In-silico and in-vitro elucidation of BH3 binding specificity towards Bcl-2

Citation: London, Nir, Stefano Gulla, Amy E. Keating, and Ora Schueler-Furman. "In Silico and in Vitro Elucidation of BH3 Binding Specificity toward Bcl-2." Biochemistry 51, no. 29 (July 24, 2012): 5841-5850.

As Published: http://dx.doi.org/10.1021/bi3003567

Publisher: American Chemical Society (ACS)

Persistent URL: http://hdl.handle.net/1721.1/84624

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Published in final edited form as:
Biochemistry. 2012 July 24; 51(29): 5841–5850. doi:10.1021/bi3003567.

In-silico and *in-vitro* elucidation of BH3 binding specificity towards Bcl-2

Nir London[a], **Stefano Gullá**[b], **Amy E. Keating**[b], and **Ora Schueler-Furman**[a,*]

[a]Department of Microbiology and Molecular Genetics, Institute for Medical Research Israel-Canada, Hadassah Medical School, The Hebrew University, POB 12272, Jerusalem, 91120 Israel

[b]Department of Biology, Massachusetts Institute of Technology, Cambridge, 02139 USA

Abstract

Interactions between Bcl-2 like proteins and BH3 domains play a key role in the regulation of apoptosis. Despite the overall structural similarity of their interaction with helical BH3 domains, Bcl-2 like proteins exhibit an intricate spectrum of binding specificities whose underlying basis is not well understood. Here, we characterize these interactions using Rosetta FlexPepBind, a protocol for the prediction of peptide binding specificity that evaluates the binding potential of different peptides based on structural models of the corresponding peptide-receptor complexes. For two prominent players, Bcl-xL and Mcl-1, we obtain good agreement with a large set of experimental SPOT array measurements and recapitulate the binding specificity of peptides derived by yeast display in a previous study. We extend our approach to a third member of this family, Bcl-2: we test our blind prediction of the binding of 180 BIM-derived peptides with a corresponding experimental SPOT array. Both prediction and experiment reveal a Bcl-2 binding specificity pattern that resembles that of Bcl-xL. Finally, we extend this application to accurately predict the specificity pattern of additional human BH3-only derived peptides. This study characterizes the distinct patterns of binding specificity of BH3-only derived peptides for the Bcl-2 like proteins Bcl-xL, Mcl-1 and Bcl-2, and provides insight into the structural basis of determinants of specificity.

Keywords

FlexPepBind; peptide-protein interactions; interaction specificity; Bcl-2 proteins; BH3 peptides; SPOT array

The Bcl-2 protein family is a key regulator of programmed cell death, tumorigenesis and cellular responses to anti-cancer therapy. Proteins of this family can be grouped into two classes: anti-apoptotic proteins that suppress cell death (such as Bcl-2, Bcl-xL, Bcl-w, Mcl-1) and pro-apoptotic proteins that promote cell death. The latter can be further classified based on their conserved domain structure into proteins containing multiple Bcl-2

*Corresponding Author: Ora Schueler-Furman, oraf@ekmd.huji.ac.il, Department of Microbiology and Molecular Genetics, Institute for Medical Research Israel-Canada, Hadassah Medical School, The Hebrew University, POB 12272, Jerusalem 91120 Israel, Tel: 972-2-675-7094, Fax: 972-2-675-7308.

Supporting Information Available

The supporting information includes one Table and two Figures. Table S1 shows the calibration of the protocol: correlations between predicted and experimental BIM-derived peptides to Bcl-xL and Mcl-1 are indicated for different sets of parameters tested. Figure S1 shows the performance of the protocol as a function of the threshold used to define binding peptide sequences in the SPOT arrays, and Figure S2 highlights the details of positions R3b and F4a on the structure of the Mcl1-Bim complex. This material is available free of charge via the Internet at http://pubs.acs.org.

homology (BH) domains (such as Bak, Bax, Bok) and proteins that contain only the BH3 domain (Bad, Bik, Bid, Bim, Puma) [1, 2].

BH3-only proteins promote apoptosis both by the activation of Bak and Bax, as well as by antagonizing pro-survival Bcl-2 family proteins [3–5]. The interaction of the amphipathic helical BH3 domain of a pro-apoptotic protein with the hydrophobic groove created by the conserved BH1, BH2 and BH3 domains of an anti-apoptotic family protein is structurally conserved and well characterized [6–9].

Over-expression of anti-apoptotic Bcl-2 family members has been associated with cancers as well as with chemotherapy resistance, and molecules targeting these interactions are consequently pursued for the development of pharmaceuticals [10–12]. As an example, ABT-737 is a selective and potent inhibitor of Bcl-2, Bcl-w and Bcl-xL [13], while gx15-070 is a pan-specific Bcl-2 family inhibitor [14]. BH3 binding specificity to different Bcl-2 family members has also been the subject of intensive biophysical and cellular biology research. Structural studies have shed light on the BH3 binding determinants [6, 7, 9, 15, 16] while other studies scanned point mutants [17–22] and random peptide libraries [21–23]. Despite all these efforts, the binding specificity of Bcl-2 family members has still not yet been fully elucidated.

