



Published in final edited form as:

J Med Chem. 2016 May 12; 59(9): 4152–4170. doi:10.1021/acs.jmedchem.5b00150.

DARC: mapping surface topography by ray-casting for effective virtual screening at protein interaction sites

Ragul Gowthaman^{1,†}, Sven A. Miller^{2,†}, Steven Rogers³, Jittasak Khowsathit², Lan Lan², Nan Bai², David K. Johnson¹, Chunjing Liu³, Liang Xu^{2,4}, Asokan Anbanandam⁵, Jeffrey Aubé^{3,6}, Anuradha Roy⁷, and John Karanicolas^{1,2,*}

¹Center for Computational Biology, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

²Department of Molecular Biosciences, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

³Center of Biomedical Research Excellence, Center for Cancer Experimental Therapeutics, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

⁴Department of Radiation Oncology, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

⁵Biomolecular NMR Laboratory, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

⁶Department of Medicinal Chemistry, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

⁷High Throughput Screening Laboratory, University of Kansas, 2030 Becker Dr., Lawrence, KS 66045-7534

Abstract

Protein-protein interactions represent an exciting and challenging target class for therapeutic intervention using small molecules. Protein interaction sites are often devoid of the deep surface pockets presented by “traditional” drug targets, and crystal structures reveal that inhibitors typically engage these sites using very shallow binding modes. As a consequence, modern virtual screening tools developed to identify inhibitors of traditional drug targets do not perform as well when they are instead deployed at protein interaction sites. To address the need for novel inhibitors of important protein interactions, here we introduce an alternate docking strategy specifically designed for this regime. Our method, termed DARC (*Docking Approach using Ray-Casting*), matches the topography of a surface pocket “observed” from within the protein to the topography “observed” when viewing a potential ligand from the same vantage point. We applied DARC to carry out a virtual screen against the protein interaction site of human anti-apoptotic protein Mcl-1, and found that 4 of the top-scoring 21 compounds showed clear inhibition in a biochemical assay. The K_i values for these compounds ranged from 1.2 to 21 μM , and each had ligand

*To whom correspondence should be addressed. johnk@ku.edu, 785-864-8298.

†indicates authors of equal contribution

efficiency comparable to promising small-molecule inhibitors of other protein-protein interactions. These hit compounds do not resemble the natural (protein) binding partner of Mcl-1, nor do they resemble any known inhibitors of Mcl-1. Our results thus demonstrate the utility of DARC for identifying novel inhibitors of protein-protein interactions.

Introduction

Protein-protein interactions comprise the underlying mechanisms for cell proliferation, differentiation, and survival; manipulation of these interactions thus represents a promising avenue for therapeutic intervention in a variety of settings. Despite the urgent need for small-molecule modulators of protein interactions, however, as recently as five years ago the dearth of success stories meant that even the druggability of this target class was very much in question [1,2]. At the same time, the challenges associated with these important targets help spur refinement of innovative new techniques, including fragment-based methods [3], peptidomimetics [4], and non-peptidic mimetics of protein secondary structural elements [5,6]. By now, these advances have enabled discovery of inhibitors against many different protein-protein interactions, and have led to an impressive array of compounds in various stages of clinical trials and pre-clinical optimization [7,8].

Modern methods for structure-based virtual screening have proven effective for a variety of applications, providing novel hits to address a very wide range of “traditional” targets such as enzymes and G protein-coupled receptors [9,10]. Unfortunately, structure-based virtual screening tools for identifying small-molecule inhibitors of protein-protein interfaces have lagged behind: due to their relatively exposed binding modes, the same docking methods that perform well for traditional drug targets struggle to correctly identify compounds that disrupt protein interactions [11]. This problem is particularly acute because the natural partners of these targets – proteins – cannot easily be used for inspiration in designing small-molecule inhibitors or as templates for ligand-based virtual screening, as is possible for many enzymes [12,13] and G protein-coupled receptors [14].

In a number of cases, protein structures have been solved both in complex with a biological protein partner and also in complex with a small-molecule inhibitor. Comparison of the unbound protein structure to the equivalent structure in complex with a small molecule shows that binding is not associated with a large conformational change; and yet, the concave pocket on the protein surface in which the small molecule binds is often absent in the unbound structure, so that the flat surface of the unbound structure must often undergo local rearrangement to reveal the small-molecule binding site [15]. This has confounded structure-based methods for identifying inhibitors – how can the protein structure be used to identify an inhibitor, if the protein structure will change upon binding?

We recently described a method for identifying protein surface pockets suitable for small-molecule binding, from either co-crystal structures or from simulations starting from the unbound protein structure [16]. Here, we introduce a method that we call DARC (Docking Approach using Ray-Casting) for matching these surface pockets to shape-complementary small molecules. Since many inhibitors of protein interactions are found to bind in shallow flat pockets on the protein surface [11,15], this method is specifically designed for such

protein-ligand interactions. Further, our method is intrinsically low-resolution, to allow for minor conformational variation of the protein upon ligand binding. We anticipate this method will provide suitable starting points for cases in which the protein structure is not perfectly optimal for the ligand, and cases in which further small conformational changes of the protein are expected to accompany ligand binding.

Computational Approach

Computational methods are implemented in the Rosetta software suite [17]; Rosetta is freely available for academic use (www.rosettacommons.org), with the new features described here included in the 3.6 release.

Scoring with DARC

To map the pocket shape on the protein surface, we use a variation of the LIGSITE [18] algorithm described fully in our earlier work [16]. In its most basic form, this method involves building a three-dimensional grid around a protein (or select region of a protein), then tagging the grid points that are covered by the protein. Next, each grid point not covered by the protein is examined to determine whether the point is bounded on two opposing sides by the protein along any direction of the grid, including diagonals. A continuous group of such special points are labeled a “pocket.” We then restrict this collection of “pocket” points to those “surface pocket” points that immediately contact the protein: these points together define the concave surface with which a ligand may interact.

From an “origin” point within the protein (determination of this point is described in Supplementary Methods), we then cast a series of rays towards the surface pocket; each ray connects the origin point to one “surface pocket” point (Figure 1a). For each ray we express the position of the “surface pocket” point relative to the origin using spherical coordinates ($\theta, \phi, \rho_{\text{pocket}}$); collectively these spherical coordinates represent a map of the binding pocket’s topography when viewed from the protein interior.

When viewed from the same vantage point (the origin), the topography of a suitably-bound ligand is expected to be highly complementary to that of the protein’s binding pocket. To evaluate the degree of complementarity, we therefore cast the same set of rays (defined by their angles θ and ϕ), and determine the distance at which each ray first intersects the ligand (Figure 1a). If the ligand is indeed complementary to the pocket, the distance to the ligand (ρ_{ligand}) for a given ray will closely match the distance to the pocket (ρ_{pocket}) for the corresponding ray.

We define the DARC “overlap score” as the degree to which the pocket and ligand topographies match, by summing over all rays as follows:

$$\text{DARC overlap score} = \sum_{\text{rays}} \left\{ \begin{array}{ll} c1 * (\rho_{\text{ligand}} - \rho_{\text{pocket}}) & \text{if } \rho_{\text{pocket}} < \rho_{\text{ligand}} \\ c2 * (\rho_{\text{pocket}} - \rho_{\text{ligand}}) & \text{if } \rho_{\text{ligand}} < \rho_{\text{pocket}} \\ c3 & \text{if ray does not intersect ligand} \\ c4 & \text{if ray does not intersect pocket} \end{array} \right\} \quad (1)$$

From a physical perspective, rays that reach the pocket before the ligand indicate underpacking between protein and ligand (*condition #1*); alternatively, rays that reach the ligand before the pocket indicate a steric clash (*condition #2*). Rays that do not intersect the small molecule at all are penalized as well (*condition #3*); from a physical perspective, this corresponds to the ligand not completely filling the pocket. Finally, we include a series of rays that do not intersect the pocket; intersection of these rays with the ligand indicates that the ligand is too large for the pocket (*condition #4*).

For a given set of rays, evaluating the intersection distance of each ray with the ligand is a task that is naturally very amenable to parallel computing; indeed, the need to efficiently carry out the related “ray tracing” operation provided motivation for development of graphics processing units (GPUs). We have recently demonstrated that the ray-casting approach we describe here can also be carried out on GPUs, leading to dramatic speed enhancement relative to the analogous calculation on a CPU alone [19].

Determination of the values for parameters $c1/c2/c3/c4$ used in these studies is described in Supplementary Methods.

Docking with DARC

Determining the optimal bound pose of a small molecule within a pocket (i.e. “docking”) entails varying the ligand position and orientation to minimize the DARC score (Equation 1). Though the first derivative of the DARC score can be calculated with respect to the ligand position and orientation, we instead carry out a derivative-free minimization using the particle-swarm optimizer implemented within the Rosetta software suite [17]. The ligand internal degrees of freedom are held fixed during this search; to account for ligand flexibility, we prebuild ligand “conformers” using OMEGA [20–22] and independently dock each of these. The final DARC score for a given ligand is taken to be the minimum DARC score from among each of its docked conformers.

Application of DARC for virtual screening is summarized through the workflow in Figure 1b. The starting pocket shape derives from a protein crystal structure, and can be an unbound structure, a bound structure in which the ligand has been discarded, or a structure generated from simulation (e.g. the “pocket optimization” simulations we have described for identifying low-energy pocket-containing conformations [16]).

Results

Building a virtual screening benchmark

Comparing DARC to other methods for virtual screening requires development of a benchmark to evaluate the performance of each method for correctly identifying known inhibitors from among “decoy” compounds. Assembling a large benchmark set is complicated by the relatively small number of available crystal structures of small-molecule inhibitors of protein interactions that have been solved in complex with their protein targets. Rather than assemble many tests with a very small number of active compounds (known inhibitors), we instead built two separate benchmark tests around protein interaction sites for which numerous chemically diverse inhibitors have been reported: Bcl-xL’s interaction with

Bak peptides, and XIAP's interaction with Smac/Diablo peptides. These two proteins (and their homologs) were excluded from parameterization of DARC (see Table S1 and *Supplementary Methods*), to ensure that DARC would not inadvertently be trained to perform well against these two protein targets.

When generating a set of “decoy” compounds for this screening experiment, it is important that these compounds be suitably matched to the “active” compounds. If the decoy compounds are systematically different in some physicochemical property, for example, a docking method may successfully pick out active compounds simply due to some implicit bias (e.g. by simply picking the largest compounds, or the most hydrophobic, etc.). On the other hand, it is important to consider that decoy compounds are not necessarily inactive, but rather “presumed inactive”: typically they have not been explicitly tested experimentally against the target proteins in the benchmark to ensure they are not active. If decoy compounds are too similar to active compounds in chemical structure, certain decoys may themselves be active, which will confound analysis of the results (because the “right answer” in the benchmark is to label decoy compounds as inactive).

To generate well-matched decoy compounds for these benchmarks, we drew from either i) a set of 85 drug-like compounds that comprise the Astex Diverse Set [23], or ii) a set of 6912 compounds reported to inhibit a different protein-protein interaction than our protein targets, compiled from the TIMBAL database [24,25].

To reduce the likelihood that “decoy” compounds might be active, we removed any decoy compounds that exhibited strong chemical similarity to a compound in the “active” set for this target (2D Fingerprint Tanimoto score [26] > 0.7). To ensure non-redundancy of the active set and non-redundancy of the decoy set, we also removed any compound with strong chemical similarity to another compound already present in the same set (2D Fingerprint Tanimoto score > 0.7).

For Bcl-xL this approach yielded 27 (diverse) active compounds, with 33 decoy compounds from the Astex set or 328 decoy compounds from the TIMBAL set. For XIAP this yielded 14 (diverse) active compounds, with 45 decoy compounds from the Astex set or 425 decoy compounds from the TIMBAL set. Though smaller than the TIMBAL decoy sets, the Astex decoys have the advantage that these compounds have typically been advanced through further optimization for their respective targets; this makes the Astex decoys less likely to exhibit (undesired) activity against Bcl-xL or XIAP.

In addition, we also generated a third set of decoy compounds using the DUD-E server [27]. In this case, we identified 50 “custom” property-matched decoys from each active compound. Thus, the 27 compounds active against Bcl-xL led to a set of 1350 DUD-E decoys, and the 14 compounds active against XIAP led to a set of 700 DUD-E decoys. The careful matching of the decoys' physicochemical properties to those of specific active compounds makes this a much more challenging benchmark; at the same time, however, there is an increased likelihood that some decoy compounds may themselves be active against the target protein.

Screening against a ligand-bound protein structure

The challenge in virtual screening entails identifying as many active compounds as possible from a library (true positives), while minimizing the number of inactive compounds incorrectly assigned as active (false positives). A very stringent score cutoff for any method will result in few compounds being assigned as active (such that some active compounds will be missed), whereas relaxing this cutoff will lead to many more compounds assigned as active (such that more false positives will be included). Since only a small fraction of the total compound library is expected to be further evaluated or validated experimentally, a typical goal for virtual screening is to rank active compounds at the very top of the sorted list (good “early” behavior), corresponding to the ability to build a small subset of the original library that is strongly enriched in active compounds [28]. The “late” behavior (the fraction of the library that must be screened to ensure no active compounds are missed) is irrelevant for virtual screening applications, since a screening tool that offers to eliminate only a small fraction of (inactive) compounds from a large library is not particularly useful.

