main domain movements present at the C-terminus and the BH3 domain. (h) Inset from the previous picture showing the most pronounced side chain movements showed by Tyr101 (magenta), Phe105 (green), and Tyr195 (red).

binding assay, wherein multiple instances of these breathing processes might be occurring simultaneously.

Screening Strategy. A number of different factors make Bcl-xL a very challenging target for structure-based drug design efforts. The binding groove of the Bcl-xL protein is very wide and shallow and does not present a well-defined shape complementarity, as seen in some other drug targets. From the viewpoint of a docking program, this structural feature of the binding site gets transformed into a sampling problem. In a more buried and well-defined site, the conformational space required to be sampled in order to attain the putative bioactive conformation within the sampled ensemble is much less, as compared to what is required in this case. Therefore, the sampling capability of the docking protocol might become a limiting factor in the determination of the putative binding conformation. The interaction sites within the binding groove are primarily hydrophobic and have very few polar interactions, such as hydrogen bonding. Presence of directional interactions, such as hydrogen bonding, helps alleviate the sampling problem to a degree vis-à-vis hydrophobic interactions, which are primarily nondirectional in nature and are not limited to individual polar atom features, such as a hydrogen-bond acceptor or donor. In addition to these, the protein shows a pronounced degree of induced-fit effect upon ligand binding. The 1YSI structure is more conducive toward the acylsulfonamide series of ligands but does not explain the binding of the some of the other Bcl-xL inhibitors of differing chemotypes. 1BXL is the more liberal of the two structures, but the lack of shape complementarity leads to a lower retrieval in enrichment studies as compared to the 1YSI structure. Apart from the protein target, the physiochemical nature of the Bcl-xL inhibitors presents certain issues which need to be handled appropriately. A major portion of the Bcl-xL inhibitors discovered until date are primarily hydrophobic and are high in molecular weight. The clogP and molecular weight of some of the most potent inhibitors such as ABT-737⁶¹ reach up to 10 and 800, respectively. The binding site of the protein is large and hydrophobic, and hence, molecules with these physiochemical property ranges seem to favor the binding affinity.

The above-mentioned factors were suitably dealt with while devising the virtual-screening strategy. Docking of large virtual screening databases is a computationally expensive exercise, and the level of calculations utilized for the docking process is a critical factor governing the computational time required for the screening process. One approach we have successfully used previously^{73–75} toward minimizing the overall down time without compromising the quality is to carry out a multistage cascade docking involving a less intensive and faster docking protocol at the early stages of the screening and a gradual increment in the computational complexity as we proceed toward the final stages. Typically, at the early stages of screening, the aim is to look for shape complementarity of the docked molecules against the binding site and for elimination of molecules with low probability of attaining a proper binding pose within the active site. On the contrary, the final stage of the screening process involves the selection of deserving hits based on their complementarity with the binding site interactions and the geometrical qualities of the binding pose itself. While the elimination stage can be handled by a less intensive docking calculation, a more

rigorous protocol incorporating exhaustive conformational sampling is a critical requirement for the final stages of the screening. 1BXL is the more liberal of the two structures and explains the binding inhibitors from various chemotypes. Therefore, it was used in the first stage of the cascade docking. If we use the more restrictive 1YSI structure, then there is a possibility of elimination of ligands at the early stage which might have bound to the more open 1BXL-type groove of the Bcl-xL protein. The molecules filtered through this step would ultimately be put through the more rigorous filtering of the cross-docking stage. Hence, this provides us an opportunity to go back a step in search of varying chemical scaffolds without having to repeat the virtual screening with the full set of ligands.

Induced-fit effects may be exhibited by targets to different degrees, and although a number of methods,⁷⁶ such as molecular dynamics (MD) simulations^{77–80} as well as induced-fit docking (IFD),⁸¹ have been utilized to explain such phenomenon, they are still computationally intensive and cannot be incorporated into VS workflows. Naturally, VS approaches which aim to explain induced-fit effects do so only to a degree, while preserving the orthogonal aim of throughput. Popular VS capable methods for explaining induced-fit effects include^{82,83} scaling of the van der Waals radii or deleting side chains to increase the cavity size and cross-docking⁸⁴ into an ensemble of protein structures capturing the various conformational states of the protein (including those generated through MD^{85,86} and IFD).⁸⁷ While scaling and side chain deletion can compensate for ligand–protein clashes generated from limited side chain movement of residues, they also may introduce artifacts into the system and generate poses which are physiologically irrelevant. The pronounced movements exhibited by Bcl-xL (side chain and backbone movements) suggested that cross-docking involving two structurally distinct and experimentally determined halo structures of the protein (1BXL, 1YSI) would provide coverage of this phenomenon to the largest degree. Furthermore, molecules which are highly ranked in both the protein structures and have similar binding poses would present a higher probability of being active against the target. The sampling problem within the binding site was taken care of by incorporating pharmacophore-based docking constraints which would guide the pose sampling and direct the search toward poses with pharmacophorically relevant interaction patterns. It was also decided to use the ‘big-n-greasy’ subset from the ZINC database.⁸⁸ As the name suggests, this subset represents a set of molecules with $2 < \text{clogP} < 6$ and $300 < \text{molwt} < 600$. This subset was chosen with the viewpoint that there is a higher probability of finding actives for this particular target within compounds with high-molecular weight and lipophilicity. Another factor that had to be taken into consideration was the fact that the enrichments obtained during the validation stage of the study for these two proteins were not very high. To complement this, additional scoring terms devised through docking pose based descriptors were incorporated into the ranking strategy. These additional knowledge-based filters helped in boosting the overall performance during the validation studies.

Screening Protocol and Validation Studies. The actual virtual screening workflow involved a multistep approach starting with a receptor-based pharmacophore prefilter, a

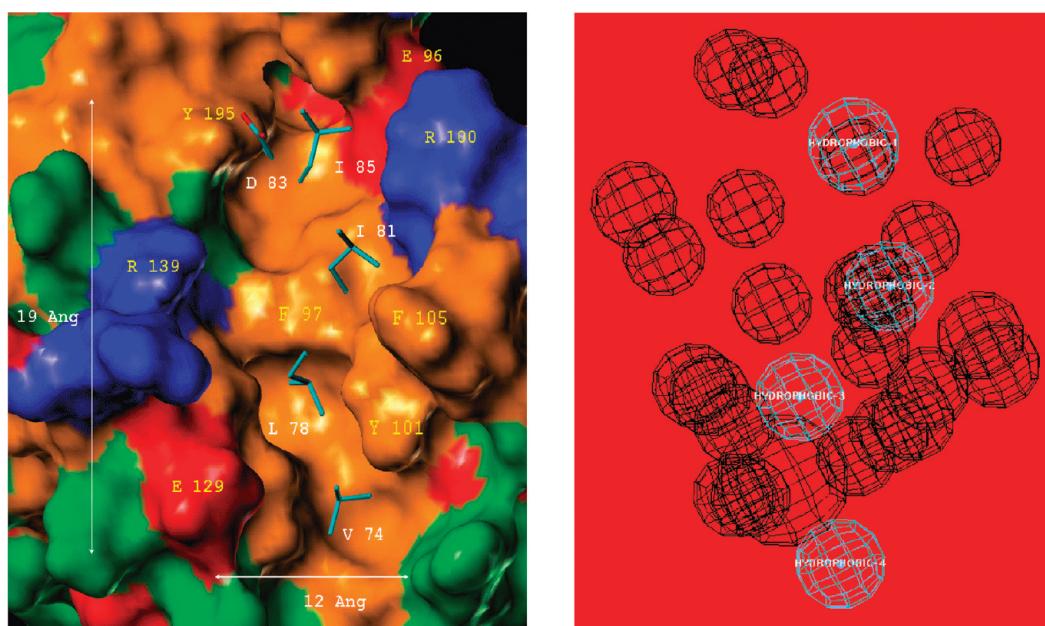


Figure 3. The receptor-based pharmacophore (on right) generated from the Bak/Bcl-xL complex (1BXL, PDB code) (on left). The cyan spheres represent the hydrophobic features defined using the side chains of the four hydrophobic residues (Val 74, Leu 78, Ile81, and Ile85) from the BAK peptide. The exclusion spheres are shown in black and represent the positions of the Bcl-xL residues.

cascade docking approach utilizing the 1BXL structure and culminating in a cross-docking step involving the two protein structures 1BXL and 1YSI. Validation studies were performed for each of the steps of this virtual screening workflow.