Computational attempts to map peptide-binding specificity are usually restricted to one specific type of peptide-binding domain, and can be roughly divided into sequence-based (e.g.[24–27]) and structure-based approaches. The latter may use a fixed backbone (e.g.[28, 29]), or incorporate backbone flexibility (at least on the peptide side; e.g. [30–35]). Incorporation of backbone conformational flexibility has generally improved computer-aided design of functional protein interactions, as well as structure-based prediction of peptide-protein and protein-protein interaction specificity [36, 37]. Several recent studies have advanced the computational structural characterization [38–42] and high-resolution modeling [43–48] of peptide-protein interactions. Consequently, these new tools now provide a platform for the prediction of peptide binding specificity.

What determines the ability of peptides to bind to proteins? Compared to regular protein-protein interactions, peptide-mediated interactions are challenged by the small interface that provides less binding enthalpy, as well as by the fact that the peptide loses a considerable amount of entropy upon binding, since in general it lacks a defined structure in its free state. Several strategies are used to address those challenges [38]. Overall entropy loss upon binding is reduced thanks to small conformational changes of the receptor that adopts the bound conformation prior to interacting with the peptide partner. On the peptide side, a considerable fraction of peptide-mediated interactions involve helical peptides that at least partially adopt the helical conformation in the free state. It therefore does not come as a surprise that a predominant part of successful designs of inhibitory peptides involve helical peptides (e.g. [49, 50]). The interaction between Bcl-2 like proteins and BH3 domains is an example of a helix-receptor interaction that represents an ideal starting point to characterize high-affinity peptide-protein interactions. Thus, this investigation can lead to helical inhibitory peptides that compete at protein interfaces [51, 52] (and can be further optimized, e.g. with staples [53, 54]).

A recent review has described the various structural and sequence features that influence and determine binding specificity [55]. A wide variety of different protein-protein interactions have been observed. Of particular importance to the complexity of the interaction network in a cell are interaction types that have been duplicated and are multiply reused. A range of different specificities can be observed for homologous protein pairs: certain proteins will bind only to one specific protein partner, while others show less specificity and bind

promiscuously to a range of different homologs. The wide variability and conservation observed in different systems has revealed both robustness and sensitivity: binding can be obtained in different ways for the same interface, while single mutations can strongly affect binding. These differences will have profound functional consequences on the cell. The interaction between Bcl-2 like proteins and BH3 derived peptides represents a model case to elucidate the structural basis of such intricate specificity effects.

In this study we have applied our previously developed FlexPepBind protocol that predicts peptide-binding specificity (e.g. [35]) to recapitulate the BH3 peptide binding preferences to Bcl-2 protein family members. After an initial calibration and application of FlexPepBind to Mcl-1 and Bcl-xL based on *in vitro* SPOT array experiments [56] for a set of BIM-derived peptides [21], we used this protocol to blindly predict the corresponding specificity profile of Bcl-2. Our prediction was shown to agree well with an additional, independently performed *in vitro* SPOT array experiment: Both prediction and experiment indicate that Bcl-2 displays a BH3 binding profile that is highly similar to the profile previously reported for Bcl-xL. Finally, we used this approach to recapitulate the distinct patterns of binding specificity of a set of BH3-like peptides to Bcl-xL and Mcl-1, and the binding of several native human BH3 peptides to all three receptors. The successful performance of our protocol on several receptors of this system bodes well for future application and suggests that, after proper calibration, it is general enough for a large variety of peptide-protein interactions.

Materials and Methods

Datasets

All datasets of peptide binding assays Bcl-xL and Mcl-1 were taken from Dutta *et al*. [21] and the numbering of the residues is according to the definition therein (see also Fig. 1): *SPOT arrays*: (a) TRAIN1 – a SPOT array displaying all possible combinations of selected Bim-BH3 mutants at the following positions: 2d: Ile/Ala/Phe; 3a: Leu/Ile/Phe/Ala; 3b: Arg/Asp; 3d: Ile/Phe/Asp/Asn/Ala; 4a: Phe/Val/Asn (n=360 sequences); (b) TRAIN2 – a SPOT array displaying Bim-BH3 single mutant variants, in which 10 interface positions of the peptide (2d,2e,2g,3a,3b,3d,3e,3f,3g and 4a) were substituted to all amino acids (excluding Met and Cys; the native Bim-BH3 sequence was probed 20 times but only used once for calculation of the correlation) (n=181). This set was probed by either 1μM or 100nM of the receptor. (2) *Yeast surface display experiments:* (a) YEAST-X – Sequences from a yeast surface display experiment that were shown to bind to Bcl-xL in preference to Mcl-1 (n=40). (b) YEAST-M – Sequences from a yeast surface display experiment that were shown to bind to Mcl-1 in preference to Bcl-xL (n=33). (c) YEAST-BOTH - Sequences from a yeast surface display experiment that were shown to bind both Bcl-xL and Mcl-1 (n=17).