We started from the crystal structures of Bcl-xL and XIAP, each solved in complex with an inhibitor that had been excluded from our library of active compounds. In addition to DARC, we used five other methods to score and rank each compound in this benchmark experiment. Four of these (DOCK [29,30], AutoDock [31], rDock [32], and PLANTS [33]) are receptor-based screening tools that we used to dock each compound in the protein interaction site of each protein. The fifth (ROCS [34–36]) is a ligand-based screening tool, which ranks compounds on the basis of how well their three-dimensional structure mimics the volume and chemical features of a template – in this case, the structure of the inhibitor solved in complex with the target protein. Parameters used for each of these other packages are included in the Supplementary Methods.

Given the rank order of each compound using each of the three methods, we prepared separate receiver operating characteristic (ROC) curves for both Bcl-xL and XIAP, corresponding to an experiment in which we screen a library comprised of active compounds and Astex decoys (Figure 2a). Each point on such a plot corresponds to a given score cutoff, and each point indicates the fraction of active compounds collected at this cutoff, as a function of the inactive compounds incorrectly assigned as active at this cutoff. On this plot, the critical “early” behavior described earlier corresponds to a steep rise at the leftmost part of the plot: a useful method for screening must pick out as many active compounds as possible while accumulating a very small fraction of the inactive compounds. All three methods outperform the behavior that would result from randomly ranking the compounds in the library (*grey curve*), especially in this “early” region. We also observe similar results when carrying out the same screening experiment using decoy compounds drawn from the TIMBAL set rather than from the Astex set (Figure S1). With the exception of ROCS applied to XIAP, meanwhile, none of the methods perform at a comparable level when used to distinguish known active compounds from their closely-matched decoys in the DUD-E set (Figure S2).

As seen in representative examples of docked active compounds (Figure 2b), the DARC-generated models exhibit clear shape complementarity that allowed these compounds to be (correctly) identified as true ligands; experimentally-derived structures of these compounds

in complex with their protein partners have not been reported to date. While low-resolution considerations of shape complementarity alone cannot be sufficient to reliably predict binding affinity or details of the interactions, it is striking that for these two benchmark tests the simple approach employed by DARC successfully distinguishes active from inactive compounds with similar accuracy as more sophisticated methods. We further note that the performance of DARC in this benchmark is remarkably robust to the location of the “origin” point from which rays emanate, provided a reasonable approach is used to define this point (Figure S3). Thus, these observations support the hypothesis that the crude pocket shape on the protein surface provides a strong restriction on the chemical “shape space” of inhibitory compounds.

Screening against a “pocket optimized” protein structure

The fact that a protein interaction site is likely to undergo a conformational change upon inhibitor binding [15] makes the choice of receptor conformation of particular interest for this target class, especially given that the performance of most approaches decreases considerably when an unbound starting structure is used in place of a ligand-bound starting structure [37]. A variety of methods exist for incorporating receptor flexibility, typically by docking against multiple receptor conformations from either crystal structures [38–42] or simulation [43–45].

Rather than start from unbound crystal structures of Bcl-xL and XIAP – which lack the surface pocket required for inhibitor binding – we instead started from conformations generated using the “pocket optimization” approach we have recently described [16]. Briefly, we developed a biasing potential that drives conformational sampling in molecular mechanics simulations towards conformations containing surface pockets suitable for small molecule binding. This biasing potential operates solely on simple geometric considerations, so that the shapes of the resulting surface pockets are influenced solely by properties of the protein surface and *not* knowledge of any particular ligand. Starting from unbound crystal structures, we carried out simulations to generate “pocket optimized” conformations of Bcl-xL and XIAP: conformations containing a surface pocket, but without the precise details that would be encoded in a ligand-bound structure (our application of this approach to generate these protein conformations is described in Supplementary Methods).

We re-screened the same compound library described earlier, this time using a “pocket optimized” conformation of Bcl-xL or XIAP (Figure 3a). We did not include ROCS in this stage of the benchmark, since there is no structure of a cognate ligand available for this protein conformation that could be used as a template. Unsurprisingly, the performance of DOCK, AutoDock, rDock, and PLANTS all suffer dramatically in this more challenging regime. In contrast, the performance of DARC is remarkably similar to the earlier experiment, demonstrating that the low-resolution nature of the underlying calculations makes DARC relatively robust to slight mismatches between protein and ligand. Indeed, representative examples of active compounds docked to these alternate conformations demonstrate that the overall pocket shapes remain complementary to these ligands (Figure 3b). We note that these observations also hold when carrying out the same experiment using decoy compounds drawn from the TIMBAL set rather than the Astex set (Figure S4), but

that again all of these methods exhibit diminished performance against the DUD-E set (Figure S5).

In summary, slight differences in these protein conformations lead to diminished performance from other receptor-based docking methods in this regime, but DARC performs equally well in this experiment as in the previous benchmark that used protein conformations from ligand-bound crystal structures. This finding is confirmed by extending this experiment to consider a series of five more distinct low-energy conformations for each protein, as generated by the “pocket optimization” approach: we find that the performance of DARC in this benchmark is notably insensitive to the precise details of the protein conformation (Figure 3c).

Collectively, these experiments point to different regimes in which each class of method is expected to prove superior. Detailed docking methods, such as DOCK, AutoDock, rDock, and PLANTS, are expected to out-perform DARC if the protein structure is solved in complex with a ligand of similar chemotype to the desired compounds. By contrast, DARC allows greater discrimination if the protein conformation is not quite optimal for the ligand: this can occur when one wishes to discover inhibitors with a radically different chemotype than known ligands, or when the protein conformation derives from simulation.

Identifying novel Mcl-1 inhibitors using DARC

To evaluate the performance of DARC in a realistic application, we next applied this approach to screen for novel classes of inhibitors for Mcl-1, a member of the Bcl-2 family of proteins. Individual members of this protein family can serve either a pro- or anti-apoptotic role, and interact with one another through a structurally-conserved binding motif. Small-molecule inhibitors of Bcl-2/Bcl-xL have shown promising efficacy in overcoming chemo/radioresistance in various tumor models including prostate cancer [46–50]. One such compound is ABT-737, a potent Bcl-2/Bcl-xL inhibitor [50]; a more recent Bcl-2-selective derivative of this compound is currently in clinical trials [51]. Recent studies have shown that cancer cells resistant to ABT-737 have high levels of Mcl-1, and that knockdown of Mcl-1 promotes ABT-737-induced apoptosis [35,48,49,52,53]. Taken together, these observations motivate the pressing need for development of potent and selective inhibitors of Mcl-1 in treating of a variety of cancers, to be used either as a single agent or in combination with inhibitors of Bcl-2/Bcl-xL [54].

The strong evidence supporting the potential impact of effective Mcl-1 inhibitors has spurred intense efforts using a variety of complementary approaches [54,55], including high-throughput screening of standard libraries using a fluorescence polarization competition assay [56,57], and of an sp^3 -rich library using differential scanning fluorimetry [58]. Aiming to build upon the success of the ABT-737 series for Bcl-2, several groups have since applied similar NMR-based fragment screening to Mcl-1; fragment hits were prioritized on the basis of ligand efficiency (binding free energy per heavyatom), first yielding weakly-binding compounds that were subsequently merged or elaborated to give inhibitors with improved potency at the expense of some ligand efficiency [59–62]. Inspired by the helical conformation adopted by Mcl-1's interaction partners, a number of groups have grafted the interaction partners' sidechains onto small-molecule scaffolds that allow them to be

presented in a similar geometry [63–65], or have used “stapled” derivatives of the helical peptides [66–68]. Such approaches predicated on mimicry of a known helical binding partner can rapidly lead to potent inhibitors, but the large chemical scaffolds required to recapitulate the helical geometry diminish from their ligand efficiency.

Like other members of the Bcl-2 family, Mcl-1’s interaction partners bind to an exposed hydrophobic groove on the protein surface [69]. We recently applied our “pocket optimization” approach to this protein surface, and generated ensembles of low-energy Mcl-1 conformations that present surface pockets suitable for small molecule binding [70]. Among these ensembles we find conformations similar to those observed in crystal structures of Mcl-1 bound to diverse small-molecule ligands, and we also observe “distinct” conformations that have not yet been observed in any experimentally-derived structures [70]. We elected to use DARC to screen chemical libraries for compounds that would complement one of two Mcl-1 conformations: either an experimentally-derived conformation, or a conformation derived from “pocket optimization” simulations. At the time we initiated these studies no unbound structure was available [71], so for the former we used a peptide-bound crystal structure (PDB ID 2pqq) [69]. For the latter, we used the lowest-energy pocket-containing conformation generated from 1000 independent trajectories.

Drawing from the ZINC database [72], we compiled a small virtual library corresponding to 62,442 highly-diverse compounds with drug-like properties (MW 500 Da, xlogP 5, etc., see Supplementary Methods) available for immediate purchase from a commercial vendor. We then applied DARC to separately dock and rank each compound using the peptide-bound structure of Mcl-1, and also separately using the “pocket optimized” conformation of Mcl-1.

The docked models produced by DARC are intrinsically low-resolution, since they rely only on matching the protein-ligand topography. To achieve further high-resolution discrimination, we therefore included an additional final step (Figure 1b): for each of the top ranked 10% of the complexes produced by DARC we carried out fullatom gradient-based minimization in Rosetta, using the standard Rosetta fullatom energy function [17] and ligand parameter definitions [73]. All internal dihedral angles of the protein were included as degrees of freedom, along with the ligand position and orientation relative to the protein. We then filtered the resulting models to remove those with no intermolecular hydrogen bonds or an abundance of buried unsatisfied polar groups. From the remaining models, we then purchased the top-scoring 21 compounds on the basis of DARC score (Table S2). Of these 21 compounds, 11 were identified by screening against the (peptide-bound) crystal structure of Mcl-1, and the other 10 were from screening against the “pocket optimized” structure.

The structural basis for the favorable scores in our DARC screen is evident from the models of these compounds in complex with Mcl-1 (Figure 4a). In each case the ligand – in its modeled conformation – exhibits exquisite shape complementarity for the protein surface; this is unsurprising, given that the DARC score is expressly built to identify the ligand position, orientation, and conformation that will maximize similarity to the topography of the protein surface. Nonetheless, there is also clear diversity amongst the models: the ligand shapes are different from one another, and they fill the protein surface pocket in distinctly different ways. In the case of M0, for example, a hydrophobic groove on the Mcl-1 surface

is neatly complemented by the spatial arrangement of the two aromatic rings in the selected ligand conformation.

The selected compounds also represent a diverse array of different chemical scaffolds (Figure 4b), with no evident similarity to one another. It must be noted that of the compounds shown, only M7 (a natural product) is devoid of potentially reactive functional groups. This is, of course, a reflection of the screening library: in a typical drug discovery application, such compounds should be removed from the library prior to screening since their potential for advancement is likely to be extremely limited [74]. For the purpose of evaluating the performance of DARC in this virtual screening experiment, however, we did not exclude such compounds.

By visual inspection, the structures of these compounds bear no obvious resemblance to any known inhibitors of Mcl-1, or to any known inhibitors of any other Bcl-2 family members. To systematically identify biologically active compounds most resembling these compounds, we applied the Similarity Ensemble Approach (SEA); this method uses ligand similarity to identify likely receptors for a query ligand [75]. In most cases SEA identifies chemical scaffolds with clear resemblance to our queries, but these related compounds have each been described in the context of very different activities (Figure 4c); even delving further into each list, we did not find similarity to other inhibitors of Bcl-2 family members. This observation highlights the fact that these compounds would have been very difficult to pick out using ligand-based screening, due to their lack of similarity to any known inhibitors of this protein family.

We next examined the ability of each compound to inhibit the interaction of Mcl-1 with one of its binding partners. We used a fluorescein-labeled peptide derived from the Noxa protein, which exhibits an increase in polarization upon Mcl-1 binding (Figure S6a). Upon addition of a known Mcl-1 inhibitor, AT-101 (i.e. R-(-)-gossypol), we observe a dose-dependent decrease in polarization that confirms this compound competes with Noxa for Mcl-1 binding, and is consistent with previously reported data [76] (Figure S6b). Using this assay we found that six of the 21 compounds from our computational screen appeared to inhibit the interaction between Mcl-1 and its cognate peptide with K_i values ranging from 1.2 to 21 μM (Figure 4d). These results were not affected by the presence of absence of detergent (CHAPS), suggesting that compound aggregation was not contributing to the observed inhibition. We then used an orthogonal label-free assay based on bio-layer interferometry (BLI) to test for direct binding of these compounds to Mcl-1, and through this assay confirmed the activity for four of these six compounds (see Supplementary Results) (Figure S7): M0, M1, M5, and M7 (Table 1). In all four cases we found that inhibition did not increase over time in the competition assay, and that the BLI signal showed a dissociation phase (decreased amplitude) upon transfer into buffer lacking the compound of interest; together, these imply that all four compounds bind in a reversible manner.

Given the chemical structures of these four most potent compounds, these observed K_i values correspond to ligand efficiency (binding free energy per heavyatom) of 0.21 to 0.27 kcal/mol•heavyatom, which is typical of other promising small-molecule inhibitors of protein-protein interactions reported in the literature [8]. With respect to other Mcl-1

inhibitors in particular, these four initial hits are less active than the most potent compounds arising from screens of libraries designed to mimic the sidechain interactions of helical cognate peptides [63–65]; however, the DARC hits exhibit greater ligand efficiency, due to the extensive size of the compounds designed for mimicry of the helix.