An enrichment test set (details in Methods Section) was generated and utilized for a majority of these validation studies. The scoring functions utilized in the evaluation of the binding poses of the ligands are either force field, empirical, or knowledge based in nature. In essence, they provide a crude approximation of a relevant biophysical property of the ligand binding process, such as the ligand binding affinity. The function is fitted into experimentally determined binding data using a statistical method, such as regression. The purpose of the scoring function is not only to produce an accurate estimation of the binding data but also to provide a computationally inexpensive approximation of the actual data, which could be utilized in a high-throughput scenario. Being an approximation, they are often marked by various pitfalls, including biases toward various physiochemical properties of the ligands. One of the parameters, molecular weight is often seen to positively correlate with the scores obtained for the ligand. Therefore, the scores obtained for the poses tend to depend on the physiochemical properties instead of the interaction and geometric profile of the ligand's binding pose. To prevent these biases and to design a proper enrichment test set, the physiochemical properties of the decoys should fall within the ranges obtained for the known actives. Therefore, appropriate selection criteria were generated for the selection of the decoy set to eliminate this bias as much as possible.

The size of the original ZINC data set to be used for docking was ~1.8 million compounds. The compound collection is made up of available chemical databases from more than 20 vendors and contains a wide variety of scaffolds covering the available chemical space. The primary criterion for the use of a receptor-based pharmacophore prefilter (Figure 3) was to devise a rational method for curtailing the

ligand data set to a reasonable size for docking. Therefore, in addition to attaining a suitable enrichment, the pharmacophore model needed to be selective and able to reduce the size of the set of filtered hits in comparison to the input database. Since this pharmacophore was the first step of the filtering process, we wished to carry forward different structural chemo types to the docking phase of the screening protocol. The more liberal 1BXL structure was, therefore, used to generate this pharmacophore. The pharmacophore consisted of four hydrophobic features defined around the side chains of the Bak peptide residues, which were found to be important for the binding interaction. These side chains were Val74, Leu78, Ile81, and Ile85. The Asp83 residue is also found to contribute to the binding process, but within the 1BXL complex, the residue is located at a height from the floor of the groove. Incorporation of this feature led to a very restrictive pharmacophore which failed to retrieve a majority of the active chemo types. While the four point pharmacophore was able to retrieve most of the chemotypes, it lacked a high degree of selectivity. Exclusion spheres were, therefore, added based on the residues of the binding site to make the pharmacophore more selective. The pharmacophore which was finally selected had an enrichment factor (EF) of 6.78 and a selectivity of 9.8. Thus the pharmacophore was able to reduce the size of the initial data set by a factor of ~1/10, while retaining a majority of the actives.

Validation studies were also conducted for the docking protocols to be used in the screening. In case of 1BXL, the bound ligand is the 16-residue Bak peptide. Since flexible docking of a molecule with so many rotatable bonds is not possible, score-based active/inactive separation and docking score vs activity correlation studies were conducted using seven congeneric series of Bcl-xL actives (Supporting Information) of desired size and range of biological activity. Score-based active/inactive separation and docking score vs activity correlation were calculated within each series. Additionally, visual inspection of docked poses was also carried out to make sure that the poses were consistent with

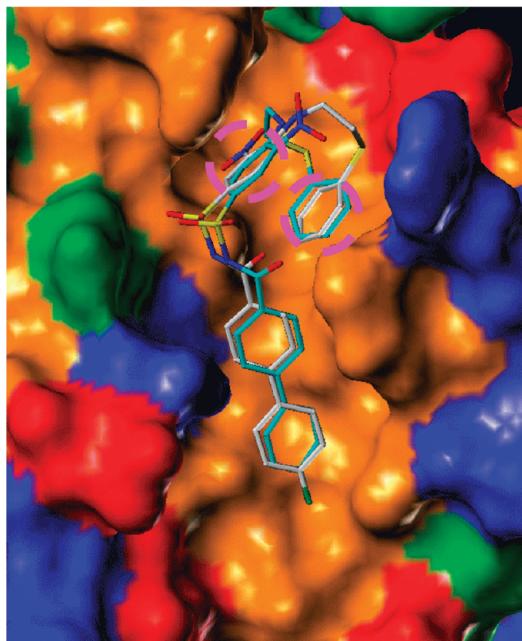


Figure 4. 1YSI pose validation study showing the docked (cyan) and the actual pose (white); the magenta circles show the location of the hydrophobic constraints.

the pharmacophoric requirements of binding site. This study was utilized to determine the best combination of binding site definition, radius for cavity detection and the scoring function utilized for the calculation of the docking poses. It was also found that because of the large size of the 1BXL cavity and the absence of any directional interactions the smaller actives produced a wide ensemble of docking poses within a congeneric series. Hydrophobic constraints were, therefore, utilized to guide the binding pose calculations and attained the desired occupancy pattern for the docked ligands. The final docking protocol selected through this study had a binding site definition of 9 Å (cavity detection radius) around an atom of the centrally placed residue Phe97 and the C scoring function for pose selection. Four hydrophobic constraints defined around the side chains of Val74, Leu78, Ile81, and Ile85 were also incorporated.

In case of 1YSI, a pose validation study was conducted using the bound pose of the N3B ligand. Variation of a set of docking parameters were evaluated, and the best results were obtained with a binding site definition of 9 Å (cavity detection radius) around an atom of the centrally placed residue Phe97 and the C scoring function for pose selection. Two hydrophobic constraints were defined around the two phenyl rings shown in Figure 4. The best-ranked docked pose obtained through this protocol had an overall heavy atom rmsd <3 Å with the actual pose. The biphenyl moiety of the molecule appears to overlap satisfactorily with that of the experimentally determined pose. The acylsulfonamide linker and the nitro-substituted phenyl ring also seem to overlap nicely with the corresponding moieties of the experimentally determined pose. The nitro group substitution at the *meta* position is pointed upward into the solvent in contrast to the experimentally determined pose, wherein it points into the cavity. This difference is expected since a charged nitro group would likely be predicted to be solvent exposed and to be away from the hydrophobic floor. Furthermore, the nitro group in the experimentally determined pose is not seen to

undergo any ligand–protein interaction and, hence, may not be crucial for the binding process. The amine linker has a slightly altered vector, but the terminal phenyl ring is again seen to overlap satisfactorily from that of the experimentally determined pose. If the position of the nitro group is omitted, then the overall heavy atom rmsd with the experimentally determined pose is <2 Å, which suggests a satisfactory pose replication.

Docking enrichment studies were conducted in triplicate for both the structures 1BXL and 1YSI. The previously described enrichment test set was utilized for conducting these studies. In case of 1BXL, the docking protocol selected from the score-based active/inactive separation and the docking score vs activity correlation studies were utilized for these studies. Validation studies were conducted for the 7–8 times speed-up mode (7S), 3 times speed-up mode (3S), and the Gold standard mode (GS) settings. The top-ranked pose for every ligand based on the C scoring function was further rescored using five different scoring functions, F, G, PMF, D, and C available from the Csore module of Sybyl 6.9. The enrichment factors were calculated at the 10% of the database screened mark for all the scoring functions. The enrichments (Figure 5) obtained over the triplicate runs were averaged to generate the final enrichment values for each scoring function and were then used for comparison. Among the scoring functions evaluated, D followed by PMF and C were found to be the better performers at the 7S, 3S, and the GS mode settings of docking accuracy. In an effort to further improve the enrichments, it was decided to combine two of the best performing functions in a range scaled format. The underlying premise being, a compound would have a higher probability of being an active if its docked pose is ranked highly by more than one scoring functions. Furthermore, a given scoring function might not be sensitive to specific structural classes of molecules which might be absent from the data set that has been used to train the scoring function. Therefore, use of multiple functions might eliminate these biases to a degree. In this case, two different combinations C/D and PMF/D were evaluated for their performance. It was found that the combination of the PMF and the D function in the range scaled format gave the best performance.

$$(PMF + D)_{RS} = PMF_{RS} + D_{RS} \quad (7)$$

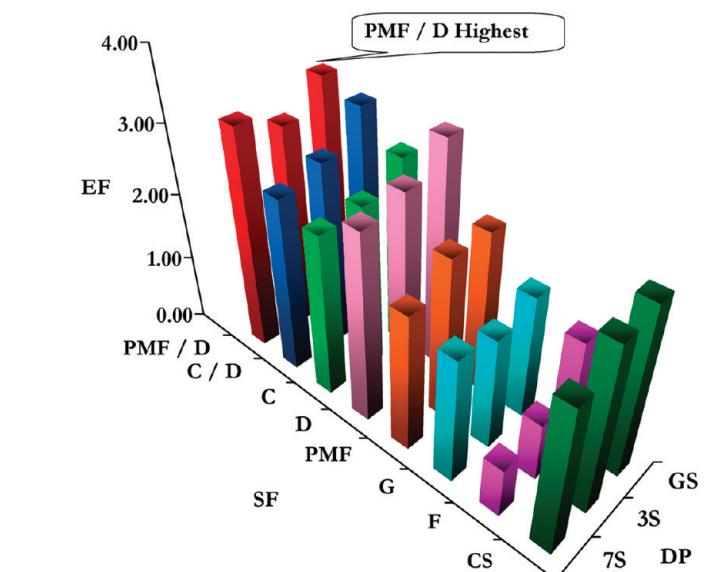
where,

$(PMF + D)_{RS}$ = Composite score obtained by addition of the PMF and D for a given pose after range scaling.