Modeling the peptide-receptor complex

General modeling scheme—Our scheme is based on the previously reported peptide-docking protocol Rosetta FlexPepDock [43]. For a given peptide, we thread the desired sequence onto the peptide in the native structure, while keeping the side chains of the receptor fixed. We then use Rosetta FlexPepDock in full-atom mode (i.e. without centroid pre-optimization) to refine the structure of the complex with the threaded target peptide (All of the peptide's side-chains as the well at the receptor's interface side-chains are flexible). We create 1000 models and score the sequence as the average score of the top 50 models (see below) to evaluate the binding ability of the specific peptide sequence.

Modeling of specific protein-peptide complexes—For each of the Bcl-xL, Mcl-1, and Bcl-2 receptors, we evaluated different available PDB structures to serve as templates to discriminate binders from non-binders in the TRAIN1 set, using a quick protocol that

involves only simple minimization of the template structure. *(1) Bcl-xL*: Two crystal structures of the Bcl-xL/Bim-BH3 peptide interaction are available (PDB id: 3io8, 3fdl). 3io8 (chains C & D) provided the best discrimination (see Supplementary Table S1), and we therefore refer to 3io8:CD as the 'native structure'. *(2) Mcl-1:* We evaluated a set of different PDB structures as templates (2nl9, 2pqk, 3d7v, 3io9, 3kj0, 3kj1, 3kj2, 3kz0). Among those, 2pqk provided the best discrimination of the TRAIN1 set. *(3) Bcl-2*: The receptor structure was taken from PDB: 2xa0, and the BH3 BIM peptide template from PDB: 2pqk.

Score function

We use the standard Rosetta score function (score12) for the simulations, with two minor adjustments: (1) the penalty for the burial of a carboxyl oxygen atom is increased (the ΔG^{free} parameter of the Lazaridis-Karplus solvation potential [57] was modified from −10 to −13.5 for this atom type). This was added to better account for a group of sequences that received poor experimental binding values in the TRAIN set, but good peptide scores. Inspection of the structural models detected a buried Asp that explains the discrepancy. (2) A weak short-range coulombic electrostatic energy term was added with weight 0.5 (as described in [58]).

We assessed each model by its "weighted score", which is the average of the Total score, Interface score and Peptide score (where Total score is the overall Rosetta energy score for the complex, Interface score is the energy of pair-wise interactions across the peptide-protein interface and Peptide score is the sum of the Rosetta energy function over the peptide residues). This score was proven useful in previous modeling studies using FlexPepDock (*e.g.* [44]).

Estimation of difference in SPOT arrays

To find the point mutations that differ the most between the Bcl-2 SPOT array experiment and prediction, the 200 scores (or −log(BLU)) were normalized to the average score of 10 wild type scores. The differences between the normalized scores were rank-ordered: the largest delta values indicate positions that differ most between the arrays.

Experimental measurement of Bcl-2 binding specificity with SPOT arrays

Human Bcl-2 protein and SPOT arrays were prepared and probed as previously described [21]. Briefly, Bcl-2 (residues 1–172) was expressed as a TEV-cleavable C-terminus fusion construct with His-tagged maltose binding protein (MBP). Purification was achieved using a Ni-NTA column followed by cleavage with TEV protease to yield N-terminally c-myc-tagged Bcl-2 product. A second Ni-NTA affinity step was used to capture His-tagged MBP and His-tagged TEV and the Bcl-2 fraction was further purified using size exclusion chromatography. SPOT arrays were synthesized on activated nitrocellulose support using Fmoc protection/deprotection chemistry by an Intavis AutoSpot robot in the MIT Biopolymers Laboratory. The 26-residue peptides were synthesized with PEG3 (three ethylene glycol units) linkers at the carboxy terminus. The array consisted of single site substitutions of 10 interfacial positions of the wild-type BIM sequence (MRPEIWIAQELRRIGDEFNAYYARRV) to all amino acids with the exception of cysteine and methionine. Resulting BIM substitution arrays were probed with either 100 nM or 1 μM Bcl-2 followed by detection with anti-c-myc-Cy3 antibody (Sigma Aldrich C6594) as previously described [21]. Fluorescence intensity for the probed SPOT arrays was measured on a Typhoon 9400 (GE Healthcare) and the resulting images were analyzed with ImageQuant (GE Healthcare) without background correction. Data from the 100 nM experiments were used in this work. Raw signals were used without background correction, measured in BLU.

Results

We have previously developed a general framework for the prediction of binding specificity of flexible peptides to protein receptors. In general, we model the structure of peptides with variable sequences to the target receptor using our high-resolution peptide docking protocol (Rosetta FlexPepDock [43, 44]) and use the energy scores of the models to assess binding ability. We have applied this protocol to determine which peptides can bind to the enzyme Farnesyltransferase (FTase), and consequently undergo farnesylation [35]. This protocol can potentially be applied to any new system after simple calibration on a set of sequences for which binding properties have experimentally been determined. Here, we first calibrate and test our protocol for the binding of BH3 derived helical peptides to Bcl-xL and Mcl-1 on different sets of experimental data reported by Dutta *et al.* [21]. We then apply the protocol to the homolog Bcl-2 and validate our predictions with a SPOT array experiment. Finally, we successfully reproduce binding specificity profiles for these three Bcl-2 derived proteins for a range of different human BH3-derived peptides.