Superposition of the modeled complexes for each of the DARC hits with the structure of Mcl-1 bound to a cognate peptide reveals that these compounds are decisively *not* recapitulating the interactions of the helical peptide (Figure 5a). While the cognate peptide uses a cluster of aliphatic residues at its N-terminus and aromatic residues at its C-terminus, compound M0 (as a representative example) occupies only the regions of the Mcl-1 surface engaged by the N-terminus of the peptide. Given that it is not constrained to fit onto the backbone of the helix, this compound also fits more deeply into the binding pocket on the Mcl-1 surface. In contrast, the most potent of the helix-mimetics was designed to recapitulate the complete set of interactions provided by the helical template [63]. The sidechains that comprise these interactions are separated by a span of 15 residues in sequence, and while this led to an inhibitor with K_i of 0.2 μM , the large chemical scaffold needed to present these interactions reduced the ligand efficiency of this compound to 0.21 kcal/mol•heavyatom.

While the underlying methodology used in our DARC screen did not explicitly select for compounds that bind in place of the N-terminus of the cognate peptide, each of the resulting hits complemented this region. This is in stark contrast to the binding modes observed in crystal structures of Bcl-xL solved in complex with potent inhibitors such as ABT-737 and WEHI-539: each of these bind to the region of the protein surface used by the C-terminus of the cognate peptide, and thus have relatively little overlap with the models of the DARC hits (Figure 5b). In the time since we carried out our computational screen, crystal structures of Mcl-1 solved in complex with four distinct classes of inhibitor have been reported [58–60,77], and in each case the ligand occupies the region of the protein surface corresponding to the N-terminus of the cognate peptide. Thus, the lack of DARC hits complementing the protein surface used by inhibitors of Bcl-xL may indicate a lack of druggability for the analogous surface of Mcl-1.

In contrast to Mcl-1 inhibitors derived by mimicry of the cognate helix, the most promising hits emerging from fragment screens initially proved less potent than those identified by DARC but had superior ligand efficiency [59–62]. One representative class of fragments inhibited Mcl-1 with K_i of only ~ 100 μM , albeit with a promising ligand efficiency of about 0.3 kcal/mol•heavyatom [59]. NMR-derived models of the binding mode for this class of compound suggested that they induce a conformational change in Mcl-1, which in turn produces a very deep pocket required for binding; upon merging with another fragment class, crystallography confirmed that the larger compound also induced this conformational change [59]. In retrospect, each of the Mcl-1 crystal structures in complex with an inhibitor (solved after completion of our screen) revealed a conformational change that allowed the ligand to access a deep pocket that was not evident from the unbound or peptide-bound structures [58–60,77] (Figure 5c). The observation that extensive ligand burial is related to high ligand efficiency (across many small-molecule inhibitors of protein-protein interactions) [11] is thus consistent with the impressive ligand efficiency of these (deeply

buried) Mcl-1 fragment hits. However, such conformations were not among those included in our DARC screen, and accordingly the DARC hits could not have taken advantage of the deep pockets presented upon conformational rearrangement of the protein.

Evaluating the DARC model of M0 binding

An important benefit that can arise from the use of structure-based modeling (such as DARC) is the prospect of immediately using the model of the complex to guide design of derivatives, once an initial hit compound is validated experimentally. In order to test whether a DARC model can indeed serve this purpose, we next sought to test whether the tighter-binding derivatives could be designed from the DARC model. Despite the fact that M0 (2',4-dihydroxy-3,4',6'-trimethoxychalcone) was the least potent of the four hits in our original screen, and concerns of promiscuity because similar chalcones have been found to exhibit other diverse activities in biological assays, the straightforward synthetic accessibility of the chalcone series prompted us to select this scaffold for further investigation. While this compound would not necessarily be an optimal starting point for advancement into a drug or biological probe, then, selecting of this compound as a starting point allowed us to make and test more derivatives than would have been possible with the other three validated DARC hits.

Starting from compound M0, we therefore we sought to ask whether such a model could be used to explain the structure-activity relationship (SAR) observed across a series of related compounds. Since our goal was to gauge the veracity of the modeled complexes that led to selection of these compounds, and not simply to identify the most potent derivatives, we further restricted our study to compounds very close to the parent compound, M0 – we reasoned that these would best allow us to test our model of the structure, reducing the likelihood of conformational changes leading to drastic and unexpected SAR.

Before designing a series of analogs, however, we noted that other chalcones are known to form covalent adducts to proteins (the chalcone α,β -unsaturated carbonyl system is a Michael acceptor and thus may react with unpaired cysteine sidechains) [78,79]. We therefore carried out analysis via surface plasmon resonance (Figure S8) and HSQC chemical shift mapping (Figure S9) to further characterize this interaction (see Supplementary Results). Taken together, the stoichiometric Hill coefficient [80] and insensitivity to the presence of detergent [81] in the fluorescence polarization assay, the fact that the SPR response did not exhibit super-stoichiometric behavior [82], and the observation of a small number of distinct chemical shift differences in the HSQC spectrum [83] all provide evidence that the observed activity of M0 is not due to compound aggregation that deactivates Mcl-1, but results from specific binding. The NMR experiment also supports two features of our model for M0 binding: first that the M0 binding mode is not completely overlapping with that of the Bcl-xL inhibitors (Figure 5b), and second that M0 binding does not induce extensive protein conformational changes as seen in other Mcl-1 inhibitors (Figure 5c).

We then proceeded to produce a variety of M0 derivatives, using a simple aldol condensation reaction (Figure S10). We were concerned that designing these analogs from our model of the M0/Mcl-1 complex might bias the resulting series towards certain regions of chemical

space, and thus not provide a fair assessment of our model. In order to evaluate the structure-activity relationship more objectively, we therefore designed a set of compounds by selecting the most readily available starting materials, based only upon the chemical structure of M0 and without consideration of our model of the complex. Using the same fluorescence polarization competition assay described earlier, we characterized a total of 27 analogs of M0 (Table 2, Figure S11). As described below, the SAR deduced from this series presents a compelling narrative when interpreted using our model of M0 binding (Figure 5d).

The most potent of these M0 analogs are D1 and D2: they each inhibit Mcl-1 about 7-fold more potently than M0. Compounds D1 and D2 differ only by a single methoxy versus fluorine substitution at the 3-position; M0 harbors a methoxy here. This methoxy group is exposed in our model of the M0/Mcl-1 complex, consistent with the observation from D1 versus D2 that its replacement with fluorine does not affect activity. Compounds D1 and D2 together share three differences relative to M0: both have a chlorine at the 5-position, both have a methoxy group at the 5'-position, and both lack the 2'-hydroxyl of M0. The latter two positions are exposed in our model of M0 binding, and are thus assumed not to strongly affect activity. In contrast, the 5-position (unsubstituted in M0) faces towards the protein, and points towards a small cavity (Figure 5d): thus, the improved packing resulting from this substitution may explain the slightly enhanced potency of D1/D2 relative to M0.

The only other two compounds in this series with potency exceeding that of M0 are D3 and D4; this equipotent pair again differs only by the same fluorine versus methoxy substitution that distinguished D1 and D2. Relative to D1/D2, compounds D3/D4 lack the 5'- and 6'-methoxy groups. While the 5'-position is unsubstituted and exposed in our model of the M0/Mcl-1 complex, the 6'-methoxy group is deeply buried in a hydrophobic cavity (Figure 5d); loss of this interaction may explain the reduced potency of D3/D4 relative to D1/D2.

Across this series of 27 M0 analogs, only compounds D1-D4 proved more potent than M0. While the SAR described above is consistent with our model of M0 binding, these results are also notably *inconsistent* with the behavior expected if these compounds were simply carrying out thia-Michael additions. Kinetic studies using a model thiol (cysteamine) demonstrate that the presence of the 6'-methoxy group decreases reactivity, and the 2'-hydroxy group increases reactivity [84]. Through the comparisons above, we find that the most potent compounds for inhibition of Mcl-1 correspond to those expected to be *least* reactive, further implying that the inhibition we observe is not due to reactivity of these compounds with Mcl-1's unpaired cysteine sidechain.

Through the results presented in Table 2 we demonstrate that chalcones are not "privileged scaffolds" [85,86] for inhibition of Mcl-1, an assertion that is best supported by specific examples. Compounds D7/D12/D13/D14 lack the key 6'-methoxy group described earlier (Figure 5d), and accordingly are less potent than M0. Compound D5 preserves each of the M0 functional groups that interact with Mcl-1 in our model (i.e. the 6'-methoxy group and the 4-hydroxyl group), and accordingly exhibits a very similar inhibition constant as M0. Compound D8 maintains the substituents of D2 on one ring, but lacks the substituents on the other ring (including the 6'-methoxy group), and is thus less potent. Compounds

D9/D10/D11 maintain the 6'-methoxy group but harbor a variety of alternate substituents on the other ring, making them less potent as well. Compounds harboring extra fused rings (D20/D21/D22/D23) are inevitably less potent, as are compounds in which the chalcone linkage has been replaced with a flavone (D24/D25/D26/D27). Overall, the assertion that chalcones are not privileged scaffolds for inhibition of Mcl-1 or generic helix-mimetics is supported by the observation that most analogs presented here – designed without consideration of our model of binding – inhibit Mcl-1 less potently than the parent compound M0. In summary then, it is the precise complementarity of select compounds for the surface of Mcl-1 that dictate their activity, and not some property of this chemical scaffold.

Ultimately, direct structural evidence will be required to confirm that these compounds are indeed binding via the designed pose. Unfortunately, our efforts to crystallize Mcl-1 in complex with a member of this chemical series have not yet proven successful. Validating the models that lead to selection of active compounds through structural biology will be important not only for retrospective evaluation of the DARC models, but also as a starting point for inspiring design of subsequent analogs. Accordingly, we used DARC to compare our model of the M0/Mcl-1 complex to a model in which M0 was replaced with D1 (Figure 5e). We find that DARC scores D1 even more favorably than M0, due to the improved packing resulting from substitution of chlorine at the 5-position. Thus, we expect that compounds with even better shape complementarity for Mcl-1 than those selected in our initial screen – compounds such as D1 – can be identified by screening larger chemical libraries with DARC. By further analogy to D1, we anticipate that such compounds will also exhibit superior potency relative to the initial hits described in this first screen. Despite the improved potency relative to M0, we nonetheless note that compound D2 does not exert the desired biological effect using an *in vitro* cellular assay (see Supplementary Results) (Figure S12). This underscores the need to test candidate inhibitors in cellular assays as well, since biochemical activity may not necessarily translate to cellular activity. Indeed, careful analysis has shown that a number of other inhibitors, each with comparable binding affinity for Mcl-1 as that of compound D2, also appear not to act on Mcl-1 in cells [57,87,88].

Discussion

Screening on the basis of surface topography

Since the protein surfaces of “traditional” drug targets have evolved to bind some cognate small-molecule ligand, they typically include a deep pocket or groove that can be targeted by inhibitors. In contrast, protein interaction sites typically do not include such surface features, and accordingly small-molecule inhibitors acting at these sites rely on shallower bound poses [11]. Because these shallow binding modes prove challenging to predict by conventional docking approaches, modern virtual screening tools exhibit diminished performance when used to predict inhibitors of protein interactions (relative to their performance when applied to “traditional” drug targets) [11]. Here we present DARC, an alternate approach to docking and virtual screening. By matching the topography of the protein surface to the buried face of the ligand, we find that DARC outperforms popular

canonical tools (DOCK, AutoDock, rDock, and PLANTS) when screening for inhibitors of protein-protein interactions.

Given the importance of shape complementation for binding, several other fast approaches have been described to rapidly evaluate poses and enable large-scale virtual screening. These include methods based on Fourier correlation theory [89,90], spherical harmonics [91,92], geometric hashing [93], and negative images [94–96]. Most recently, others have used variations of ray-casting to compare internal pockets in proteins to one another [97,98]. In addition to their intended usage, a subtle – but very important – difference between the latter work and DARC is the origin from which rays emanate. This other study casts rays from within the pocket (at the center of mass), resulting in a shape description that is most useful when the pocket is mostly (or completely) enclosed by the protein [97,98]. In contrast, DARC casts rays that originate from “behind” the pocket (i.e. inside the protein): this instead emphasizes shape complementarity at the deepest regions of the binding groove, and is thus better suited for describing the shallow bound poses typical of small-molecule inhibitors acting at protein interaction sites.

While shape complementarity is clearly a necessary feature of ligands that will bind to a protein surface, this alone obviously cannot be sufficient. In the present study, we used DARC to optimize and evaluate surface shape complementarity without consideration of electrostatic complementarity or solvation effects; for this reason, we included a “re-ranking” step in our virtual screen against Mcl-1 to filter out models with no intermolecular hydrogen bonds or an abundance of unsatisfied buried polar groups. In retrospect, the hydrophobicity of the Mcl-1 surface pocket provided a convenient testing ground for the ability of DARC to identify shape-complementary ligands, without the complication of polar groups on the protein surface. In future, however, we expect that incorporation of electrostatic complementarity into DARC lead to improved performance in virtual screens against protein surfaces that display surface polar groups [99].