PMF_{RS} = PMF score for a given ligand pose after range scaling.

D_{RS} = D score for a given ligand pose after range scaling.

Validation studies were also performed for the final stage of the screening involving cross-docking into 1BXL and 1YSI. In case of 1BXL, the docking protocols applied in the earlier validation studies were utilized here. The docking protocol from the pose validation study was utilized in case of 1YSI. The enrichment test set being used for the validation studies consist of 54 actives out of which 13 were in the nanomolar range (<1 μM), while the rest fell in the activity range of 1 to 100 μM. The early stage of a cascade screening involves an elimination process of molecules which have very little probability of binding. On the other hand, the final stage involves a selection process for deserving poses



Docking protocol	CS	F	G	PMF	D	C	C/D	PMF/D
7S	2.16	0.74	1.85	2.04	2.78	2.41	2.59	3.27
3S	2.47	0.86	1.67	2.41	2.94	2.47	2.78	3.02
GS	2.53	1.60	1.85	2.41	3.33	2.84	3.27	3.46

Figure 5. A 3D graph showing the enrichment factors (EF) obtained for the various scoring functions or combinations (SF) with three different docking protocols (DP) of 7–8 times speed up (7S) and three times speed-up (3S) and the gold standard mode (GS) speed-up settings. The table shows the EF values (average of three runs) used for plotting the graph.

generated on the basis of a more exhaustive docking protocol. Since the more potent inhibitors with nanomolar potency fulfill the pharmacophoric requirements of ligand binding to a greater extent as compared to the less potent compounds with micromolar potency, the location of their poses within a ranked database of compounds has significant importance. Therefore, in addition to the evaluation of enrichments obtained for the whole database, enrichments were also calculated for the 13 nanomolar compounds.

A flowchart describing the protocol for the cross-docking stage of the validation study is shown in Figure 6, and the results are shown in Figures 6 and 7. As a starting point, enrichments were compared at the 25% of database screened mark for the various scoring functions. It was found that at this level the overall enrichment (Figure 7) was highest for D, followed closely by the PMF function, for both 1BXL and 1YSI. The trend between the scoring functions was reversed when comparing the nanomolars, wherein the PMF function retrieved the maximum number of nanomolar compounds followed by D. Further validation studies were carried out using the PMF function since it gave better enrichment for the nanomolar compounds. The next step of the protocol was to find out the number of molecules which were common within the top 25% of the ranked databases from 1BXL and 1YSI. The true actives would have a better probability of attaining a good binding pose within both the structures, as compared to the false actives. The true positives would be placed in the top 25% of the ranked database in both cases, while the false positives might be absent from the top 25% of either of the ranked databases and would be eliminated. It was found that the data set of common hits was ~10% of the size of the enrichment test set. This ~2.5 times reduction in size led to further enhancement in the enrichments obtained for all actives as well as the nanomolars.

Another additional step of filtration was incorporated using the score obtained through DPBDs. A scoring function for use in virtual screening has the capabilities of distinguishing between a bad vs a reasonably good pose. However, being an approximation, there is a limitation to its accuracy and resolving power, particularly when the geometric and interaction profiles of the poses are marginally different from each other. Various modeling methods, such as force-field-based ligand–protein interaction energy, energy difference, MM-GBSA, MD simulations, etc., have been used to resolve these finer differences. However, the ability of these methods have been found to be more accurate when predicting within a congeneric series of ligands as opposed to virtual screening hits belonging to diverse scaffolds. As an alternative, we decided to use a knowledge-based approach through the use of DPBDs. In this approach, we identified some of the key pharmacophoric features essential for ligand binding to Bcl-xL and delineated them in the form of descriptors (see Methods Section). The same number and type of descriptors were defined for both 1BXL and 1YSI (Figure 8). The descriptor set consisted of four hydrophobic occupancy descriptors, one hydrogen-bond acceptor descriptor, and a buried hydrophobic surface area descriptor. The hydrophobic occupancy descriptors were defined in 1BXL (Figure 8a) based on the location of the four BAK side chains Val74, Leu78, Ile81, and Ile85, which have been shown to be important for binding interaction. A hydrogen-bond acceptor descriptor was defined based on the polar interaction with Arg139. The corresponding descriptors for 1YSI (Figure 8b) were defined using structural groups of the N3B ligand. An additional buried hydrophobic surface area descriptor was also included, since the binding groove is primarily hydrophobic and requires that no hydrophobic/hydrophilic mismatches occur in this region. This descriptor set provided

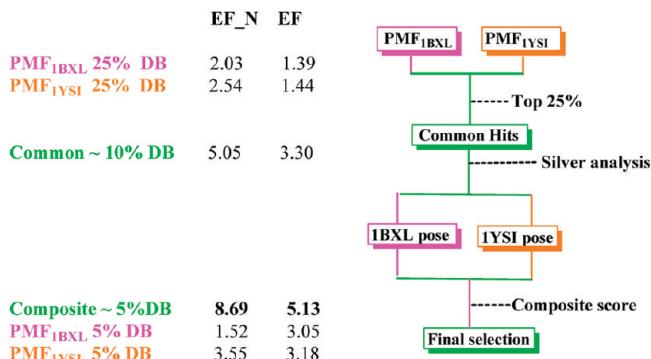
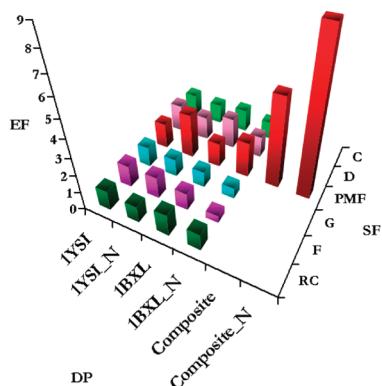


Figure 6. A flowchart utilized for the validation of the cross-docking step of the screening workflow. The improvements obtained in terms of EFs at the various stages of the screening are shown on the left. EF is for the full set of actives, and EF_N is the EF for nanomolar actives only.



	RC	F	G	PMF	D	C
1YSI	1.11	1.26	1.14	1.46	1.58	1.28
1YSI_N	0.92	1.23	1.03	2.56	1.33	1.13
1BXL	1.23	0.99	0.91	1.41	1.78	1.28
1BXL_N	0.92	0.41	0.62	2.05	1.23	1.03
Common				3.30		
Common_N				5.05		
Composite				5.14		
Composite_N				8.70		

Figure 7. A 3D graph showing the EF obtained for the various SFs with different DPs in the cross-docking phase of the validation study. The table shows the EF values (mean of triplicate runs) used for plotting the graph.

us with a greater control over the ligand selection process in a number of ways. First, the descriptor-based score could be used as an additional metric in the ranking of the database and, therefore, could supplement the role of the scoring functions. Second, the metric is able to segregate poses based on the fulfillment of the pharmacophoric requirements specific to the target in question. Third, the descriptors help in performing a “virtual” visual inspection of the database and, hence, allow one to dig deeper into the deck and to look for desired docked poses than what would be possible with manual observation.

The hits common to both the ranked databases at the 25% screened mark were further reranked using a composite score.

$$\text{Composite score} = \text{PMF}_{1\text{YSI}(\text{RS})} + \text{SILVER}_{1\text{YSI}(\text{RS})} + \text{PMF}_{1\text{BXL}(\text{RS})} + \text{SILVER}_{1\text{BXL}(\text{RS})} \quad (6)$$

where,

$\text{PMF}_{1\text{YSI}}$ (RS) = PMF score for ligand pose in 1YSI after range scaling.

$\text{SILVER}_{1\text{YSI}}(\text{RS})$ = DPBD for ligand pose in 1YSI after range scaling.

$\text{PMF}_{1\text{BXL}}(\text{RS})$ = PMF score for ligand pose in 1BXL after range scaling.

$\text{SILVER}_{1\text{BXL}}(\text{RS})$ = DPBD for ligand pose in 1BXL after range scaling.

Using this composite scoring, a higher enrichment could be obtained for all the actives ($\text{EF} = 5.13$) as well as the nanomolars ($\text{EF} = 8.98$) at $\sim 5\%$ of the screened database. At the same 5% mark, the enrichments obtained for all the actives as well as the nanomolars by using the PMF score alone were comparatively less (Figure 6). Furthermore, the size of the database was reduced by a factor of $\sim 1/20$ and would prove particularly beneficial in the actual screening workflow by reducing the number of molecules for visual inspection to a reasonable number.