(1) Calibration of the protocol

Dutta *et al.* performed two different SPOT peptide array experiments to determine the binding of different peptides to Bcl-xL and Mcl-1 [21]. TRAIN1 (n=360) is a combinatorial library of mutations at 5 positions of the BIM BH3 peptide. TRAIN2 (n=181) is a library of all possible point mutations at 10 different positions of the BIM BH3 peptide (See Fig. 1, and Methods for detailed description of the sets; TRAIN2 was probed by either 1μM or 100nM of the receptor). We modeled each of the sequences in these peptide libraries in complex with Bcl-xL and with Mcl-1. The binding energy of each of these models was then used to define their relative binding abilities to Bcl-xL and Mcl-1. We describe in the following how these sets were used to calibrate the sampling and scoring schemes of our protocol.

Several solved structures are available for both Bcl-xL and Mcl-1. To select the most suitable template for our protocol, we modeled the TRAIN1 set onto each of the different templates and minimized the threaded peptides in the context of the complex structure. To assess the performance of each template, we examined the Pearson correlation between 5 different scoring schemes (see Materials and Methods) and the experimental estimation of binding (represented as the logarithm of the SPOT binding intensity, BLU). The weighted score (“Wsc”), a scoring scheme including the total score of the complex while emphasizing the interactions across the interface and the internal energy of the peptide (see Materials and Methods) was shown useful in previous peptide modeling studies [44], and is indeed also adequate for both receptors in this case. In the following we report results obtained using this score. For Bcl-xL, the score of the minimized peptides with the template structure 3io8 [16] in the Protein Data Bank (PDB[59]) showed better correlation (r=0.54) than the other available template, while for Mcl-1 among 8 tested templates PDB: 2pqk [9] resulted in the best correlation (r=0.46).

We next applied the full FlexPepBind protocol on these templates: As opposed to the fast and simple minimization-based protocol, we created for each peptide sequence 200 models using our flexible peptide docking protocol FlexPepDock, and selected the top-scoring models as representatives for that sequence. This exhaustive protocol samples extensively the peptide flexibility within the binding pocket. Indeed, the top-scoring model of each peptide sequence improved the correlation (r=0.62 for both receptors). We note that results based on minimization of different templates correlate with the performance of these templates using the full FlexPepDock protocol (see Supplementary Table S1), and conclude that simple minimization is a good indicator for the initial selection of the best structural template.

Manual inspection of several false positive peptide sequences (*i.e.* sequences that in the experiment did not bind, but were nevertheless predicted to bind well by our protocol) revealed a common I3dD mutation in which the aspartic acid was buried in a hydrophobic pocket (numbering of peptide positions according the scheme used in Dutta *et al.* [21], see Fig. 1). This indicates that the burial of charges and polar atoms may not be penalized strongly enough in the current default Rosetta energy function. We therefore modified the Rosetta solvation model parameters (that are based on the Lazaridis-Karplus solvent exclusion model [57]) to further penalize buried carboxyl groups (see Materials and Methods). Application of our protocol with this adapted energy function slightly improved the correlation to experiment (Bcl-xL r=0.66; Mcl-1 r=0.68). Addition of a simple Coulombic electrostatic energy term[58] further improved performance for Bcl-xL (r=0.76), without any influence on the Mcl-1 predictions.

The final calibration of our protocol involved the investigation of two additional features that in previous applications improved predictions, namely the number of models created by the docking protocol, and the number of top-scoring models used to evaluate the binding energies. Overall, marginal improvement could be observed when the number of models was increased (from n=200 to n=1000). Scoring based on averaging the top 10- or top 50-scoring models did not significantly improve performance.

We note that significant improvement was observed for the Bcl-xL predictions for some of the tested parameter sets, but at the expense of reduced performance on Mcl-1. Since we aim at one general protocol that can be applied to this whole family, we focused on a protocol that performs best for both Bcl-xL and Mcl-1 (see Supplementary Table S1).

Overall, Table 1 and Fig. 2A & B show a good agreement between our predicted values and the logarithm of the measured SPOT intensities, both for the TRAIN1 set, as well as for the subsequently evaluated TRAIN2 set.

(2) Prediction of binding for the TRAIN 1 and TRAIN 2 SPOT arrays

Examination of the TRAIN correlation plots revealed two distinct regions: a group of non-binding sequences, which can be detected by high energy scores, and a group of binding sequences that correlate with our predicted score. This observation led us to inspect the performance of the protocol in a binary discrimination task. That is, we defined a threshold for the experimental binding value, below which a peptide is considered a ‘binder’ and above – a ‘non-binder’. If one wishes to utilize this protocol to detect high-affinity binders, this threshold should be higher (strict), whereas if the objective is to detect any type of binding this threshold should be lower (permissive). Given this binary labeling of the data points, we created Receiver Operating Characteristic (ROC) curves and calculated the area under the curve (AUC), which indicates the discrimination ability of our scoring. For both receptors, and on both training sets, we obtained AUC values above 0.83 (See Fig. 2C & D and Table 2). We report the results using an arbitrary threshold of −7.5 (logarithm of the normalized fluorescence intensity in the SPOT experiment) to define ‘binders’. However, we note that the predictions would improve with increasing threshold stringency, at the expense of detecting fewer binders (see Supplementary Fig. S1). This indicates that our protocol is particularly suitable for the selection of high affinity binding peptides.