Approaches for identifying inhibitors of protein-protein interactions

In many cases, artificial ligands that bind to a specific protein surface bear some resemblance to an endogenous ligand that also binds to that surface: analysis of precedent drug targets shows that knowledge of a substrate, product, or effector with “drug-like” physicochemical properties is a good predictor of druggability for a given protein surface [100]. Analogs of endogenous ligands can also provide a starting point for designing inhibitors [101]. In the case of protein interaction sites, however, the natural ligand is not a small molecule and thus does not provide an obvious template from which to start. This has driven advances in new methodologies for identifying inhibitors of these “non-traditional” targets, most notably including mimicry of key interacting groups of the protein partner [63–65,102–104] and fragment-based approaches [59,105–107].

In addition to chemical scaffolds that mimic secondary structural elements [63–65], exciting new computational approaches have facilitated identification of compounds that instead mimic the geometric orientation of key “hotspot” or “anchor” sidechains in a protein-protein interface [102–104]. While either mimicry-based approach may provide a robust starting point for recapitulating the interactions of a protein partner using a small-molecule scaffold,

both are likely to encounter the same intrinsic limitation: ligand efficiency is unlikely to exceed that of the interacting groups from the protein partner, which in turn is usually less than that those achieved by small-molecule inhibitors that do not explicitly mimic interactions of the protein partner [15]. Accordingly, an important advantage of *de novo* structure-based screening methods that do not rely on mimicry, such as DARC, is the potential to identify inhibitors that achieve superior ligand efficiency through deeply buried interactions not available to a protein-based ligand.

In contrast, fragment-based approaches often prioritize small compounds with very high ligand efficiency from the outset: this allows exploration of deep crevices that would not necessarily be evident from the structure of the protein-protein complex. Here, the challenge often lies in elaborating initial fragment hits (by growing, merging, or linking them) into larger compounds that maintain these highly productive interactions [108,109]. In our DARC screen against Mcl-1, we focused on compounds larger than traditional fragments: this led to initial hits with easily detectable activity, albeit at some expense of ligand efficiency. Much as the “multiple solvent crystal structures” method [110] uses very small fragments to probe the protein surface for productive interactions with isolated functional groups, screening a commercially-available library of “prototype” compounds – as we have done here – allows rapid evaluation of specific protein-ligand interactions predicted *in silico*, and can provide potential chemical scaffolds for further optimization.

Screening against multiple protein conformations

The compounds identified as “hits” by DARC, and by structure-based screening methods in general, naturally depend on the conformation of the protein target. With respect to inhibitors of protein-protein interactions in particular, a number of examples have been shown to bind concave surface pockets that are absent in the corresponding unbound protein structures: these “cryptic” pockets are revealed by local conformational changes associated with inhibitor binding [15]. In the case of Bcl-2 family members, the protein surface can also adopt a number of different conformations to accommodate ligands with radically different shapes and chemical properties [70].

Variation in the protein structure is most commonly included in virtual screening by first generating an ensemble of relevant conformations from crystal structures [38–42] or simulation [43–45], and then separately screening against each of these conformations. Nonetheless, identifying the optimal conformations to include in these ensembles is still a challenging task [111–113]. By preferentially exploring conformations that contain a surface pocket suitable for binding some (unspecified) ligand [16], the “pocket optimization” approach provides a natural complement to DARC screening. In principle, each of these low-energy pocket-containing conformations represents a starting point for screening, and capturing the potential diversity of pockets on the protein surface will be essential for fully realizing the available diversity of potential inhibitors.

In our Mcl-1 screen, DARC identified inhibitors both when screening against the peptide-bound conformation and when screening against the pocket-optimized conformation. Due to the conformations selected for these screens, however, our models did not take advantage of deep pockets that have been observed in recent inhibitor-bound crystal structures of Mcl-1

[58–60,77] (Figure 5c): the particular pocket-optimized conformation used in our screen was much more similar to the peptide-bound conformation, making these deep pockets unavailable to DARC. Careful analysis showed that similar conformations to those observed in the inhibitor-bound crystal structures were indeed represented in the ensemble produced by pocket optimization [70], but were not included here because we only screened against a single pocket-optimized conformation. In future, then, inclusion of additional pocket-containing protein conformations may allow identification of inhibitors that bind more deeply, and exhibit improved ligand efficiency.

Advantages to structure-based virtual screening

Enhanced and reliable tools for structure-based virtual screening will enable development of new tool compounds, particularly in academic settings where the costs associated with large-scale biochemical (or phenotypic) screening can be prohibitive. However, the ability to precisely complement a specific surface pocket also makes structure-based virtual screening particularly attractive in a variety of other contexts that may prove challenging for traditional biochemical screening. These include building “conformation-selective” inhibitors (such as compounds that are sensitive to the phosphorylation state of a kinase activation loop [114]), targeting sites that enable unique binding kinetics (such as a newly-discovered pocket on ERK1/2 [115]), and building allosteric inhibitors that address both the wild-type and drug-resistant isoforms of a target (such as BCR-ABL [116] or HIV-1 protease [117]).

In light of the plasticity of many protein interaction sites [15], these tools may also prove particularly useful for designing inhibitors that exhibit specific selectivity profiles. We have shown that the activity of a given compound against various members of the Bcl-2 family can be predicted from whether or not each protein samples a complementary surface pocket [70]. In other words, a given compound will inhibit those Bcl-2 family members that include conformations to suitably accommodate this compound, but will not inhibit other family members that cannot adopt such conformations. By identifying “common” pockets harbored on the surfaces of multiple family members, it may be possible to identify inhibitors designed to act against multiple family members (pan-inhibitors). “Pocket optimization” simulations also reveal highly unique pockets for each family member, that are sampled by one family member but are not accessible to any other family member [70]. Thus, DARC’s ability to identify ligands that precisely address a given family member’s unique “signature” pockets may also provide a means to identify highly selective compounds at a very early stage of development.

Methods

Detailed descriptions of computational and experimental methods are provided in the Supplementary Methods section, along with sample command-lines and instructions for using DARC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Amy Keating for providing the plasmid encoding Mcl-1, and we thank Anne Cooper and Philip Gao for assistance with protein production. We are grateful to OpenEye Scientific Software (Santa Fe, NM) for providing an academic license for the use of OMEGA and QuacPac. We are grateful to the developers of the DOCK, AutoDock, rDock, and PLANTS software for providing academic licenses for the use of these programs. This work was supported by grants from the National Institute of General Medical Sciences (1R01GM099959, 5P50GM069663, 8P30GM103495, and 8P20GM103420), and the National Center for Research Resources (5P30RR030926 and 5P20RR017708). This work was also supported by the National Science Foundation through XSEDE allocation MCB130049, the University of Kansas Undergraduate Research Awards (S.M.), and the Alfred P. Sloan Fellowship (J.K.).

References

1. Domling A. Small molecular weight protein-protein interaction antagonists: an insurmountable challenge? *Curr Opin Chem Biol.* 2008; 12(3):281–291. [PubMed: 18501203]
2. Whitty A, Kumaravel G. Between a rock and a hard place? *Nat Chem Biol.* 2006; 2(3):112–118. [PubMed: 16484997]
3. Murray CW, Rees DC. The rise of fragment-based drug discovery. *Nat Chem.* 2009; 1(3):187–192. [PubMed: 21378847]
4. Akram ON, DeGraff DJ, Sheehan JH, Tilley WD, Matusik RJ, Ahn JM, Raj GV. Tailoring peptidomimetics for targeting protein-protein interactions. *Mol Cancer Res.* 2014; 12(7):967–978. [PubMed: 24642350]
5. Cummings CG, Hamilton AD. Disrupting protein-protein interactions with non-peptidic, small molecule alpha-helix mimetics. *Curr Opin Chem Biol.* 2010; 14(3):341–346. [PubMed: 20430687]
6. Whitby LR, Boger DL. Comprehensive peptidomimetic libraries targeting protein-protein interactions. *Acc Chem Res.* 2012; 45(10):1698–1709. [PubMed: 22799570]
7. Cierpicki T, Grembecka J. Targeting protein-protein interactions in hematologic malignancies: still a challenge or a great opportunity for future therapies? *Immunol Rev.* 2015; 263(1):279–301. [PubMed: 25510283]
8. Arkin MR, Tang Y, Wells JA. Small-Molecule Inhibitors of Protein-Protein Interactions: Progressing toward the Reality. *Chem Biol.* 2014; 21(9):1102–1114. [PubMed: 25237857]
9. Kar S, Roy K. How far can virtual screening take us in drug discovery? *Expert Opin Drug Discov.* 2013; 8(3):245–261. [PubMed: 23330660]
10. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem.* 2014; 14(16):1923–1938. [PubMed: 25262799]
11. Gowthaman R, Deeds EJ, Karanicolas J. Structural properties of non-traditional drug targets present new challenges for virtual screening. *J Chem Inf Model.* 2013; 53(8):2073–2081. [PubMed: 23879197]
12. Cihlar T, Ray AS. Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. *Antiviral Res.* 2010; 85(1):39–58. [PubMed: 19887088]
13. Hodgson DR, Schroder M. Chemical approaches towards unravelling kinase-mediated signalling pathways. *Chem Soc Rev.* 2011; 40(3):1211–1223. [PubMed: 21152652]
14. Jacobson KA. Structure-based approaches to ligands for G-protein-coupled adenosine and P2Y receptors, from small molecules to nanoconjugates. *J Med Chem.* 2013; 56(10):3749–3767. [PubMed: 23597047]
15. Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature.* 2007; 450(7172):1001–1009. [PubMed: 18075579]
16. Johnson DK, Karanicolas J. Druggable protein interaction sites are more predisposed to surface pocket formation than the rest of the protein surface. *PLoS Comput Biol.* 2013; 9(3):e1002951. [PubMed: 23505360]
17. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman K, Renfrew PD, Smith CA, Sheffler W, Davis IW, Cooper S, Treuille A, Mandell DJ, Richter F, Ban YE, Fleishman SJ, Corn JE, Kim DE, Lyskov S, Berrondo M, Mentzer S, Popovic Z, Havranek JJ, Karanicolas J,

- Das R, Meiler J, Kortemme T, Gray JJ, Kuhlman B, Baker D, Bradley P. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* 2011; 487:545–574. [PubMed: 21187238]
18. Hendlich M, Rippmann F, Barnickel G. LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *J Mol Graph Model.* 1997; 15(6):359–363. 89. [PubMed: 9704298]
19. Khar KR, Goldschmidt L, Karanicolas J. Fast docking on graphics processing units via Ray-Casting. *PLoS One.* 2013; 8(8):e70661. [PubMed: 23976948]
20. Hawkins PC, Skillman AG, Warren GL, Ellingson BA, Stahl MT. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *Journal of chemical information and modeling.* 2010; 50(4):572–584. [PubMed: 20235588]
21. Hawkins PC, Nicholls A. Conformer generation with OMEGA: learning from the data set and the analysis of failures. *Journal of chemical information and modeling.* 2012; 52(11):2919–2936. [PubMed: 23082786]
22. OMEGA version 2.4.3. OpenEye Scientific Software SF, NM. <http://www.eyesopen.com>
23. Hartshorn MJ, Verdonk ML, Chessari G, Brewerton SC, Mooij WT, Mortenson PN, Murray CW. Diverse, high-quality test set for the validation of protein-ligand docking performance. *J Med Chem.* 2007; 50(4):726–741. [PubMed: 17300160]
24. Higuero AP, Schreyer A, Bickerton GR, Pitt WR, Groom CR, Blundell TL. Atomic interactions and profile of small molecules disrupting protein-protein interfaces: the TIMBAL database. *Chem Biol Drug Des.* 2009; 74(5):457–467. [PubMed: 19811506]
25. Higuero AP, Jubb H, Blundell TL. TIMBAL v2: update of a database holding small molecules modulating protein-protein interactions. *Database (Oxford).* 2013; 2013 bat039.
26. Willett P. Similarity-based approaches to virtual screening. *Biochem Soc Trans.* 2003; 31(Pt 3): 603–606. [PubMed: 12773164]
27. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUDE): better ligands and decoys for better benchmarking. *J Med Chem.* 2012; 55(14):6582–6594. [PubMed: 22716043]
28. Nicholls A. What do we know and when do we know it? *J Comput Aided Mol Des.* 2008; 22(3–4): 239–255. [PubMed: 18253702]
29. Lang PT, Brozell SR, Mukherjee S, Pettersen EF, Meng EC, Thomas V, Rizzo RC, Case DA, James TL, Kuntz ID. DOCK 6: combining techniques to model RNA-small molecule complexes. *RNA.* 2009; 15(6):1219–1230. [PubMed: 19369428]
30. Brozell SR, Mukherjee S, Balius TE, Roe DR, Case DA, Rizzo RC. Evaluation of DOCK 6 as a pose generation and database enrichment tool. *J Comput Aided Mol Des.* 2012; 26(6):749–773. [PubMed: 22569593]
31. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009; 30(16):2785–2791. [PubMed: 19399780]
32. Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N, Garmendia-Doval AB, Juhos S, Schmidtke P, Barril X, Hubbard RE, Morley SD. rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol.* 2014; 10(4):e1003571. [PubMed: 24722481]
33. Korb O, Stutzle T, Exner TE. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J Chem Inf Model.* 2009; 49(1):84–96. [PubMed: 19125657]
34. Rush TS 3rd, Grant JA, Mosyak L, Nicholls A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J Med Chem.* 2005; 48(5):1489–1495. [PubMed: 15743191]
35. Hawkins PC, Skillman AG, Nicholls A. Comparison of shape-matching and docking as virtual screening tools. *J Med Chem.* 2007; 50(1):74–82. [PubMed: 17201411]
36. ROCS version 3.2.0.3. OpenEye Scientific Software SF, NM. <http://www.eyesopen.com>