Screening Workflow. The actual workflow for the virtual screening is depicted in Figure 9 and used parameters obtained from the validation studies. The screening funnel was initiated by passing the ~ 1.8 million molecules (ZINC ‘big-n-greasy’ subset) through the receptor-based pharmacophore prefilter leaving ~ 0.4 million compounds (Stage I), which corresponded to a ~ 4.5 time’s reduction in data set size. Cascade docking was initiated by docking Stage I into the 1BXL structure using the Gold 7S mode, which performs approximately 10 000 genetic algorithm (GA) operations per molecule (as opposed to 100 000 GA operations at the more rigorous GS mode). The top 10% of the ranked database based on the $(\text{PMF} + \text{D})_{\text{RS}}$ score, corresponding to ~ 73 000 molecules (~ 5.5 times data set reduction from the previous step), was selected for Stage II. The next step involved docking the Stage II database into the 1BXL structure in the Gold 3S mode (performs approximately a third of the GA operations compared to the more rigorous GS mode) and selecting the top 10% (~ 12 000 molecules) for Stage III based on the $(\text{PMF} + \text{D})_{\text{RS}}$ score, achieving a ~ 6 fold reduction in data set size from the previous step.

The molecules selected in the Stage III were then submitted to the cross-docking phase of filtration by docking into the 1BXL and 1YSI structures at the Gold GS mode setting. The top 25% of the ranked databases (Stage IV), based on the PMF function, were selected based on the score-based cutoffs obtained from the enrichment study and corresponded to ~ 7400 and ~ 6300 molecules for the 1BXL and 1YSI structures, respectively. An intersection of the two Stage IV databases resulted in the selection of ~ 4400 molecules (Stage V), which were common to both Stage IV databases. The 1BXL and 1YSI poses for these common molecules were evaluated using the DPBDs calculated using the Silver program, and the molecules were then reranked by the DPBD based composite score described in the validation studies. Following this, the top 5% of the ranked database was selected based on a score-based cutoff and resulted in 277 molecules (Stage VI).

At this final stage, there were a number of ways of doing the final selection for biological evaluation of compounds. The primary choice would be to conduct a selection of compounds which attain a desirable pose in both structures, thereby enhancing their probability of being an active. In addition, one could also conceive of two secondary selections from the top 5% of the ranked databases of the 1BXL and

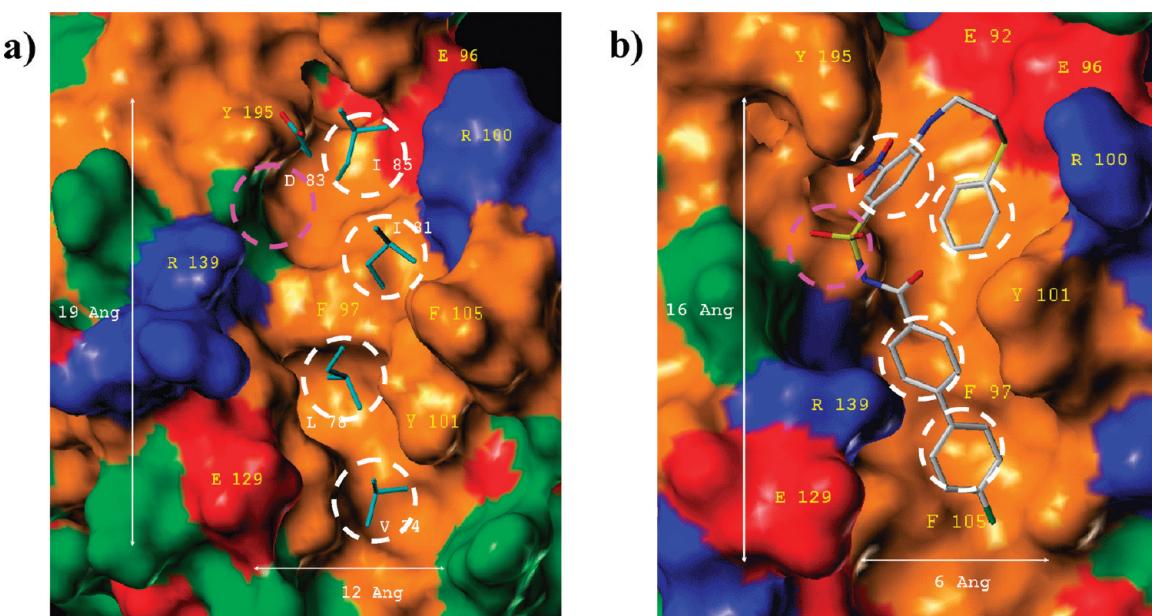


Figure 8. The figure shows the definitions of the docking pose based descriptors generated for (a) 1BXL and (b) 1YSI using Silver. The hydrophobic descriptors are shown in white dotted circles, while the hydrogen-bond acceptor descriptor is shown using magenta dotted circles.

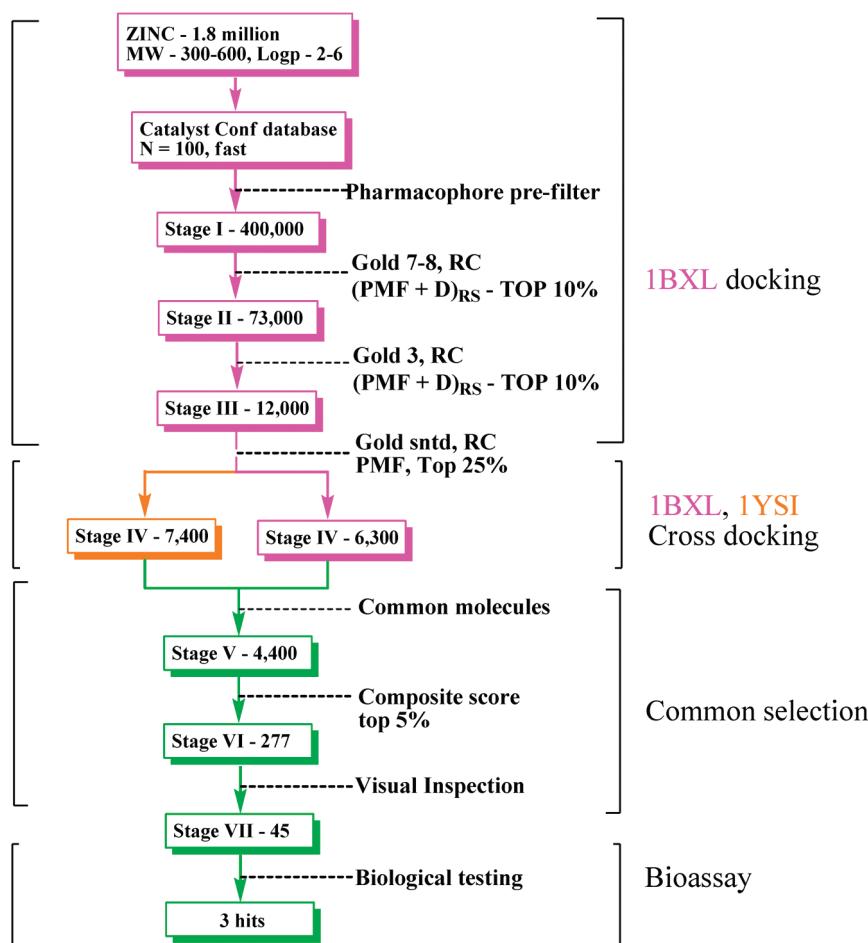


Figure 9. A workflow describing the protocol utilized for the virtual screening study.

1YSI alone, which were not common between the two ranked databases but satisfied the individual DPBDs. In this fashion, we would also be able to evaluate molecules which have a high probability of binding to either one of the protein structures. In our current study, we decided to make a

selection of molecules from the molecules common to the two structures.

Visual inspection is the final and one of the most crucial stages of a virtual screening protocol. In our case, the molecules selected for the visual inspection stage had already

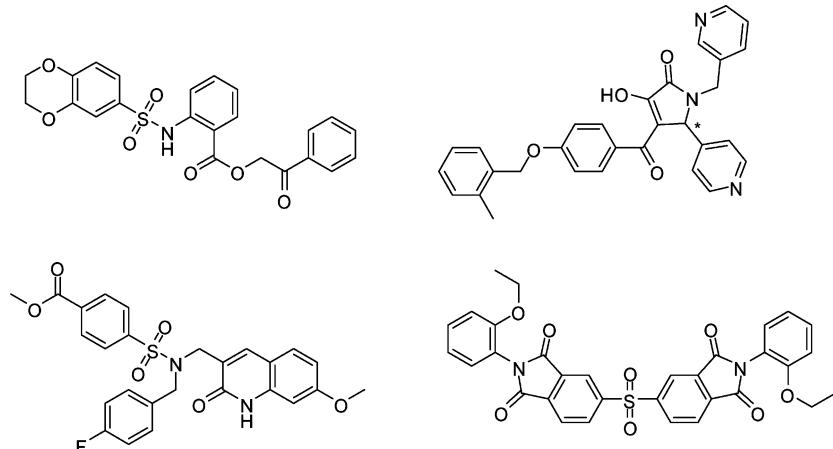


Figure 10. Representative molecules from the various structural classes selected for the first phase of virtual screening.

been filtered through the docking pose based descriptors and, hence, fulfilled the basic pharmacophoric requirements for binding. However, the poses still needed to be checked for: (i) complementarity of the interaction features of the ligand; (ii) improper atom- and bond-type definitions of the interaction features which might have affected the calculation of the pose-based descriptors; (iii) improper geometries of ligand groups; (iv) occupancy and shape complementarity with the binding pocket; and (v) hydrophobic, lipophilic mismatches (solvent exposed hydrophobic groups, polar groups buried within hydrophobic regions) as well as (vi) similarity of binding poses of a given ligand for the two protein structures. On the basis of our evaluation, around ~150 molecules were short listed. The docked poses of these molecules were subjected to an eMBRACE calculation, wherein the energy difference was calculated to evaluate the energetic stability of the bound poses.