We selected a score threshold of −223 for Bcl-xL and −325 for Mcl-1, corresponding to a false positive rate (FPR) of 10% on the TRAIN1 dataset (see ROC plots in Fig. 2C & D). These thresholds achieve a true positive rate (TPR) of 86% and 54% for Bcl-xL and Mcl-1 respectively on the TRAIN1 set. On the TRAIN2 sets, the same threshold gives a much higher TPR for Mcl-1 (Bcl-xL 88%; Mcl-1 96%) together with a corresponding increase in FPR (Bcl-xL 32%; Mcl-1 59%). This could be attributed to different backgrounds of the

SPOT arrays. It is also reasonable that the TRAIN1 and TRAIN2 sets contain different numbers of binders. TRAIN2 contains only single point mutations, which are expected to perturb binding less than the combinations of mutations explored in TRAIN1 (see Materials and Methods for descriptions of the datasets).

Although we sometimes refer below to “binding” and “non-binding” peptides, this is a system- and assay-specific designation and the interpretation should be that these are higher-affinity / lower-affinity binding peptides. Further, numerical score thresholds should not be applied to predictions of peptides from different sequence regimes. Rather, the relative ranking of scores within data sets serves as a better indicator for comparison.

I3d and F4a positions—Of particular interest in the TRAIN2 set are point mutations in BIM positions 3d and 4a. Mcl-1 binds almost singularly to peptides with a native isoleucine at position 3d, and any point mutation abolished binding on the SPOT arrays (at least at 100 nM concentration). In turn, Bcl-xL is promiscuous at the 3d position and allows almost any amino-acid point mutation. Position 4a shows the reverse behavior, *i.e.* many mutations can be introduced without impairing binding to Mcl-1, but binding to Bcl-xL is very specific. A plot of predicted scores for these point mutants largely recapitulates this behavior (see Fig. 3).

(3) Characterization of binding specificity of peptides isolated from a yeast surface display

In addition to the SPOT arrays, Dutta *et al.* also report a third set of sequences of peptides that were recovered from a yeast surface display library [21]. This set contains peptides that bind specifically to Bcl-xL (YEAST-X; n=40), to Mcl-1 (YEAST-M; n=33), or to both (YEAST-BOTH; n=17). We used this set to test the ability of our computational protocol to characterize binding specificity. For each of the peptides, we calculated binding scores for both Bcl-xL and Mcl-1. Fig. 4 shows that the Bcl-xL and Mcl-1 specific peptides are clearly separable using these two values. Our Bcl-xL score distinguishes Bcl-xL binders from non-binders, and detects most sequences that bind both proteins. However, FlexPepBind fails to differentiate Mcl-1 binders vs. non-binders using the Mcl-1 score. Mcl-1 specific peptides obtain a similar range of scores as the peptides that bind solely to Bcl-xL. This behavior was also observed by Dutta *et al.*, who used a SPOT-derived PSSM based on point mutation data (TRAIN2 in this study) to predict the yeast data. The PSSM was subsequently improved by assessing an additional set of mutants (TRAIN1 in this study) to create a modified model with better discrimination (see Fig. 6E in [21]).

Examination of our structural models suggests that the problem can be traced back to a set of peptides that contain a mutation at position 3b. According to the TRAIN1 set, most peptides containing an R3bD mutation are non-binders, but those that do bind are rescued by a F4aV mutation. Indeed, many of the YEAST-M sequences contain such an R3b to D/E mutation (13/33), or R3b to G mutation (9/33). 14 of these 22 are combined with the F4aV mutation (actually, most YEAST-M sequences, 19/33, contain valine at position 4a). Structurally, these positions are not spatially adjacent, and coupling between these positions is thus not expected (see Supplementary Fig. S2). It seems that our protocol fails to properly weight the contribution of the 4a valine, and therefore predicts the wrong outcome for these sequences. In contrast, the experiment-based PSSM approach overcomes this failure by using data from TRAIN1, which captures the valine preference (see Fig. 6F in [21]).

(4) Characterization of Bcl-2 binding specificity: SPOT array for Bcl-2

In order to evaluate the general ability of our protocol to characterize BH3 binding proteins, we extended our applications to another member of this family: Bcl-2. We applied the same protocol to a crystal structure of Bcl-2 in complex with a BH3 peptide (PDB: 2xa0 [15]), and

constructed an *in silico* SPOT array analogue to TRAIN2. Independently of the blind prediction, a SPOT array displaying 26-residue Bim-BH3 variants representing the TRAIN2 set was constructed using solid-phase synthesis in the same way as described by Dutta *et al.* [21] and probed with 100 nM Bcl-2. Fig. 5 displays the predicted Bcl-2 spot array side by side with the experimental results. Overall there is good agreement between the predicted binding and experimental results (r=0.65 which is similar to the values obtained for both Mcl-1 and Bcl-xL for this dataset; see Figure S3 in the supplementary material), indicating that indeed this approach might be extensible and general for additional receptors.