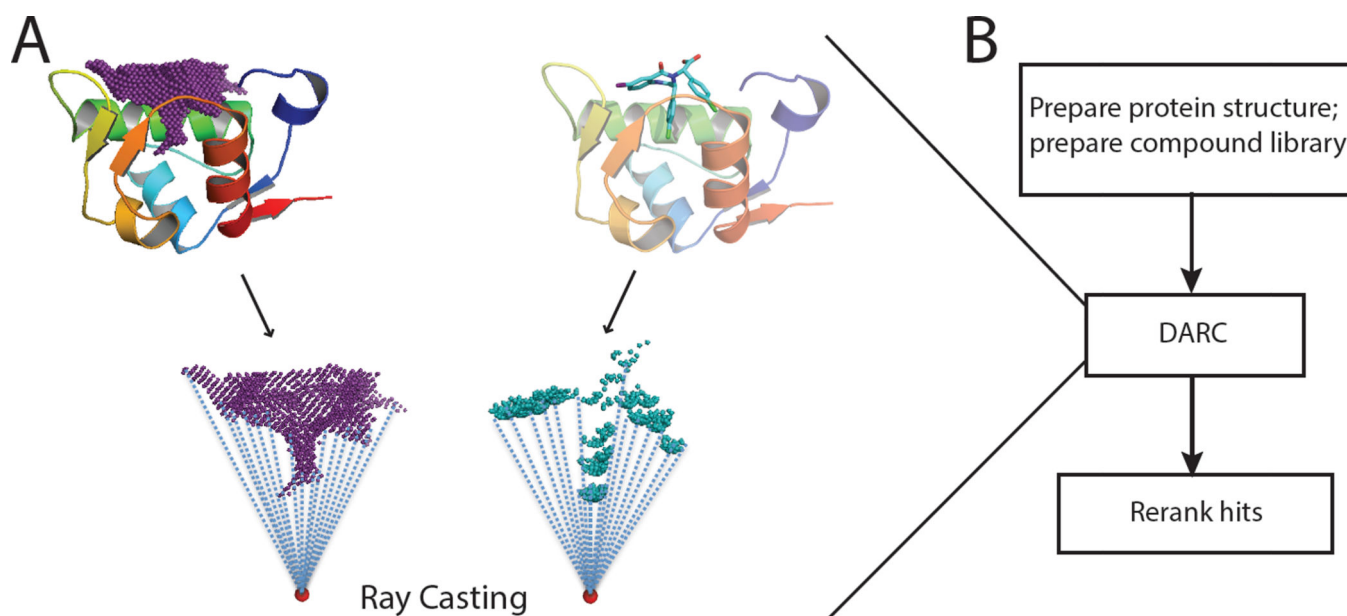
37. McGovern SL, Shoichet BK. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J Med Chem.* 2003; 46(14):2895–2907. [PubMed: 12825931]
38. Claussen H, Buning C, Rarey M, Lengauer T. FlexE: efficient molecular docking considering protein structure variations. *J Mol Biol.* 2001; 308(2):377–395. [PubMed: 11327774]
39. Cavasotto CN, Abagyan RA. Protein flexibility in ligand docking and virtual screening to protein kinases. *J Mol Biol.* 2004; 337(1):209–225. [PubMed: 15001363]
40. Osguthorpe DJ, Sherman W, Hagler AT. Exploring protein flexibility: incorporating structural ensembles from crystal structures and simulation into virtual screening protocols. *J Phys Chem B.* 2012; 116(23):6952–6959. [PubMed: 22424156]
41. Huang SY, Zou X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins.* 2007; 66(2):399–421. [PubMed: 17096427]
42. Cosconati S, Marinelli L, Di Leva FS, La Pietra V, De Simone A, Mancini F, Andrisano V, Novellino E, Goodsell DS, Olson AJ. Protein flexibility in virtual screening: the BACE-1 case study. *J Chem Inf Model.* 2012; 52(10):2697–2704. [PubMed: 23005250]
43. Cavasotto CN, Kovacs JA, Abagyan RA. Representing receptor flexibility in ligand docking through relevant normal modes. *J Am Chem Soc.* 2005; 127(26):9632–9640. [PubMed: 15984891]
44. Dixit A, Verkhivker GM. Integrating ligand-based and protein-centric virtual screening of kinase inhibitors using ensembles of multiple protein kinase genes and conformations. *J Chem Inf Model.* 2012; 52(10):2501–2515. [PubMed: 22992037]
45. Cheng LS, Amaro RE, Xu D, Li WW, Arzberger PW, McCammon JA. Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. *J Med Chem.* 2008; 51(13):3878–3894. [PubMed: 18558668]
46. Shoemaker AR, Oleksijew A, Bauch J, Belli BA, Borre T, Bruncko M, Deckwirth T, Frost DJ, Jarvis K, Joseph MK, Marsh K, McClellan W, Nellans H, Ng S, Nimmer P, O'Connor JM, Oltersdorf T, Qing W, Shen W, Stavropoulos J, Tahir SK, Wang B, Warner R, Zhang H, Fesik SW, Rosenberg SH, Elmore SW. A small-molecule inhibitor of Bcl-XL potentiates the activity of cytotoxic drugs in vitro and in vivo. *Cancer Research.* 2006; 66(17):8731–8739. [PubMed: 16951189]
47. Wang G, Nikolovska-Coleska Z, Yang CY, Wang R, Tang G, Guo J, Shangary S, Qiu S, Gao W, Yang D, Meagher J, Stuckey J, Krajewski K, Jiang S, Roller PP, Abaan HO, 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. [PubMed: 17034116]
48. van Delft MF, Wei AH, Mason KD, Vandenberg CJ, Chen L, Czabotar PE, Willis SN, Scott CL, Day CL, Cory S, Adams JM, Roberts AW, Huang DC. The BH3 mimetic ABT-737 targets selective Bcl-2 proteins and efficiently induces apoptosis via Bak/Bax if Mcl-1 is neutralized. *Cancer Cell.* 2006; 10(5):389–399. [PubMed: 17097561]
49. Konopleva M, Contractor R, Tsao T, Samudio I, Ruvolo PP, Kitada S, Deng X, Zhai D, Shi YX, Sneed T, Verhaegen M, Soengas M, Ruvolo VR, McQueen T, Schober WD, Watt JC, Jiffar T, Ling X, Marini FC, Harris D, Dietrich M, Estrov Z, McCubrey J, May WS, Reed JC, Andreeff M. Mechanisms of apoptosis sensitivity and resistance to the BH3 mimetic ABT-737 in acute myeloid leukemia. *Cancer Cell.* 2006; 10(5):375–388. [PubMed: 17097560]
50. Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, Bruncko M, Deckwerth TL, Dinges J, Hajduk PJ, Joseph MK, Kitada S, Korsmeyer SJ, Kunzer AR, Letai A, Li C, Mitten MJ, Nettesheim DG, Ng S, Nimmer PM, O'Connor JM, Oleksijew A, Petros AM, Reed JC, Shen W, Tahir SK, Thompson CB, Tomaselli KJ, Wang B, Wendt MD, Zhang H, Fesik SW, Rosenberg SH. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature.* 2005; 435(7042):677–681. [PubMed: 15902208]
51. Souers AJ, Levenson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, Dayton BD, Ding H, Enschede SH, Fairbrother WJ, Huang DC, Hymowitz SG, Jin S, Khaw SL, Kovar PJ, Lam LT, Lee J, Maecker HL, Marsh KC, Mason KD, Mitten MJ, Nimmer PM, Oleksijew A, Park CH, Park CM, Phillips DC, Roberts AW, Sampath D, Seymour JF, Smith ML, Sullivan GM, Tahir SK, Tse C, Wendt MD, Xiao Y, Xue JC, Zhang H, Humerickhouse RA, Rosenberg SH, Elmore SW. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med.* 2013; 19(2):202–208. [PubMed: 23291630]

52. Chen S, Dai Y, Harada H, Dent P, Grant S. Mcl-1 down-regulation potentiates ABT-737 lethality by cooperatively inducing Bak activation and Bax translocation. *Cancer Research*. 2007; 67(2): 782–791. [PubMed: 17234790]
53. Lemmen C, Lengauer T, Klebe G. FLEXS: a method for fast flexible ligand superposition. *J Med Chem*. 1998; 41(23):4502–4520. [PubMed: 9804690]
54. Belmar J, Fesik SW. Small molecule Mcl-1 inhibitors for the treatment of cancer. *Pharmacol Ther*. 2015; 145C:76–84. [PubMed: 25172548]
55. Bajwa N, Liao C, Nikolovska-Coleska Z. Inhibitors of the anti-apoptotic Bcl-2 proteins: a patent review. *Expert Opin Ther Pat*. 2012; 22(1):37–55. [PubMed: 22195752]
56. Abulwerdi F, Liao C, Liu M, Azmi AS, Aboukameel A, Mady AS, Gulappa T, Cierpicki T, Owens S, Zhang T, Sun D, Stuckey JA, Mohammad RM, Nikolovska-Coleska Z. A novel small-molecule inhibitor of mcl-1 blocks pancreatic cancer growth in vitro and in vivo. *Mol Cancer Ther*. 2014; 13(3):565–575. [PubMed: 24019208]
57. Bruncko M, Wang L, Sheppard GS, Phillips DC, Tahir SK, Xue J, Erickson S, Fidanze S, Fry E, Hasvold L, Jenkins GJ, Jin S, Judge RA, Kovar PJ, Madar D, Nimmer P, Park C, Petros AM, Rosenberg SH, Smith ML, Song X, Sun C, Tao ZF, Wang X, Xiao Y, Zhang H, Tse C, Levenson JD, Elmore SW, Souers AJ. Structure-guided design of a series of MCL-1 inhibitors with high affinity and selectivity. *J Med Chem*. 2015; 58(5):2180–2194. [PubMed: 25679114]
58. Fang C, D'Souza B, Thompson CF, Clifton MC, Fairman JW, Fulroth B, Leed A, McCarren P, Wang L, Wang Y, Feau C, Kaushik VK, Palmer M, Wei G, Golub TR, Hubbard BK, Serrano-Wu MH. Single Diastereomer of a Macrolactam Core Binds Specifically to Myeloid Cell Leukemia 1 (MCL1). *ACS Med Chem Lett*. 2014; 5(12):1308–1312. [PubMed: 25516789]
59. Friberg A, Vigil D, Zhao B, Daniels RN, Burke JP, Garcia-Barrantes PM, Camper D, Chauder BA, Lee T, Olejniczak ET, Fesik SW. Discovery of potent myeloid cell leukemia 1 (Mcl-1) inhibitors using fragment-based methods and structure-based design. *J Med Chem*. 2013; 56(1):15–30. [PubMed: 23244564]
60. Petros AM, Swann SL, Song D, Swinger K, Park C, Zhang H, Wendt MD, Kunzer AR, Souers AJ, Sun C. Fragment-based discovery of potent inhibitors of the anti-apoptotic MCL-1 protein. *Bioorg Med Chem Lett*. 2014; 24(6):1484–1488. [PubMed: 24582986]
61. Rega MF, Wu B, Wei J, Zhang Z, Cellitti JF, Pellecchia M. SAR by interligand nuclear overhauser effects (ILOEs) based discovery of acylsulfonamide compounds active against Bcl-x(L) and Mcl-1. *J Med Chem*. 2011; 54(17):6000–6013. [PubMed: 21797225]
62. Prakesch M, Denisov AY, Naim M, Gehring K, Arya P. The discovery of small molecule chemical probes of Bcl-X(L) and Mcl-1. *Bioorg Med Chem*. 2008; 16(15):7443–7449. [PubMed: 18603434]
63. Ding X, Li Y, Lv L, Zhou M, Han L, Zhang Z, Ba Q, Li J, Wang H, Liu H, Wang R. De novo design, synthesis and evaluation of benzylpiperazine derivatives as highly selective binders of Mcl-1. *ChemMedChem*. 2013; 8(12):1986–2014. [PubMed: 24124106]
64. Cao X, Yap JL, Newell-Rogers MK, Peddaboina C, Jiang W, Papaconstantinou HT, Jupiter D, Rai A, Jung KY, Tubin RP, Yu W, Vanommeslaeghe K, Wilder PT, MacKerell AD Jr, Fletcher S, Smythe RW. The novel BH3 alpha-helix mimetic JY-1-106 induces apoptosis in a subset of cancer cells (lung cancer, colon cancer and mesothelioma) by disrupting Bcl-xL and Mcl-1 protein-protein interactions with Bak. *Mol Cancer*. 2013; 12(1):42. [PubMed: 23680104]
65. Kazi A, Sun J, Doi K, Sung SS, Takahashi Y, Yin H, Rodriguez JM, Becerril J, Berndt N, Hamilton AD, Wang HG, Sebt SM. The BH3 alpha-helical mimic BH3-M6 disrupts Bcl-X(L), Bcl-2, and MCL-1 protein-protein interactions with Bax, Bak, Bad, or Bim and induces apoptosis in a Bax- and Bim-dependent manner. *J Biol Chem*. 2011; 286(11):9382–9392. [PubMed: 21148306]
66. Stewart ML, Fire E, Keating AE, Walensky LD. The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nat Chem Biol*. 2010; 6(8):595–601. [PubMed: 20562877]
67. Muppidi A, Doi K, Edwardraja S, Drake EJ, Gulick AM, Wang HG, Lin Q. Rational design of proteolytically stable, cell-permeable peptide-based selective Mcl-1 inhibitors. *J Am Chem Soc*. 2012; 134(36):14734–14737. [PubMed: 22920569]
68. Muppidi A, Doi K, Ramil CP, Wang HG, Lin Q. Synthesis of cell-permeable stapled BH3 peptide-based Mcl-1 inhibitors containing simple aryl and vinylaryl cross-linkers. *Tetrahedron*. 2014; 70(42):7740–7745. [PubMed: 25267861]