Biological Evaluation and Analysis of Hits. Finally, the selected molecules (Figure 10) were put through a clustering analysis, and a set of 45 molecules were selected for purchase. One of the compounds **J042** (Figure 11a) was found to have an IC₅₀ of 2.58 μM ($K_{i(\text{calc})}$) 0.38 μM, see Supporting Information, Figure 4).

The predicted docking pose of **J042** in the 1BXL and 1YSI is shown in Figure 11d and e. Comparisons of the 1BXL pose of **J042** with the Bak/Bcl-xL complex (Figure 11b) suggest that **J042** undergoes interactions which are similar to the Bak residues Val 74, Leu 78, Ile 81, Asp 83, and Ile 85, which were found to be important for binding affinity. The terminal phenyl group attached to the two position of the furan ring is located approximately in the lower hydrophobic pocket and appeared to substitute for the Val 74 residue in the Bak peptide. Similarly, the furan ring can substitute for the interaction undergone by the Leu 78 residue. The phenyl ring at the 3 position of the thiazolidine-4-one ring appears to occupy a portion of the upper hydrophobic pocket where the Ile 81 residue of the Bak peptide is observed to bind. The ethyl phenyl group at the extremity of the molecule protrudes into the upper hydrophobic pocket and can substitute for the interaction undergone by the Ile 85 residue. The cyano group points upward into the solvent and can interact with the Tyr 195 and Arg 139 residues. The Asp 83 residue of the Bak peptide undergoes an interaction of the same nature.

The binding pose of **J042** in the 1YSI structure also suggests an interaction pattern similar to the acylsulfonamide inhibitor **N3B** in complex with Bcl-xL (Figure 11c). The 2-phenyl furan group of **J042** is oriented similar to the biphenyl system of **N3B**. The phenyl ring substitution on the thiazolidine-4-one ring system occupies a portion of the upper hydrophobic pocket occupied by the thiophenyl group. The ethyl phenyl substitution extends into the upper hydrophobic pocket and compensates for the interaction undergone by the nitro phenyl group of **N3B**. The sulfone group of the **N3B** ligand is replaced by the cyano group of **J042**. A polar group at this position can interact with the hydroxyl group of Tyr195 and the side chain of Arg139. The nitro group on the lower phenyl ring is solvent exposed and interacts with the Arg139 residue. The predicted binding poses of **J042** give us some idea about the possible binding mode of this inhibitor. However, we should keep in mind that this is a highly malleable binding site with profound induced-fit effects. Therefore, it is highly possible that the actual binding site configuration might be different from the two shown here and unique to the chemical class of the hit **J042**.

The two weaker hits **B29** (IC₅₀ = 620.50 μM, $K_{i(\text{calc})}$ 90.41 μM) and **B30** (IC₅₀ = 491.00 μM, $K_{i(\text{calc})}$ 71.54 μM) have binding poses very similar to each other in both the structures 1BXL and 1YSI. The predicted binding pose of **B30** in the two structures is shown in Figure 11f and g. The naphthalene moiety partially overlaps with the biphenyl group of **N3B**, while the two terminal phenyl groups occupy the upper region of the hydrophobic pocket which is occupied by the nitro phenyl and the thiophenyl groups of the **N3B**. The oxygen atoms of the sulfonamide moiety are located at a position to undergo polar interactions with the hydroxyl group of Tyr195 as well as the charged Arg139 residue. The ethoxy group of **B30** extends further down the binding pocket and is likely to enhance the binding affinity in comparison to **B29**, whose predicted pose is very similar but does not possess this hydrophobic tail group.

An overall analysis of the binding poses suggests that the upper groove acts as a necessary hydrophobic anchor point for ligand binding and is found to be important for the ligand poses of all the three hits. While all the hits project two hydrophobic groups toward the hotspots occupied by the Ile81 and Ile85 of the Bak peptide, it might be most advantageous for **J042** where the phenyl group connected

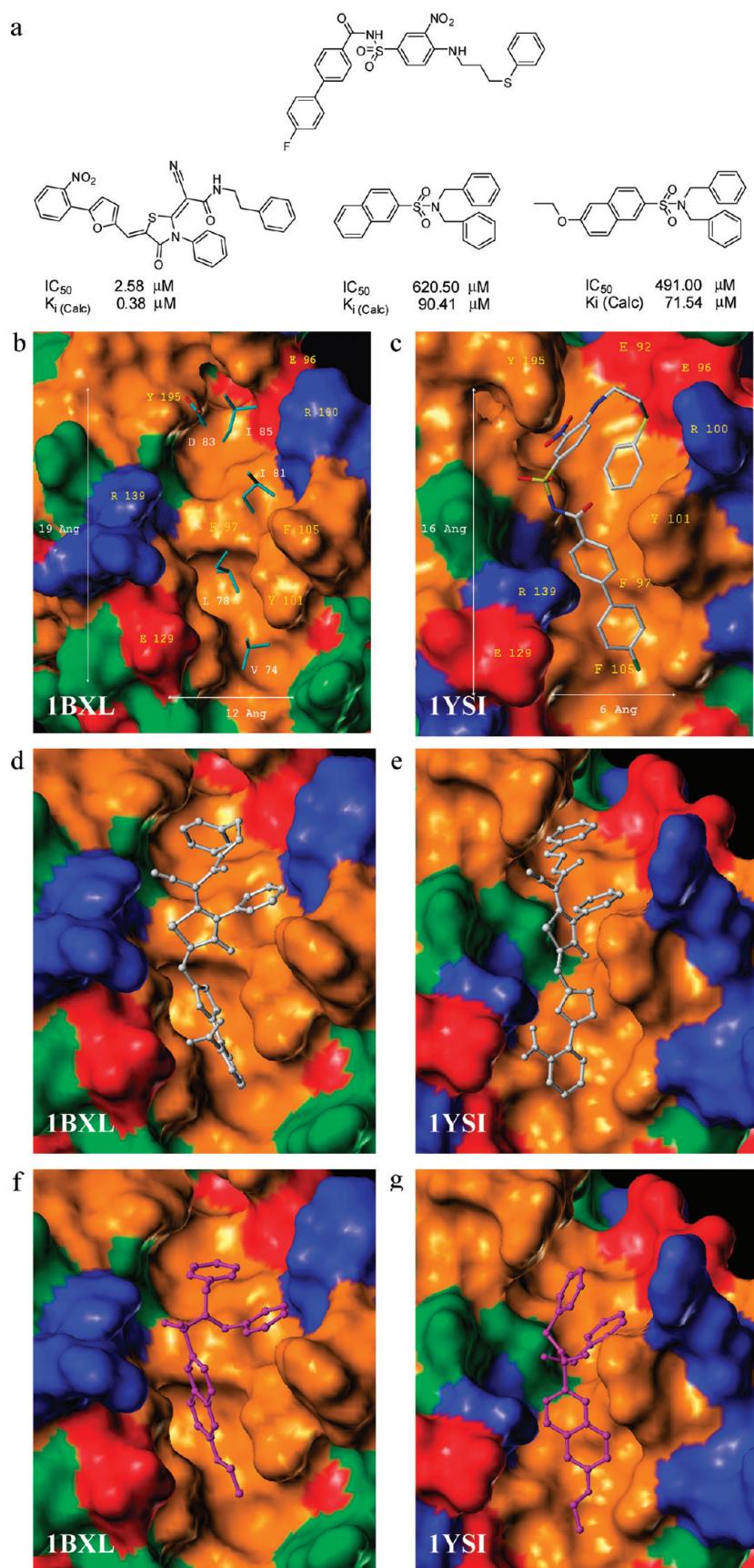


Figure 11. (a) Chemical structure and Bcl-xL inhibitory activity of the hits. The regions of the molecule which are observed/predicted to undergo a similar interaction profile with the 1YSI ligand are color coded. (b) The binding groove of the Bak/Bcl-xL structure (1BXL) showing the key amino acid side chains (cyan) of the Bak peptide. (c) The binding groove of the N3B/Bcl-xL structure showing the key interactions of the N3B ligand which surrogates for the Bak peptide interactions. The predicted binding pose of **J042** in (d) 1BXL and (e) 1YSI. The binding pose of **B30** in (f) 1BXL and (g) 1YSI. In all of these figures, the hydrophobic residues are colored in orange, while the basic and acidic residues are colored in blue and red, respectively.