Both computational and experimental assessments indicated that the overall specificity profile emerging from this set for Bcl-2/BIM interaction is similar to that of Bcl-xL (r=0.88 between the two experiments). Nevertheless, some differences can be identified. Most notable is position 4a that previously showed divergent behavior for Mcl-1 and Bcl-xL [17, 21]. Similar to Mcl-1, Bcl-2 also shows promiscuity with regard to this position and is able to bind a range of different substitutions that do not bind Bcl-xL (compare last row of Fig. 5A & B). This trend was to some extent predicted a-priori (compare the last row of Fig. 5C & D). An additional position that was well predicted to show better binding to Bcl-2 over Bcl-xL across the array is position 2g. Substitutions at this site are more promiscuous with regards to Bcl-2 in comparison to Bcl-xL, which prefers mostly polar residues at this position. Lastly, Bcl-2 is a little more permissive for substitutions at position 2e. While this position still prefers Ala (native) or Gly, other small amino acids (Ser,Pro) bind equally well, and even bulkier residues are better accommodated than for Bcl-xL.

The similarity between the binding profiles of Bcl-2 and Bcl-xL can be identified using the *predicted* Bcl-2 array: Comparison of this array to the published experimental Bcl-xL array performed by Dutta et al. (Fig. 5C & B) reveals considerable similarities: e.g., the “Wild Type” column is among the darkest in both arrays, proline can hardly be incorporated into any position, and position 3e permits only A,G,S for both. Thus, even though quantitatively the Pearson correlation between the predicted Bcl-2 array and the experimental Bcl-xL array is lower than for the two experimental arrays (r=0.67 when compared to the 1μM Bcl-xL array), the overall agreement is surprisingly good.

(5) The binding specificity of additional BH3 peptides

Encouraged by the performance of FlexPepBind in this blind test for an additional Bcl-2 like protein, we wanted to assess its predictions for additional BH3 peptides. While until now all modeled peptide sequences were derived from the BIM BH3 peptide, the receptors encounter in the cell other BH3 sequences, such as those in proteins Bid, Bad, Bik, Noxa, Puma, Hrk and Bmf. We modeled these BH3 sequences and predicted their binding profiles to Bcl-xL, Bcl2 and Mcl-1. Table 3 compares the predicted binding to experimental IC50 values determined by Chen *et al.* [60]. Even though no linear correlation is observed, there is a good correspondence between the experimental and predicted values. All three receptors bind tightest to the BIM peptide, and BIM indeed obtains for all three receptors the best score among the BH3 peptides. Mcl-1 is the only receptor to bind NOXA, and indeed NOXA obtains a good binding value for Mcl-1, but very poor values for Bcl-xL and Bcl-2. The reverse is seen for BAD, which binds Bcl-xL and Bcl-2, but not Mcl-1 - and again the prediction results in intermediate binding values for the first two, and a very poor value for the latter. This is an example of the generalization ability of our protocol, which can be calibrated based on significantly less experimental data compared to corresponding sequence based prediction methods.

Discussion

The range of interactions mediated by the binding of Bcl-2 like proteins to BH3 domains in their partners plays an important role in the regulation of apoptosis[5]. This interaction has been intensively studied in the past using a plethora of different approaches[18, 21]. In addition, the detailed analysis of different solved crystal structures, and drug design into the BH3 binding pocket have probed important structural characteristics of this interaction [6–9]. This well-characterized system therefore represents an ideal platform for the validation and further improvement of approaches that attempt prediction of binding specificity in peptide-mediated interactions, such as the structure-based approach presented in this study. In turn, insights from successful modeling of peptide binding specificity can also further improve our understanding of the important but complex system of regulation of apoptosis.

Computational structural specificity prediction for flexible peptides

We have previously used structure-based prediction of binding specificity to successfully identify both known and novel protein Farnesyltransferase (FTase) substrate peptides [35]. The structure-based characterization of BH3 peptides binding to Bcl-2 - like proteins in this study demonstrates that this approach can indeed be generalized and applied to additional biological peptide-protein interactions. The two systems are markedly different: While for FTase, the binding motif is very short (4 residues) and constrained by 3 specific interactions (conserved hydrogen bonds and Zn coordination), BH3 peptides are much longer (~ 20 residues) and although constrained by their helical conformation, they display a much larger overall degree of flexibility. It is probably for this reason that the latter required more intensive modeling (in comparison to a minimization-only protocol that provided good results for FTase [35]).

Modifications in the protocol during calibration that improved its performance for the system at hand highlight general problems mainly in the scoring function used for the modeling of the peptide-receptor complex structure and the ranking of the different sequences. The main improvements in the performance of the protocol (measured as Pearson correlation to the experimental SPOT array binding values for the training sets) were the modification of the solvation model to further penalize desolvation of Asp and Glu, and the addition of an explicit electrostatic term. Both modifications make sense in light of the amphiphilic nature of the peptide-protein interface, and should now be assessed on a larger scale to see how they impact the overall modeling accuracy of a variety of Rosetta protocols. Current discrepancies between predictions and experiment, such as the possible underestimation of the binding energy of valine at position 4a, emphasize future directions to explore for improvements.