69. Fire E, Gulla SV, Grant RA, Keating AE. Mcl-1-Bim complexes accommodate surprising point mutations via minor structural changes. *Protein Sci.* 2010; 19(3):507–519. [PubMed: 20066663]
70. Johnson DK, Karanicolas J. Selectivity by small-molecule inhibitors of protein interactions can be driven by protein surface fluctuations. *PLoS Comput Biol.* in press
71. Liu G, Poppe L, Aoki K, Yamane H, Lewis J, Szyperski T. High-quality NMR structure of human anti-apoptotic protein domain Mcl-1(171-327) for cancer drug design. *PLoS One.* 2014; 9(5):e96521. [PubMed: 24789074]
72. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of chemical information and modeling.* 2012
73. Davis IW, Baker D. RosettaLigand docking with full ligand and receptor flexibility. *J Mol Biol.* 2009; 385(2):381–392. [PubMed: 19041878]
74. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem.* 2010; 53(7): 2719–2740. [PubMed: 20131845]
75. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol.* 2007; 25(2):197–206. [PubMed: 17287757]
76. Wolter KG, Verhaegen M, Fernandez Y, Nikolovska-Coleska Z, Riblett M, de la Vega CM, Wang S, Soengas MS. Therapeutic window for melanoma treatment provided by selective effects of the proteasome on Bcl-2 proteins. *Cell Death Differ.* 2007; 14(9):1605–1616. [PubMed: 17541428]
77. Tanaka Y, Aikawa K, Nishida G, Homma M, Sogabe S, Igaki S, Hayano Y, Sameshima T, Miyahisa I, Kawamoto T, Tawada M, Imai Y, Inazuka M, Cho N, Imaeda Y, Ishikawa T. Discovery of potent Mcl-1/Bcl-xL dual inhibitors by using a hybridization strategy based on structural analysis of target proteins. *J Med Chem.* 2013; 56(23):9635–9645. [PubMed: 24215352]
78. Srinivasan B, Johnson TE, Lad R, Xing C. Structure-activity relationship studies of chalcone leading to 3-hydroxy-4,3',4',5'-tetramethoxychalcone and its analogues as potent nuclear factor kappaB inhibitors and their anticancer activities. *J Med Chem.* 2009; 52(22):7228–7235. [PubMed: 19883086]
79. Battenberg OA, Yang Y, Verhelst SH, Sieber SA. Target profiling of 4-hydroxyderricin in *S. aureus* reveals seryl-tRNA synthetase binding and inhibition by covalent modification. *Mol Biosyst.* 2013; 9(3):343–351. [PubMed: 23295910]
80. Shoichet BK. Interpreting steep dose-response curves in early inhibitor discovery. *J Med Chem.* 2006; 49(25):7274–7277. [PubMed: 17149857]
81. Ryan AJ, Gray NM, Lowe PN, Chung CW. Effect of detergent on “promiscuous” inhibitors. *J Med Chem.* 2003; 46(16):3448–3451. [PubMed: 12877581]
82. Giannetti AM, Koch BD, Browner MF. Surface plasmon resonance based assay for the detection and characterization of promiscuous inhibitors. *J Med Chem.* 2008; 51(3):574–580. [PubMed: 18181566]
83. Coan KE, Maltby DA, Burlingame AL, Shoichet BK. Promiscuous aggregate-based inhibitors promote enzyme unfolding. *J Med Chem.* 2009; 52(7):2067–2075. [PubMed: 19281222]
84. Amslinger S, Al-Rifai N, Winter K, Wormann K, Scholz R, Baumeister P, Wild M. Reactivity assessment of chalcones by a kinetic thiol assay. *Org Biomol Chem.* 2013; 11(4):549–554. [PubMed: 23224077]
85. DeSimone RW, Currie KS, Mitchell SA, Darrow JW, Pippin DA. Privileged structures: applications in drug discovery. *Comb Chem High Throughput Screen.* 2004; 7(5):473–494. [PubMed: 15320713]
86. Welsch ME, Snyder SA, Stockwell BR. Privileged scaffolds for library design and drug discovery. *Curr Opin Chem Biol.* 2010; 14(3):347–361. [PubMed: 20303320]
87. Varadarajan S, Vogler M, Butterworth M, Dinsdale D, Walensky LD, Cohen GM. Evaluation and critical assessment of putative MCL-1 inhibitors. *Cell Death Differ.* 2013; 20(11):1475–1484. [PubMed: 23832116]
88. Eichhorn JM, Alford SE, Hughes CC, Fenical W, Chambers TC. Purported Mcl-1 inhibitor marinopyrrole A fails to show selective cytotoxicity for Mcl-1-dependent cell lines. *Cell Death Dis.* 2013; 4:e880. [PubMed: 24157874]

89. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci U S A*. 1992; 89(6):2195–2199. [PubMed: 1549581]
90. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol*. 1997; 272(1):106–120. [PubMed: 9299341]
91. Morris RJ, Najmanovich RJ, Kahraman A, Thornton JM. Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons. *Bioinformatics*. 2005; 21(10):2347–2355. [PubMed: 15728116]
92. Cai W, Shao X, Maigret B. Protein-ligand recognition using spherical harmonic molecular surfaces: towards a fast and efficient filter for large virtual throughput screening. *J Mol Graph Model*. 2002; 20(4):313–328. [PubMed: 11858640]
93. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*. 2005; 33(Web Server issue):W363–W367. [PubMed: 15980490]
94. Lee HS, Lee CS, Kim JS, Kim DH, Choe H. Improving virtual screening performance against conformational variations of receptors by shape matching with ligand binding pocket. *J Chem Inf Model*. 2009; 49(11):2419–2428. [PubMed: 19852439]
95. Ebalunode JO, Ouyang Z, Liang J, Zheng W. Novel approach to structure-based pharmacophore search using computational geometry and shape matching techniques. *J Chem Inf Model*. 2008; 48(4):889–901. [PubMed: 18396858]
96. Hoffmann B, Zaslavskiy M, Vert JP, Stoven V. A new protein binding pocket similarity measure based on comparison of clouds of atoms in 3D: application to ligand prediction. *BMC Bioinformatics*. 2010; 11:99. [PubMed: 20175916]
97. Li B, Turuvekere S, Agrawal M, La D, Ramani K, Kihara D. Characterization of local geometry of protein surfaces with the visibility criterion. *Proteins*. 2008; 71(2):670–683. [PubMed: 17975834]
98. Chikhi R, Sael L, Kihara D. Real-time ligand binding pocket database search using local surface descriptors. *Proteins*. 2010; 78(9):2007–2028. [PubMed: 20455259]
99. Gowthaman R, Lyskov S, Karanicolas J. DARC 2.0 : Improved docking and virtual screening at protein interaction sites. manuscript in preparation.
100. Fauman EB, Rai BK, Huang ES. Structure-based druggability assessment--identifying suitable targets for small molecule therapeutics. *Curr Opin Chem Biol*. 2011; 15(4):463–468. [PubMed: 21704549]
101. Naylor E, Arredouani A, Vasudevan SR, Lewis AM, Parkesh R, Mizote A, Rosen D, Thomas JM, Izumi M, Ganesan A, Galione A, Churchill GC. Identification of a chemical probe for NAADP by virtual screening. *Nat Chem Biol*. 2009; 5(4):220–226. [PubMed: 19234453]
102. Koes D, Khoury K, Huang Y, Wang W, Bista M, Popowicz GM, Wolf S, Holak TA, Domling A, Camacho CJ. Enabling large-scale design, synthesis and validation of small molecule protein-protein antagonists. *PLoS One*. 2012; 7(3):e32839. [PubMed: 22427896]
103. Koes DR, Camacho CJ. Small-molecule inhibitor starting points learned from protein-protein interaction inhibitor structure. *Bioinformatics*. 2012; 28(6):784–791. [PubMed: 22210869]
104. Metz A, Schanda J, Grez M, Wichmann C, Gohlke H. From determinants of RUNX1/ETO tetramerization to small-molecule protein-protein interaction inhibitors targeting acute myeloid leukemia. *J Chem Inf Model*. 2013; 53(9):2197–2202. [PubMed: 23957251]
105. Tsao DH, Sutherland AG, Jennings LD, Li Y, Rush TS 3rd, Alvarez JC, Ding W, Dushin EG, Dushin RG, Haney SA, Kenny CH, Malakian AK, Nilakantan R, Mosyak L. Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. *Bioorg Med Chem*. 2006; 14(23):7953–7961. [PubMed: 16919463]
106. Petros AM, Huth JR, Oost T, Park CM, Ding H, Wang X, Zhang H, Nimmer P, Mendoza R, Sun C, Mack J, Walter K, Dorwin S, Gramling E, Lador U, Rosenberg SH, Elmore SW, Fesik SW, Hajduk PJ. Discovery of a potent and selective Bcl-2 inhibitor using SAR by NMR. *Bioorg Med Chem Lett*. 2010; 20(22):6587–6591. [PubMed: 20870405]
107. Frank AO, Feldkamp MD, Kennedy JP, Waterson AG, Pelz NF, Patrone JD, Vangamudi B, Camper DV, Rossanese OW, Chazin WJ, Fesik SW. Discovery of a potent inhibitor of replication

- protein a protein-protein interactions using a fragment-linking approach. *J Med Chem.* 2013; 56(22):9242–9250. [PubMed: 24147804]
108. Villemagne B, Flipo M, Blondiaux N, Crauste C, Malaquin S, Leroux F, Piveteau C, Villeret V, Brodin P, Villoutreix BO, Sperandio O, Soror SH, Wohlkonig A, Wintjens R, Deprez B, Baulard AR, Willand N. Ligand efficiency driven design of new inhibitors of Mycobacterium tuberculosis transcriptional repressor EthR using fragment growing, merging, and linking approaches. *J Med Chem.* 2014; 57(11):4876–4888. [PubMed: 24818704]
109. Murray CW, Blundell TL. Structural biology in fragment-based drug design. *Curr Opin Struct Biol.* 2010; 20(4):497–507. [PubMed: 20471246]
110. Mattos C, Bellamacina CR, Peisach E, Pereira A, Vitkup D, Petsko GA, Ringe D. Multiple solvent crystal structures: probing binding sites, plasticity and hydration. *J Mol Biol.* 2006; 357(5):1471–1482. [PubMed: 16488429]
111. Sperandio O, Mouawad L, Pinto E, Villoutreix BO, Perahia D, Miteva MA. How to choose relevant multiple receptor conformations for virtual screening: a test case of Cdk2 and normal mode analysis. *Eur Biophys J.* 2010; 39(9):1365–1372. [PubMed: 20237920]
112. Korb O, Olsson TS, Bowden SJ, Hall RJ, Verdonk ML, Liebeschuetz JW, Cole JC. Potential and limitations of ensemble docking. *J Chem Inf Model.* 2012; 52(5):1262–1274. [PubMed: 22482774]
113. Sorensen J, Demir O, Swift RV, Feher VA, Amaro RE. Molecular docking to flexible targets. *Methods Mol Biol.* 2015; 1215:445–469. [PubMed: 25330975]
114. Hari SB, Perera BG, Ranjitkar P, Seeliger MA, Maly DJ. Conformation-selective inhibitors reveal differences in the activation and phosphate-binding loops of the tyrosine kinases Abl and Src. *ACS Chem Biol.* 2013; 8(12):2734–2743. [PubMed: 24106839]
115. Chaikuad A, Tacconi EM, Zimmer J, Liang Y, Gray NS, Tarsounas M, Knapp S. A unique inhibitor binding site in ERK1/2 is associated with slow binding kinetics. *Nat Chem Biol.* 2014; 10(10):853–860. [PubMed: 25195011]
116. Gray NS, Fabbro D. Discovery of allosteric BCR-ABL inhibitors from phenotypic screen to clinical candidate. *Methods Enzymol.* 2014; 548:173–188. [PubMed: 25399646]
117. Ung PM, Dunbar JB Jr, Gestwicki JE, Carlson HA. An allosteric modulator of HIV-1 protease shows equipotent inhibition of wild-type and drug-resistant proteases. *J Med Chem.* 2014; 57(15):6468–6478. [PubMed: 25062388]

**Figure 1. Docking Approach using Ray-Casting**

(A) DARC first casts a set of rays emanating from an origin within the protein (*red dot*), and maps the topography of the surface pocket by monitoring intersection of these rays with the pocket (*left*). To evaluate the shape complementarity of a given ligand for this pocket, DARC casts the same rays (from the same origin) and monitors their intersection with the ligand (*right*). If ligand (in its current position and orientation) is perfectly shape-complementary to the pocket, each ray will intersect the ligand at the same distance from the origin as it intersected the protein surface pocket. By moving the ligand to maximize this shape complementarity, a ligand can be docked into a protein surface pocket. (B) A schematic diagram of the complete workflow split into three stages: pre-DARC preparation, DARC, and post-DARC re-ranking/filtering of the docked models.