Table 1. eMBRACE Energy Terms Calculated for the Binding Poses of the Three Hits in the 1YSI Structure

srno	total energy	valence	van der Waals	electrostatic	solvation	VS rank ^a
J042	-139.93	18.47	-196.33	-112.56	150.49	1
B29	-51.43	3.11	-106.76	59.67	-7.46	3
B30	-103.9	21.5	-161.15	-69.8	105.56	2

^a Relative rank of the compounds among the 45 compounds selected for biological evaluation.

by the ethyl linker provides enhanced hydrophobic interaction and better shape complementarity with the upper region of the binding groove. Also of interest might be the cyano group which reaches up further into the solvent than the sulfone oxygens, thereby attaining two things: (i) a lower hydrophilic/lipophilic mismatch with the primarily hydrophobic groove and (ii) perhaps providing a better interaction with the Arg139 residue which interacts with the Asp 83 of the Bak peptide. Alternatively, as seen in the 1YSI structure, the cyano group can form a hydrogen bond with the Tyr101 residue, while the nitro group can interact with the Arg139 side chain. While the interaction profile in the upper hydrophobic groove is essential for ligand binding, it is the interaction profile in the lower hydrophobic groove that clearly demarcates the three hits based on their inhibitory profile. The weakest hit **B29** has the naphthalene moiety over the region occupied by Leu74 of the Bak peptide, which demarcates the upper and the lower grooves but fails to occupy the region occupied by the Val74 residue of the Bak peptide. The ethoxy moiety of **B30**, the next most potent hit, occupies this region and provides additional hydrophobic interaction. However, the ethoxy group being flexible would have a degree of entropy that might counteract the stability attained through the additional hydrophobic interaction. In case of **J042**, the furan connected to the ortho nitro phenyl group forms a relatively rigid system that goes deep into the lower hydrophobic groove providing a highly enhanced hydrophobic interaction profile that results in its higher activity. The polar nitro group on the phenyl ring points into the solvent, thereby further reducing the conformational entropy of this portion of the ligand, and provides an amphiphilic character similar to the Bak peptide bound to Bcl-xL, thereby making it the most potent hit of the three. The eMBRACE energy differences (Table 1) for the docking poses of these three hits (based on 1YSI poses) corroborates with their biological activity providing the correct rank ordering based on their van der Waals, electrostatic, and total energy differences. The Pearson correlation coefficient between the biological activities and the total energy as well as the van der Waals term is ~0.80, while that with the electrostatic term is ~0.70, suggesting that the stability of the binding poses is primarily driven by hydrophobic interactions.

CONCLUSIONS

Bcl-xL, an antiapoptotic member of the BCL-2 protein family, plays a critical role in the cells internal apoptotic pathway and is an ideal target for the development of apoptosis-based anticancer therapy. In this present study, a comprehensive virtual screening protocol was developed to identify inhibitors of Bcl-xL. A number of innovative filters were incorporated into the various stages of the virtual screening funnel. Extensive validation studies were con-

ducted for each step of the protocol to ensure its filtering efficiency. First, a receptor-based pharmacophore prefilter was developed from the more lenient 1BXL structure and incorporated at the top of the screening funnel, which led to a rational curtailing (~4.5 times reduction) of the initial database. The curtailed database, now of a suitable size for docking, was subjected to a cascade docking approach using the more lenient 1BXL structure. This involved a series of docking filters based on increasing levels of docking complexity, resulting in a curtailed set ~33 times smaller than that of the predocking data set. The next filtering step involved a unique cross-docking strategy coupled with a composite scoring scheme based on DPBDs which showed enhancement in enrichment factors during validation studies. This unique modification was able to incorporate the effects of the conformational flexibility of the binding site and to filter hits with a higher probability of fitting into this highly plastic binding site. This step helped in a further reduction of the data set and allowed us to conduct a visual inspection to carry out the final selection of compounds for screening. The final selection was focused on compounds which showed desirable binding poses in both the 1BXL and the 1YSI structures. The overall protocol was utilized to screen the ~1.8 million ZINC ‘big-n-greasy’ subset and to select a set of 45 compounds for biological evaluation in a Bcl-xL inhibition assay. Thus, an ~40 000-fold reduction in size of the database was carried out using a set of rational filters incorporated within the screening funnel. One of the compounds **J042** showed inhibitory activity of IC_{50} 2.48 μM (K_i (calc) 0.38 μM). Two other weakly active compounds **B29** (IC_{50} = 620.50 μM , K_i (calc) 90.41 μM) and **B30** (IC_{50} = 491.00 μM , K_i (calc) 71.54 μM) were also identified through the screening process. While the use of a ‘big-n-greasy’ subset to identify preliminary hits would be thought of as a disadvantage from traditional medchem optimization perspectives, the nature of the protein–protein interaction interfaces makes it imperative that these cutoffs be recalibrated for this particular group of targets. The ligand efficiency (LE)⁸⁹ of the identified hits based on their calculated K_i s were 0.21, 0.19, and 0.18, respectively. Along similar lines, the LE’s (based on K_i) for some the compounds currently in preclinical and clinical studies are >0.21, >0.18 (K_i s < 1 nM) and 0.22 for ABT-737, ABT-263, and Gossypol, respectively. Our current hits are not yet optimized for potency, and hence, if we are able to carry out hit optimization and SAR exploration while preserving LE in the same range, then we would be able to have leads in the same property space as the current established Bcl-xL inhibitors. A decomposition of the substructures of the identified within their predicted docking poses suggested similarities in their interaction profile with the Bak peptide as well as the corresponding moieties from the **N3B** ligand

from the 1YSI structures. The predicted binding poses of the identified hits were also indicative of the importance of the interaction of the ligand groups with both the upper and the lower hydrophobic grooves of the Bcl-xL pocket. The future directions from this point would be to conduct a similarity/substructure search using the scaffolds of the hits to retrieve active structural analogues with the highest LEs. Additional retro synthetic analysis could be carried out to evaluate structural modifications not accessible through the commercially available analogues. Evaluation of their inhibitory activity would help us in understanding the SAR, and the gain in potency would likely be associated with the gain of cellular activity which, in other cases, have been shown to trail behind the target activity.