The analysis of binding specificity of peptides from yeast surface display highlights the advantages of a sequence-based approach rooted in experiments, where the experiments can be fine-tuned for a specific system by assessment of additional sequences for correction of background effects. This has been nicely demonstrated also in a study just published that has significantly extended the amount of sequenced peptides that bind to different bcl-2 proteins, and thus could report improved characterization of bcl-2 binding specificities[22]. Our structure-based approach in contrast uses experimental binding data merely to define a threshold. Despite its limitation in this specific case, this approach is expected to be more general in comparison to the PSSM-based methodology that will always be biased towards similar sequences (Indeed, we showed in our previous work on FTase substrates that our approach can be used to identify novel substrates not detectable based on sequence similarity [35]).

The conformational flexibility of the protein receptor might also play a major role in determining binding specificity. Structural studies have previously shown that Mcl-1 is able to accommodate surprising point mutations in the BIM BH3 sequence, without significant effect on its low nanomolar binding affinity. This is facilitated in part by side chain repacking (e.g. I2dA mutation [9]), as well as backbone movement of both the peptide and the Mcl-1 binding pocket (mainly the a3-a4 helices, that are 1.5 Å RMSD apart [9]). While our protocol is able to capture side chain movements and the peptide flexibility (which for different mutants can differ up to 1 Å RMSD from the native conformation [9]), we do not model the receptor backbone flexibility, which is critical for accommodation of some mutants (e.g. the binding of the BIM I2dY mutant to Mcl-1 [9]). Indeed, our predictions correctly identify I2dA and F4aE as binders, while the mutation I2dY that involves backbone changes achieves a score that does not pass the defined threshold. Targeted incorporation of receptor flexibility in the future will therefore improve performance, provided that the new False Positive conformations accessible thanks to increased flexibility can efficiently be screened.

Finally, the structural features of the unbound BH3 peptides might also contribute to their binding ability. The unstructured regions in BH3 containing proteins play an important regulatory role, and the length of these regions is regulated by alternative splicing. The BH3 peptides themselves are unstructured in solution, and adopt a helical conformation upon binding to Bcl-2 like family proteins [61]. Helical propensity itself might play a major role in the binding process: Indeed, studies have shown that mutations that mainly affect helical propensity (either by replacement with helix-favoring amino acids, or by elongation/shortening of the helix) can dramatically affect binding [62]. By itself however, helical propensity might not be the main determinant of binding, as assessment of helical propensity of the different mutants analyzed in the present study (using AGADIR [63]) did not show any clear and consistent correlation between the ease of helix formation and the ability of BIM-derived peptides to bind to Bcl-2 like proteins (the correlation coefficients of AGADIR scores with TRAIN1/TRAIN2 binding data were: Bcl-xL r=−0.6/0.12 and Mcl-1 r= −0.23/0.33, for the 100nM binding assays).

Large-scale identification of BH3 domains in humans and pathogens

We have shown that our approach is not limited to a specific receptor, nor to sequences derived from a specific BH3 peptide. We therefore anticipate that FlexPepBind can be used for the identification of novel BH3 binding sequences in the human genome and in other genomes, similar to its application to the identification of potential novel human FTase targets in our previous study [35]. BH3-only proteins are not limited to the Bcl-2 like protein family. Such functional sequences were found for example e.g. in an E3 ubiquitin ligase that utilizes a BH3 domain to bind specifically to its target Mcl-1 and mark it for degradation [64]. Application of our protocol to the putative BH3 domain of this protein (mule) indeed predicted a high-affinity interaction with Mcl-1 (data not shown). Another validated BH3 sequence can be found in nuclear Clusterin (CLU), which is a multifunctional glycoprotein that is overexpressed in prostate and breast cancers [65, 66]. Due to the poor sequence identity of known BH3 domains, sequence based predictors are not suitable for exact determination of the binding properties of BH3 like sequences but can rather serve as an initial fast screen to offer candidate sequences later to be filtered by FlexPepBind.

Apoptosis is part of the cell's combat arsenal against invading viruses. Consequently, many viruses have evolved strategies to undermine cell suicide. One such mechanism is the expression of Bcl-2-like anti-apoptotic proteins: For example, Epstein-Barr and Kaposi's sarcoma viruses both encode Bcl-2 homologs (BHRF-1 and KSHV Bcl-2, respectively). It was shown that these proteins are not only able to bind BH3 domain peptides found in the cell, but also bear unique specificity profiles: BHRF-1 binds Bim, Bid and Puma, but not

other BH3 domains [67], while KSHV Bcl-2 shows a specificity profile similar to Mcl-1 [68], binding most BH3 domains except for Bad [69]. Application of FlexPepBind on a solved crystal structure of BHRF1 bound to Bim (PDB: 2wh6 [67]) recovers the main binding preferences of BHRF1 to cellular BH3 domains: Bim and Bid are identified as tight binders (ranked 1 and 3), while Puma is ranked less favorably (rank 6 out of 8 BH3 peptides). The structure of KSHV Bcl-2 in complex with a ligand has not been solved yet, but an obvious extension of our protocol would be to predict binding using the free structure or a homology model. This would however require further calibration and testing. These preliminary results indicate that FlexPepBind might be a valuable tool for the detection and characterization of additional viral receptors.