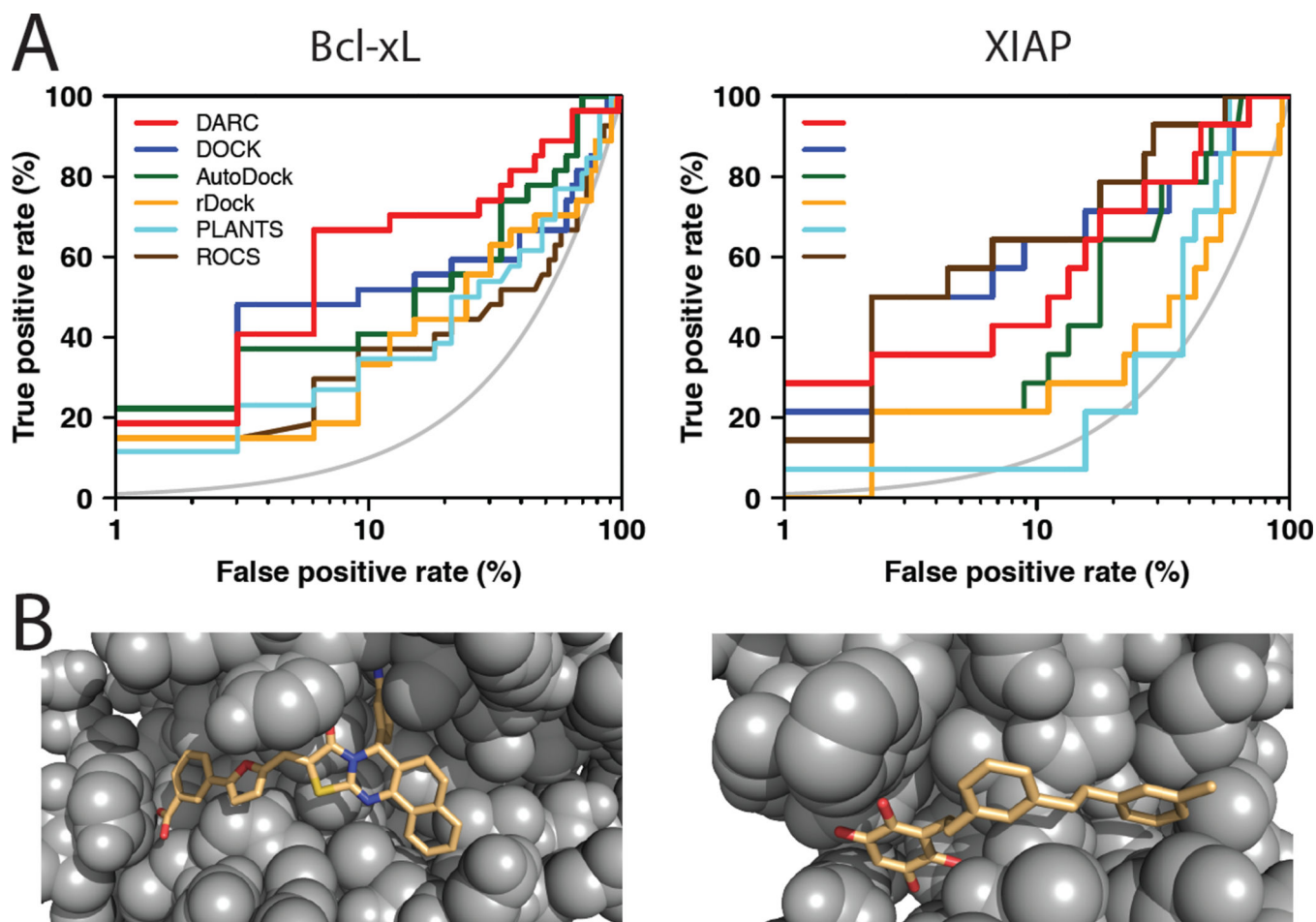


Figure 2. Virtual screening benchmark experiment using an inhibitor-bound protein structure (A) This receiver operating characteristic (ROC) plot compares the performance of various methods (DARC, DOCK [29,30], AutoDock [31], rDock [32], PLANTS [33], and ROCS [34–36]) for predicting whether a given compound is active against a particular protein interaction site. Drug-like “decoy” compounds were drawn from the Astex diverse set [23], then further filtered to remove any compounds that are similar in chemical structure to an active compound or any other decoy compound. The results are presented on a semi-log plot to highlight the “early” performance of the methods; the grey curve indicates the random retrieval of compounds (i.e. a random predictor). *Left*: discriminating 27 diverse compounds active against Bcl-xL from among 33 matched decoy compounds. *Right*: discriminating 14 diverse compounds active against XIAP from among 45 matched decoy compounds. (B) DARC-docked models of representative active compounds against Bcl-xL (*left*) and XIAP (*right*). These compounds were not chemically similar to those used in solving these crystal structures (such compounds were excluded from this benchmark).

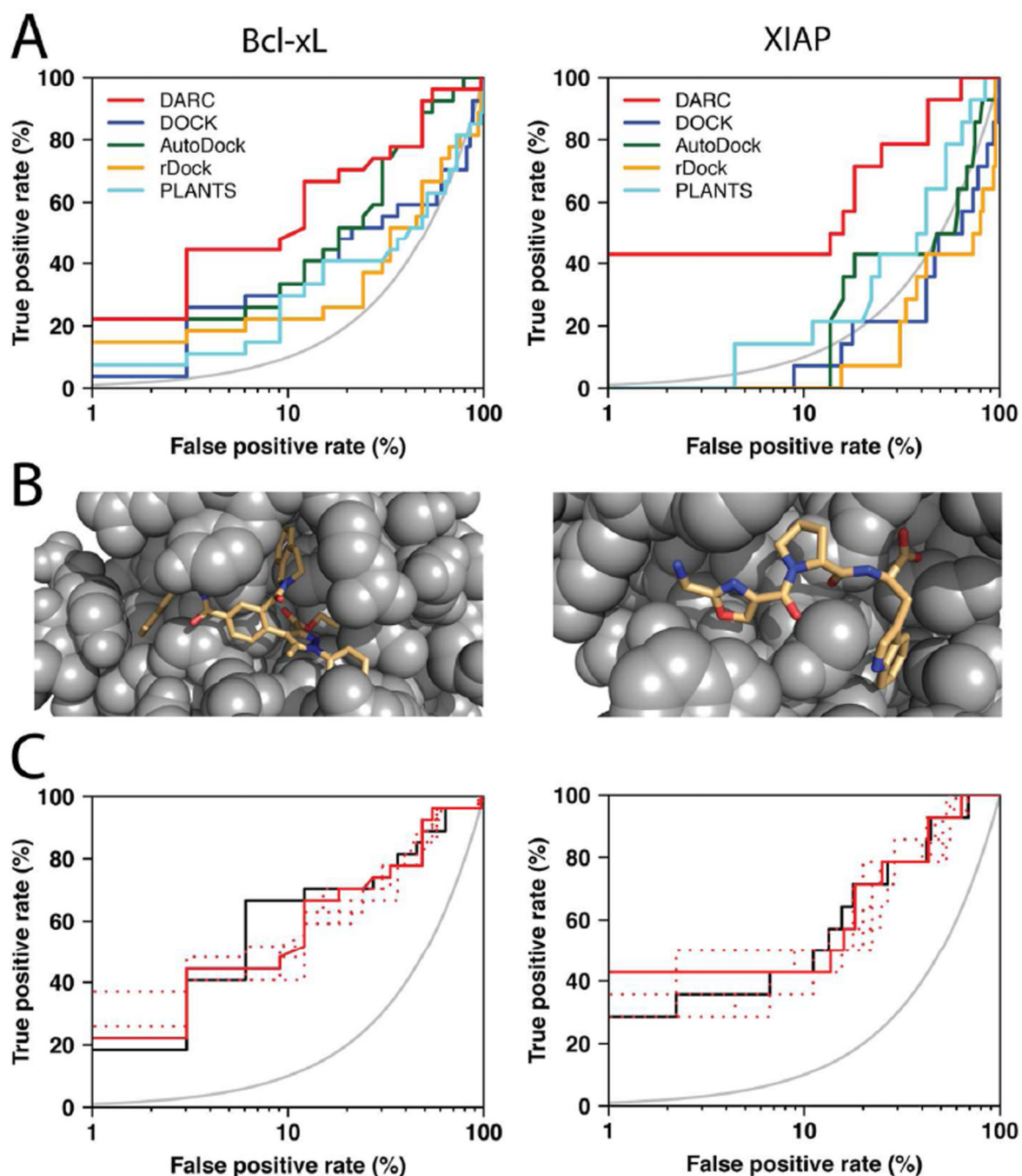


Figure 3. Virtual screening benchmark experiment using protein structures from “pocket optimization” simulations

(A) This benchmark experiment involves discrimination of the same active versus “decoy” compounds as in Figure 2, however this time compounds were screened against a protein conformation generated via “pocket optimization” simulations initiated from an unbound crystal structure (instead of protein conformations from an inhibitor-bound crystal structure) [16]. The results are again presented on a semi-log plot to highlight the “early” performance of the methods; the grey curve indicates the random retrieval of compounds (i.e. a random

predictor). **(B)** Models of representative active compounds docked by DARC to the “pocket optimized” conformation of Bcl-xL (*left*) or XIAP (*right*). **(C)** The performance of DARC in this benchmark is essentially equivalent when screening against a known ligand-bound structure (*solid black line*), the lowest-energy “pocket optimized” conformations (i.e. as described in the previous panels) (*solid red line*), or any of five other low-energy “pocket optimized” conformations (*dashed red lines*); this observation demonstrates the insensitivity of the method to details of the protein structure.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

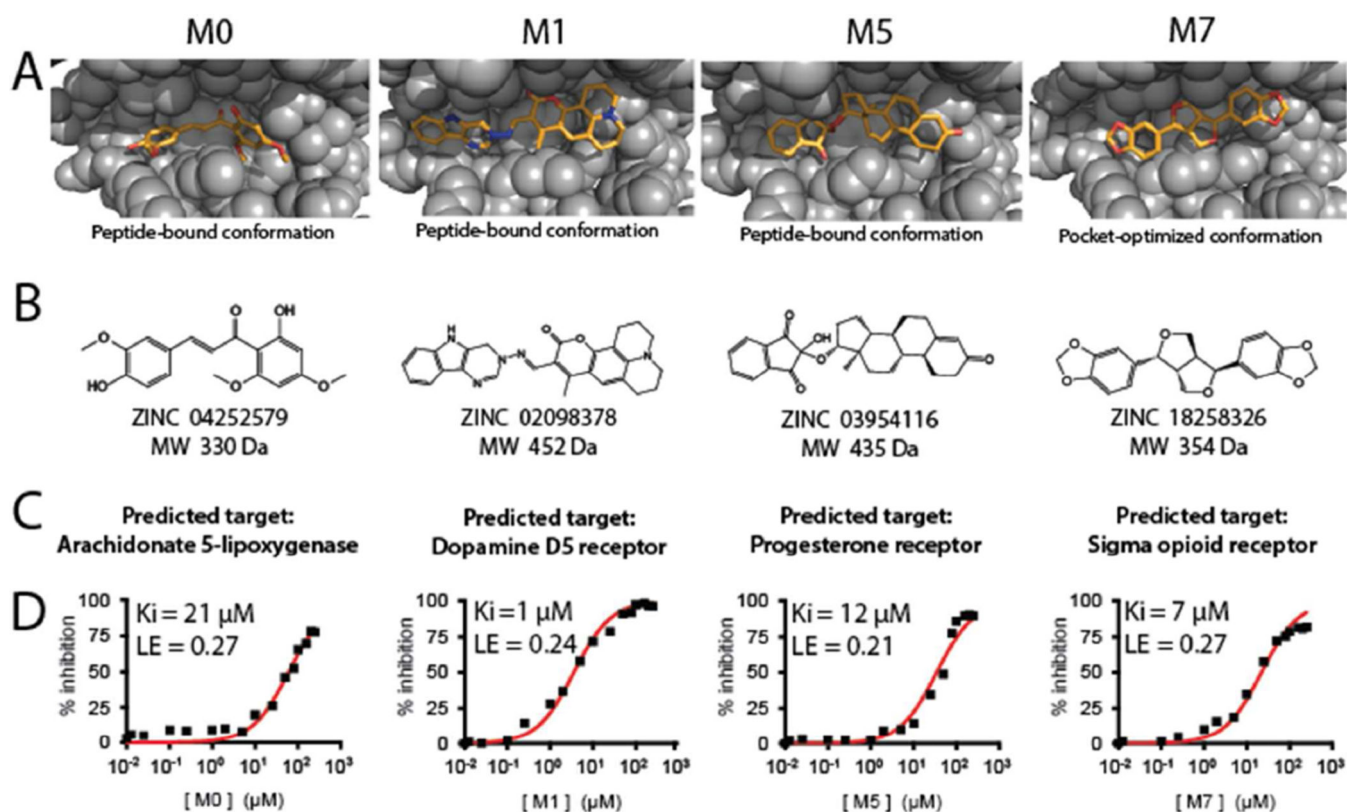


Figure 4. DARC can be used to identify novel inhibitors of human Mcl-1

Screening was carried out against the protein interaction site of Mcl-1, and 21 of the top-scoring compounds were tested in a biochemical assay. (A) Docked models generated by DARC for four compounds emerging from our screen. Compounds were screened either against an experimentally-derived peptide-bound structure of Mcl-1, or against an Mcl-1 conformation generated via “pocket optimization” simulations [16]. (B) The chemical structures of the compounds in each model are shown. We note that some of these have reactive functional groups: such compounds could be removed from the library prior to screening, in applications where such moieties are undesirable. (C) The most likely activity of each compound was predicted using the Similarity Ensemble Approach (SEA) [75]. None of these compounds are assigned as likely inhibitors of Bcl-2-family proteins, underscoring their lack of similarity to known inhibitors of this family. (D) Each of these four compounds inhibit Mcl-1’s interaction with a FITC-labeled cognate peptide, as determined via fluorescence polarization. CHAPS detergent (0.1%) was included in this assay, to ensure that these results were not due to compound aggregation; we also observe equivalent inhibition in the absence of CHAPS. Each curve is fit to a one-site inhibition model, with a single free parameter in the fitting. Direct binding of each of these compounds to Mcl-1 was subsequently confirmed via bio-layer interferometry (see Supplementary Results).