Supporting Information Available: Compounds used in the enrichment study, plots, and dose response curves from the biological evaluation and NMR spectra of active compounds. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- Fulda, S.; Debatin, K. M. Modulation of apoptosis signaling for cancer therapy. *Arch. Immunol. Ther. Exp.* **2006**, *54* (3), 173–175.
- Thompson, C. B. Apoptosis in the pathogenesis and treatment of disease. *Science* **1995**, *267* (5203), 1456–1462.
- Hanahan, D.; Weinberg, R. A. The hallmarks of cancer. *Cell* **2000**, *100* (1), 57–70.
- Daniel, N. N.; Korsmeyer, S. J. Cell death: critical control points. *Cell* **2004**, *116* (2), 205–219.
- Nicholson, D. W. From bench to clinic with apoptosis-based therapeutic agents. *Nature* **2000**, *407* (6805), 810–816.
- Kakehi, Y.; Zeng, Y. Receptors of TRAIL as the target of apoptosis-inducing therapy for genitourinary cancers. *New Perspect. Cancer Res. Ther.* **2005**, 377–391.
- McConkey, D. J. Resistance to apoptosis in cancer therapy. *Death Recept. Cancer Ther.* **2005**, 43–64.
- Bremer, E.; van Dam, G.; Kroesen, B. J.; de Leij, L.; Helfrich, W. Targeted induction of apoptosis for cancer therapy: current progress and prospects. *Trends Mol. Med.* **2006**, *12* (8), 382–393.
- Meng, X. W.; Lee, S. H.; Kaufmann, S. H. Apoptosis in the treatment of cancer: a promise kept. *Curr. Opin. Cell Biol.* **2006**, *18* (6), 668–676.
- Walters, W. P.; Namchuk, M. Designing screens: how to make your hits a hit. *Nat. Rev. Drug. Discovery* **2003**, *2* (4), 259–266.
- Secchiero, P.; Vaccarezza, M.; Gonelli, A.; Zauli, G. TNF-related apoptosis-inducing ligand (TRAIL): a potential candidate for combined treatment of hematological malignancies. *Curr. Pharm. Des.* **2004**, *10* (29), 3673–3681.
- Merino, D.; Lalaoui, N.; Morizot, A.; Solary, E.; Micheau, O. TRAIL in cancer therapy: present and future challenges. *Expert Opin. Ther. Targets* **2007**, *11* (10), 1299–1314.
- Johnstone, R. W.; Frew, A. J.; Smyth, M. J. The TRAIL apoptotic pathway in cancer onset, progression and therapy. *Nat. Rev. Cancer* **2008**, *8* (10), 782–798.
- Kirkin, V.; Joos, S.; Zornig, M. The role of Bcl-2 family members in tumorigenesis. *Biochim. Biophys. Acta* **2004**, *1644* (2–3), 229–249.
- Baell, J. B.; Huang, D. C. Prospects for targeting the Bcl-2 family of proteins to develop novel cytotoxic drugs. *Biochem. Pharmacol.* **2002**, *64* (5–6), 851–863.
- Deveraux, Q. L.; Takahashi, R.; Salvesen, G. S.; Reed, J. C. X-linked IAP is a direct inhibitor of cell-death proteases. *Nature* **1997**, *388* (6639), 300–304.
- Du, C.; Fang, M.; Li, Y.; Li, L.; Wang, X. Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell* **2000**, *102* (1), 33–42.
- Verhagen, A. M.; Ekerdt, P. G.; Pakusch, M.; Silke, J.; Connolly, L. M.; Reid, G. E.; Moritz, R. L.; Simpson, R. J.; Vaux, D. L.; Identification of, DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. *Cell* **2000**, *102* (1), 43–53.
- Li, L.; Thomas, R. M.; Suzuki, H.; De Brabander, J. K.; Wang, X.; Harran, P. G. A small molecule Smac mimetic potentiates TRAIL- and TNFalpha-mediated cell death. *Science* **2004**, *305* (5689), 1471–1474.
- Cossu, F.; Milani, M.; Mastrangelo, E.; Vachette, P.; Servida, F.; Lecis, D.; Canevari, G.; Delia, D.; Drago, C.; Rizzo, V.; Manzoni, L.; Seneci, P.; Scolastico, C.; Bolognesi, M. Structural basis for bivalent smac-mimetics recognition in the IAP protein family. *J. Mol. Biol.* **2009**, *392* (3), 630–644.
- Chen, D. J.; Huerta, S. Smac mimetics as new cancer therapeutics. *Anti-Cancer Drugs* **2009**, *20* (8), 646–658.
- Chene, P. Inhibiting the p53-MDM2 interaction: an important target for cancer therapy. *Nat. Rev. Cancer* **2003**, *3* (2), 102–109.
- Vousden, K. H. Outcomes of p53 activation—spoilt for choice. *J. Cell Sci.* **2006**, *119* (Pt 24), 5015–5020.
- Vazquez, A.; Bond, E. E.; Levine, A. J.; Bond, G. L. The genetics of the p53 pathway, apoptosis and cancer therapy. *Nat. Rev. Drug. Discovery* **2008**, *7* (12), 979–987.
- Rayburn, E. R.; Ezell, S. J.; Zhang, R. Recent advances in validating MDM2 as a cancer target. *Anti-Cancer Agents Med. Chem.* **2009**, *9* (8), 882–903.
- Yu, S.; Qin, D.; Shangary, S.; Chen, J.; Wang, G.; Ding, K.; McEachern, D.; Qiu, S.; Nikolovska-Coleska, Z.; Miller, R.; Kang, S.; Yang, D.; Wang, S. Potent and Orally Active Small-Molecule Inhibitors of the MDM2-p53 Interaction. *J. Med. Chem.* **2009**.
- Oksanen, E.; Ahonen, A. K.; Tuominen, H.; Tuominen, V.; Lahti, R.; Goldman, A.; Heikinheimo, P. A complete structural description of the catalytic cycle of yeast pyrophosphatase. *Biochemistry* **2007**, *46* (5), 1228–1239.
- Yin, H.; Hamilton, A. D. Strategies for targeting protein-protein interactions with synthetic agents. *Angew. Chem., Int. Ed. Engl.* **2005**, *44* (27), 4130–4163.
- Toogood, P. L. Inhibition of protein-protein association by small molecules: approaches and progress. *J. Med. Chem.* **2002**, *45* (8), 1543–1558.
- Chene, P. Drugs targeting protein-protein interactions. *ChemMedChem* **2006**, *1* (4), 400–411.
- Fletcher, S.; Hamilton, A. D. Protein-protein interaction inhibitors: small molecules from screening techniques. *Curr. Top. Med. Chem.* **2007**, *7* (10), 922–927.
- DeLano, W. L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.* **2002**, *12* (1), 14–20.
- Chakrabarti, P.; Janin, J. Dissecting protein-protein recognition sites. *Proteins* **2002**, *47* (3), 334–343.
- Petros, A. M.; Olejniczak, E. T.; Fesik, S. W. Structural biology of the Bcl-2 family of proteins. *Biochim. Biophys. Acta* **2004**, *1644* (2–3), 83–94.
- Mikhailov, V.; Mikhailova, M.; Pulkrabek, D. J.; Dong, Z.; Venkatachalam, M. A.; Saikumar, P. Bcl-2 prevents Bax oligomerization in the mitochondrial outer membrane. *J. Biol. Chem.* **2001**, *276* (21), 18361–18374.
- Dlugosz, P. J.; Billen, L. P.; Annis, M. G.; Zhu, W.; Zhang, Z.; Lin, J.; Leber, B.; Andrews, D. W. Bcl-2 changes conformation to inhibit Bax oligomerization. *EMBO J.* **2006**, *25* (11), 2287–2296.
- Yano, M.; Terada, K.; Gotoh, T.; Mori, M. In vitro analysis of Bcl-2 proteins in mitochondria and endoplasmic reticulum: similarities in antiapoptotic functions and differences in regulation. *Exp. Cell Res.* **2007**, *313* (17), 3767–3778.
- Mikhailov, V.; Mikhailova, M.; Degenhardt, K.; Venkatachalam, M. A.; White, E.; Saikumar, P. Association of Bax and Bak homooligomers in mitochondria. Bax requirement for Bak reorganization and cytochrome c release. *J. Biol. Chem.* **2003**, *278* (7), 5367–5376.
- An, J.; Chen, Y.; Huang, Z. Critical upstream signals of cytochrome C release induced by a novel Bcl-2 inhibitor. *J. Biol. Chem.* **2004**, *279* (18), 19133–19140.
- Degenhardt, K.; Sundararajan, R.; Lindsten, T.; Thompson, C.; White, E. Bax and Bak independently promote cytochrome C release from mitochondria. *J. Biol. Chem.* **2002**, *277* (16), 14127–14134.
- Ruvkun, G. Molecular biology. Glimpses of a tiny RNA world. *Science* **2001**, *294* (5543), 797–799.
- Mohammad, R.; Giri, A.; Goustin, A. S. Small-molecule inhibitors of Bcl-2 family proteins as therapeutic agents in cancer. *Recent Pat. Anti-Cancer Drug Discovery* **2008**, *3* (1), 20–30.
- Amundson, S. A.; Myers, T. G.; Scudiero, D.; Kitada, S.; Reed, J. C.; Fornace, A. J., Jr. An informatics approach identifying markers of chemosensitivity in human cancer cell lines. *Cancer Res.* **2000**, *60* (21), 6101–6110.
- Reed, J. C. Promise and problems of Bcl-2 antisense therapy. *J. Natl. Cancer Inst.* **1997**, *89* (14), 988–990.
- Letai, A.; Sorcinelli, M. D.; Beard, C.; Korsmeyer, S. J. Antiapoptotic BCL-2 is required for maintenance of a model leukemia. *Cancer Cell* **2004**, *6* (3), 241–249.
- Klasa, R. J.; Gillum, A. M.; Klem, R. E.; Frankel, S. R. Oblimersen Bcl-2 antisense: facilitating apoptosis in anticancer treatment. *Antisense Nucleic Acid Drug Dev.* **2002**, *12* (3), 193–213.

- (47) Minko, T.; Dharap, S. S.; Fabbricatore, A. T. Enhancing the efficacy of chemotherapeutic drugs by the suppression of antiapoptotic cellular defense. *Cancer Detect. Prev.* **2003**, *27* (3), 193–202.
- (48) Chan, S. L.; Lee, M. C.; Tan, K. O.; Yang, L. K.; Lee, A. S.; Flotow, H.; Fu, N. Y.; Butler, M. S.; Soejarto, D. D.; Buss, A. D.; Yu, V. C. Identification of chelerythrine as an inhibitor of BclXL function. *J. Biol. Chem.* **2003**, *278* (23), 20453–20456.
- (49) Degterev, A.; Lugovskoy, A.; Cardone, M.; Mulley, B.; Wagner, G.; Mitchison, T.; Yuan, J. Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. *Nat. Cell Biol.* **2001**, *3* (2), 173–182.
- (50) Rega, M. F.; Leone, M.; Jung, D.; Cotton, N. J.; Stebbins, J. L.; Pellecchia, M. Structure-based discovery of a new class of Bcl-xL antagonists. *Bioorg. Chem.* **2007**, *35* (4), 344–353.
- (51) Wang, J. L.; Liu, D.; Zhang, Z. J.; Shan, S.; Han, X.; Srinivasula, S. M.; Croce, C. M.; Alnemri, E. S.; Huang, Z. Structure-based discovery of an organic compound that binds Bcl-2 protein and induces apoptosis of tumor cells. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (13), 7124–7129.
- (52) Kitada, S.; Leone, M.; Sareth, S.; Zhai, D.; Reed, J. C.; Pellecchia, M. Discovery, characterization, and structure-activity relationships studies of proapoptotic polyphenols targeting B-cell lymphocyte/leukemia-2 proteins. *J. Med. Chem.* **2003**, *46* (20), 4259–4264.
- (53) Petros, A. M.; Dinges, J.; Augeri, D. J.; Baumeister, S. A.; Betebenner, D. A.; Bures, M. G.; Elmore, S. W.; Hajduk, P. J.; Joseph, M. K.; Landis, S. K.; Nettesheim, D. G.; Rosenberg, S. H.; Shen, W.; Thomas, S.; Wang, X.; Zanze, I.; Zhang, H.; Fesik, S. W. Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. *J. Med. Chem.* **2006**, *49* (2), 656–663.
- (54) Tzung, S. P.; Kim, K. M.; Basanez, G.; Giedt, C. D.; Simon, J.; Zimmerberg, J.; Zhang, K. Y.; Hockenberry, D. M. Antimycin A mimics a cell-death-inducing Bcl-2 homology domain 3. *Nat. Cell Biol.* **2001**, *3* (2), 183–191.
- (55) Leone, M.; Zhai, D.; Sareth, S.; Kitada, S.; Reed, J. C.; Pellecchia, M. Cancer prevention by tea polyphenols is linked to their direct inhibition of antiapoptotic Bcl-2-family proteins. *Cancer Res.* **2003**, *63* (23), 8118–8121.
- (56) Real, P. J.; Cao, Y.; Wang, R.; Nikolovska-Coleska, Z.; Sanz-Ortiz, J.; Wang, S.; Fernandez-Luna, J. L. Breast cancer cells can evade apoptosis-mediated selective killing by a novel small molecule inhibitor of Bcl-2. *Cancer Res.* **2004**, *64* (21), 7947–7953.
- (57) Wang, G.; Nikolovska-Coleska, Z.; Yang, C. Y.; Wang, R.; Tang, G.; Guo, J.; Shangary, S.; Qiu, S.; Gao, W.; Yang, D.; Meagher, J.; Stuckey, J.; Krajewski, K.; Jiang, S.; Roller, P. P.; Abaan, H. O.; Tomita, Y.; Wang, S. Structure-based design of potent small-molecule inhibitors of anti-apoptotic Bcl-2 proteins. *J. Med. Chem.* **2006**, *49* (21), 6139–6142.
- (58) Wang, Y. Obatoclax Mesylate. *Drugs Future* **2007**, *32* (3), 228–233.
- (59) Lugovskoy, A. A.; Degterev, A. I.; Fahmy, A. F.; Zhou, P.; Gross, J. D.; Yuan, J.; Wagner, G. A novel approach for characterizing protein ligand complexes: molecular basis for specificity of small-molecule Bcl-2 inhibitors. *J. Am. Chem. Soc.* **2002**, *124* (7), 1234–1240.
- (60) Bruncko, M.; Oost, T. K.; Belli, B. A.; Ding, H.; Joseph, M. K.; Kunzer, A.; Martineau, D.; McClellan, W. J.; Mitten, M.; Ng, S. C.; Nimmer, P. M.; Oltersdorf, T.; Park, C. M.; Petros, A. M.; Shoemaker, A. R.; Song, X.; Wang, X.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H.; Elmore, S. W. Studies leading to potent, dual inhibitors of Bcl-2 and Bcl-xL. *J. Med. Chem.* **2007**, *50* (4), 641–662.
- (61) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **2005**, *435* (7042), 677–681.
- (62) Park, C. M.; Bruncko, M.; Adickes, J.; Bauch, J.; Ding, H.; Kunzer, A.; Marsh, K. C.; Nimmer, P.; Shoemaker, A. R.; Song, X.; Tahir, S. K.; Tse, C.; Wang, X.; Wendt, M. D.; Yang, X.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H.; Elmore, S. W. Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. *J. Med. Chem.* **2008**, *51* (21), 6902–6915.
- (63) Posner, B. A. High-throughput screening-driven lead discovery: meeting the challenges of finding new therapeutics. *Curr. Opin. Drug Discovery Dev.* **2005**, *8* (4), 487–494.
- (64) Fox, S.; Farr-Jones, S.; Sopchak, L.; Boggs, A.; Comley, J. High-throughput screening: searching for higher productivity. *J. Biomol. Screening* **2004**, *9* (4), 354–358.
- (65) Uciechowska, U.; Schemies, J.; Neugebauer, R. C.; Huda, E. M.; Schmitt, M. L.; Meier, R.; Verdin, E.; Jung, M.; Sippl, W. Thiobarbiturates as sirtuin inhibitors: virtual screening, free-energy calculations, and biological testing. *ChemMedChem* **2008**, *3* (12), 1965–1976.
- (66) Park, H.; Jung, S. K.; Jeong, D. G.; Ryu, S. E.; Kim, S. J. Discovery of VHR phosphatase inhibitors with micromolar activity based on structure-based virtual screening. *ChemMedChem* **2008**, *3* (6), 877–880.
- (67) Larsen, S. B.; Omkvist, D. H.; Brodin, B.; Nielsen, C. U.; Steffansen, B.; Olsen, L.; Jorgensen, F. S. Discovery of ligands for the human intestinal Di-/Tripeptide transporter (hPEPT1) using a QSAR-assisted virtual screening strategy. *ChemMedChem* **2009**, *4* (9), 1439–1445.
- (68) Heinke, R.; Spannhoff, A.; Meier, R.; Trojer, P.; Bauer, I.; Jung, M.; Sippl, W. Virtual screening and biological characterization of novel histone arginine methyltransferase PRMT1 inhibitors. *ChemMedChem* **2009**, *4* (1), 69–77.
- (69) Wu, S.; Bottini, M.; Rickert, R. C.; Mustelin, T.; Tautz, L. In silico screening for PTPN22 inhibitors: active hits from an inactive phosphatase conformation. *ChemMedChem* **2009**, *4* (3), 440–444.
- (70) Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S. Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal. Biochem.* **2004**, *332* (2), 261–273.
- (71) Fang, X.; Wang, R. The Ki Calculator; http://sw16.im.med.umich.edu/software/calc_ki. Accessed March 15, 2010.
- (72) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; Thompson, C. B.; Fesik, S. W. Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. *Science* **1997**, *275* (5302), 983–986.
- (73) Desai, P. V.; Patny, A.; Sabnis, Y.; Tekwani, B.; Gut, J.; Rosenthal, P.; Srivastava, A.; Avery, M. Identification of novel parasitic cysteine protease inhibitors using virtual screening. 1. The ChemBridge database. *J. Med. Chem.* **2004**, *47* (26), 6609–6615.
- (74) Desai, P. V.; Patny, A.; Gut, J.; Rosenthal, P. J.; Tekwani, B.; Srivastava, A.; Avery, M. Identification of novel parasitic cysteine protease inhibitors by use of virtual screening. 2. The available chemical directory. *J. Med. Chem.* **2006**, *49* (5), 1576–1584.
- (75) Mukherjee, P.; Desai, P.; Ross, L.; White, E. L.; Avery, M. A. Structure-based virtual screening against SARS-3CL(pro) to identify novel non-peptidic hits. *Bioorg. Med. Chem.* **2008**, *16* (7), 4138–4149.
- (76) Carlson, H. A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6* (4), 447–452.
- (77) Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H. Flexible docking of a ligand peptide to a receptor protein by multicanonical molecular dynamics simulation. *Chem. Phys. Lett.* **1997**, *278* (4,5,6), 297–301.
- (78) Wasserman, Z. R.; Hodge, C. N. Fitting an inhibitor into the active site of thermolysin: a molecular dynamics case study. *Proteins* **1996**, *24* (2), 227–237.
- (79) Di Nola, A.; Roccatano, D.; Berendsen, H. J. Molecular dynamics simulation of the docking of substrates to proteins. *Proteins* **1994**, *19* (3), 174–182.
- (80) Mangoni, M.; Roccatano, D.; Di Nola, A. Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation. *Proteins* **1999**, *35* (2), 153–162.
- (81) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49* (2), 534–553.
- (82) Carlson, H. A.; McCammon, J. A. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* **2000**, *57* (2), 213–218.
- (83) Jiang, F.; Kim, S. H. Soft docking": matching of molecular surface cubes. *J. Mol. Biol.* **1991**, *219* (1), 79–102.
- (84) Knegtel, R. M.; Kuntz, I. D.; Oshiro, C. M. Molecular docking to ensembles of protein structures. *J. Mol. Biol.* **1997**, *266* (2), 424–440.
- (85) Broughton, H. B. A method for including protein flexibility in protein-ligand docking: improving tools for database mining and virtual screening. *J. Mol. Graph. Model.* **2000**, *18* (3), 247–257, 302–244.
- (86) Osterberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* **2002**, *46* (1), 34–40.
- (87) Sherman, W.; Beard, H. S.; Farid, R. Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug. Des.* **2006**, *67* (1), 83–84.
- (88) Irwin, J. J.; Shoichet, B. K. ZINC--a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177–182.
- (89) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9* (10), 430–431.