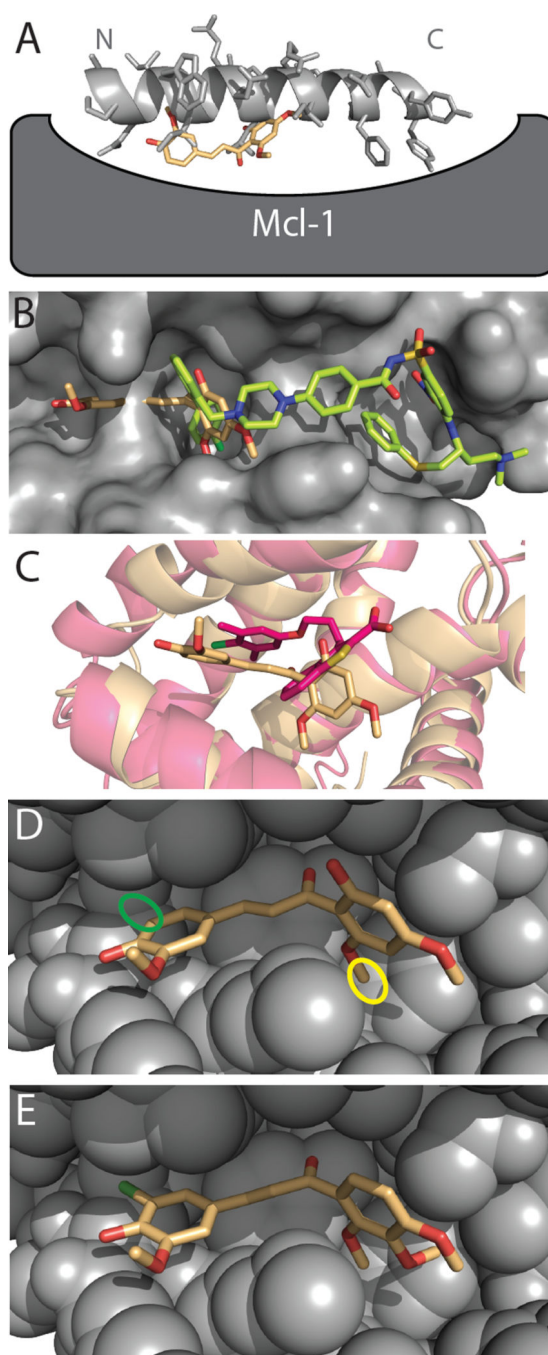


Figure 5. Evaluating DARC's model of the M0/Mcl-1 complex

(A) Though M0 was identified in a screen against a peptide-bound conformation of Mcl-1, M0 (*peach*) is distinctly *not* mimicking the interactions of the cognate peptide (*light gray*). In the DARC-generated model M0 competes with the N-terminus of the peptide for Mcl-1 binding, but fits more deeply into the Mcl-1 surface pocket than the peptide sidechains. (B) This M0 binding mode is distinctly different from that of known Bcl-xL inhibitors such as ABT-737 (*green*), which instead overlap with the C-terminus of the peptide-binding site. The surface of the Bcl-xL protein is shown (*grey*), solved in complex with ABT-737. (C)

Other recently-described inhibitors of Mcl-1, including the representative example shown here (*dark pink*) [59], bind at a similar location to M0; however, each distinct series of inhibitors induce their own unique Mcl-1 conformational change to allow deeper ligand burial. **(D)** The locations of two M0 positions that play key roles in our model of the complex: the 6'-methoxy group (*yellow*) and the unsubstituted 5-position (*green*). **(E)** A model of the D1, the most potent M0 derivative, in complex with Mcl-1. This compound preserves the crucial 6'-methoxy group, and embellishes the 5-position with a chlorine substitution to improve packing. Closely analogous compounds to D1 that lack the methoxy group at the 6'-position (D3) or that lack the chlorine at the 5-position (D5) exhibit reduced activity.

Table 1
DARC screening hits confirmed experimentally as inhibitors of Mcl-1

The complete set of DARC hits (including those that did not inhibit Mcl-1) are included as Table S2. K_i values were determined via a fluorescence polarization competition assay (see Supplementary Methods), and then bio-layer interferometry was used to directly confirm binding to Mcl-1 (see Supplementary Results).

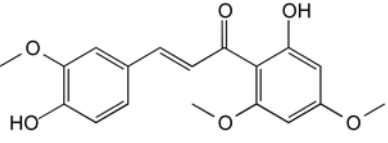
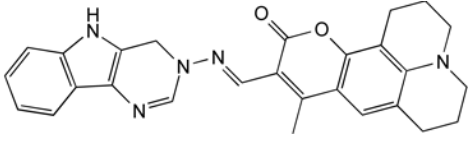
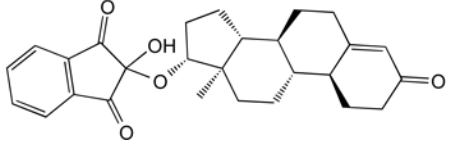
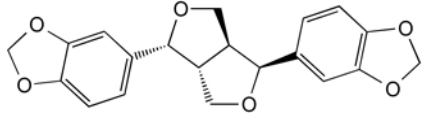
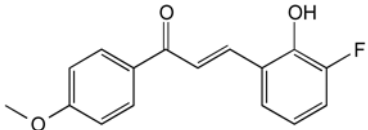
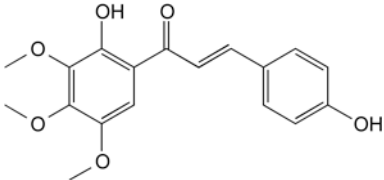
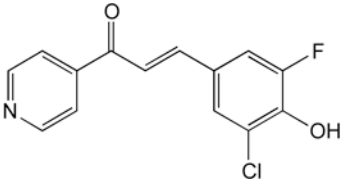
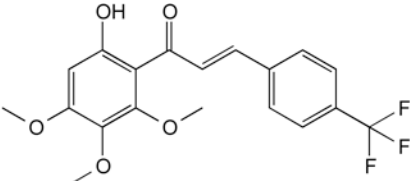
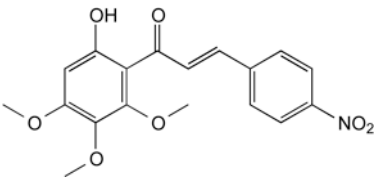
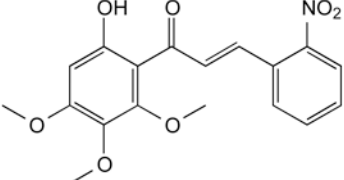
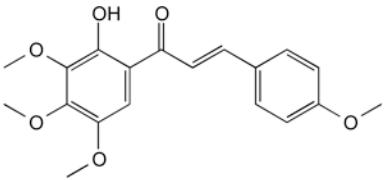
Compound	Chemical structure	K_i (μM)	L.E. (kcal/mol-heavyatom)
M0		21	0.27
M1		1.2	0.24
M5		12	0.21
M7		7	0.27

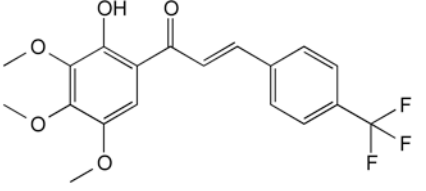
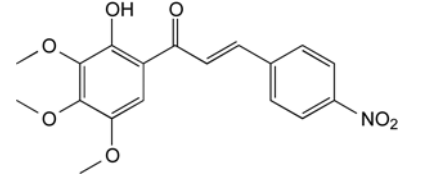
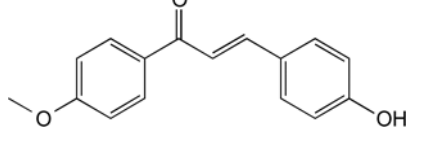
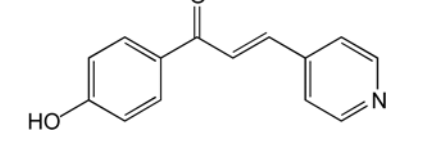
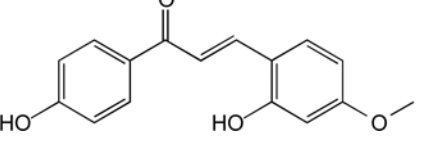
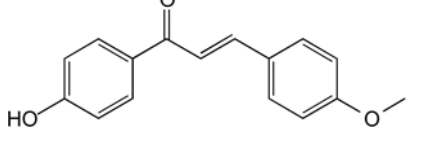
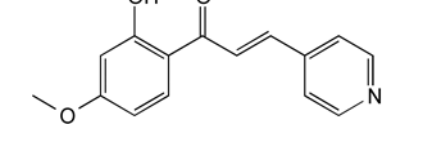
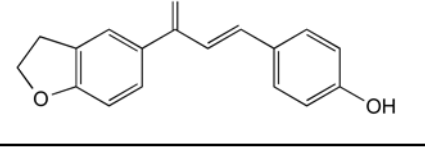
Table 2

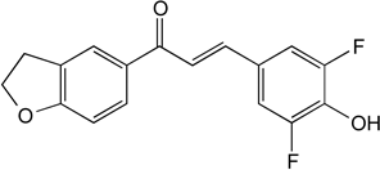
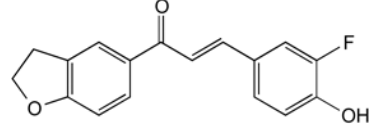
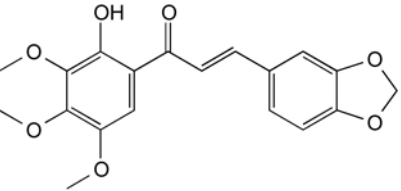
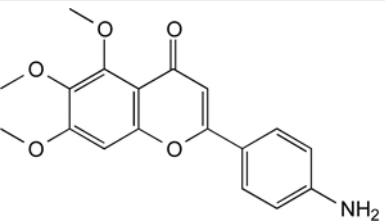
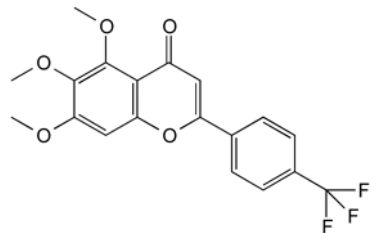
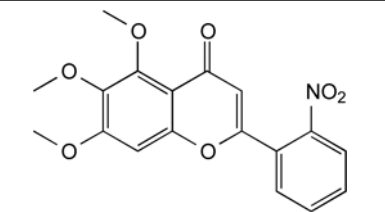
Chemical analogs of M0

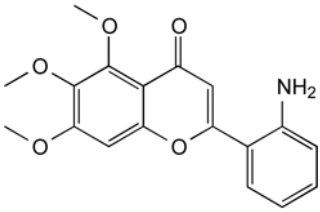
We designed and synthesized a series of 27 M0 derivatives, and tested each of these for inhibition of human Mcl-1 using a fluorescence polarization competition assay (see Supplementary Methods). This series of analogs was designed without consideration of DARC's model of the M0/Mcl-1 complex, to allow unbiased evaluation of the model.

Compound	Chemical structure	Ki (μM)	L.E. (kcal/mol \cdot heavyatom)
M0		21	0.27
D1		2	0.29
D2		3	0.30
D3		7	0.33
D4		10	0.31
D5		28	0.33

Compound	Chemical structure	Ki (μM)	L.E. (kcal/mol·heavyatom)
D6		20	0.32
D7		40	0.25
D8		>50	N.D.
D9		>50	N.D.
D10		>50	N.D.
D11		>50	N.D.
D12		>50	N.D.

Compound	Chemical structure	Ki (μM)	L.E. (kcal/mol-heavyatom)
D13		>50	N.D.
D14		>50	N.D.
D15		>50	N.D.
D16		>50	N.D.
D17		>50	N.D.
D18		>50	N.D.
D19		>50	N.D.
D20		>50	N.D.

Compound	Chemical structure	Ki (μM)	L.E. (kcal/mol·heavyatom)
D21		>50	N.D.
D22		>50	N.D.
D23		>50	N.D.
D24		>50	N.D.
D25		>50	N.D.
D26		>50	N.D.

Compound	Chemical structure	Ki (μM)	L.E. (kcal/mol \cdot heavyatom)
D27	 <chem>COC1=C(OC)C(=O)C2=C(C=C1)OC(C2)C3=CC=C(N)C=C3</chem>	>50	N.D.