Specificity profiles of Bcl-2 and Bcl-xL

Our results demonstrate that Bcl-2 and Bcl-xL share a very similar specificity profile for BIM-BH3 derived point mutants. This is surprising in light of the fact that the two proteins are not highly conserved: only 14 of the 21 interface residues in Bcl-xL are conserved in Bcl-2 (i.e. 66% sequence identity at the interface; overall sequence identity of 40%). On the other hand, previous studies have shown that the binding preferences for biological BH3 sequences of these two proteins are indeed similar. What then is the biological role of this apparent redundancy? One possibility is robustness: owing to their cardinal role in determining cell fate, they act as a backup for each other (this motif is found in many signal transduction pathways). Another possibility is that the small differences in binding preferences (such as a ~100-fold preference for Bcl-xL to bind Bik or Bid) might lead to an overall different outcome when integrating the cellular network of interactions. Indeed, in a large-scale assessment of squamous cell carcinoma, most cell lines upregulated either Bcl-2 or Bcl-xL, but not both. Moreover the survival profiles of the different subgroups differed significantly [70]. This and other studies indicate that these two proteins play similar but not identical roles [71]. Further investigation of these differences with tools such as the one presented in this study may help to obtain a more comprehensive view of the underlying features of the cellular network of interactions and their contribution to cellular behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank S. Chen for thoughtful comments on the work, and R. Cook and the MIT Biopolymers Laboratory for help with SPOT array synthesis.

This work was funded by the Israel Science Foundation, founded by the Israel Academy of Science and Humanities (grant number 319/11) and the USA-Israel Binational Science Foundation (grant number 2009418) to OSF, and by NIH R01GM84181 to AK. NL was supported by a Converging Technologies Scholarship funded by the Planning and Budgeting Committee of the Israeli Council for higher education.

Abbreviations

AUC	Area Under roc Curve
Bcl2	B cell lymphoma 2
BH	Bcl2-homology domain
BIM	Bcl-2-interacting mediator of cell death
BLU	Boehringer Light Unit

FPR	False Positive Rate
FTase	FarnesylTransferase
Mcl1	myeloid leukemia cell differentiation protein
PDB	Protein Data Bank
PSSM	Position Specific Scoring Matrix
RMSD	Root Mean Square Deviation
ROC	Receiving Operating Characteristic
TPR	True Positive Rate
Wsc	Weighted Score

References

1. Youle RJ, Strasser A. The BCL-2 protein family: opposing activities that mediate cell death. Nat Rev Mol Cell Biol. 2008; 9:47–59. [PubMed: 18097445]
2. Adams JM, Cory S. The Bcl-2 protein family: arbiters of cell survival. Science. 1998; 281:1322–1326. [PubMed: 9735050]
3. Shamas-Din A, Brahmbhatt H, Leber B, Andrews DW. BH3-only proteins: Orchestrators of apoptosis. Biochim Biophys Acta. 2011; 1813:508–520. [PubMed: 21146563]
4. Giam M, Huang DC, Bouillet P. BH3-only proteins and their roles in programmed cell death. Oncogene. 2008; 27(Suppl 1):S128–S136. [PubMed: 19641498]
5. Lomonosova E, Chinnadurai G. BH3-only proteins in apoptosis and beyond: an overview. Oncogene. 2008; 27(Suppl 1):S2–S19. [PubMed: 19641503]
6. Liu X, Dai S, Zhu Y, Marrack P, Kappler JW. The structure of a Bcl-xL/Bim fragment complex: implications for Bim function. Immunity. 2003; 19:341–352. [PubMed: 14499110]
7. Petros AM, Olejniczak ET, Fesik SW. Structural biology of the Bcl-2 family of proteins. Biochim Biophys Acta. 2004; 1644:83–94. [PubMed: 14996493]
8. Lama D, Sankararamakrishnan R. Identification of core structural residues in the sequentially diverse and structurally homologous Bcl-2 family of proteins. Biochemistry. 2010; 49:2574–2584. [PubMed: 20141168]
9. Fire E, Gulla SV, Grant RA, Keating AE. Mcl-1-Bim complexes accommodate surprising point mutations via minor structural changes. Protein Sci. 2010; 19:507–519. [PubMed: 20066663]
10. Letai AG. Diagnosing and exploiting cancer's addiction to blocks in apoptosis. Nat Rev Cancer. 2008; 8:121–132. [PubMed: 18202696]
11. Kang MH, Reynolds CP. Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. Clin Cancer Res. 2009; 15:1126–1132. [PubMed: 19228717]
12. Mohammad R, Giri A, Goustin AS. Small-molecule inhibitors of Bcl-2 family proteins as therapeutic agents in cancer. Recent Pat Anticancer Drug Discov. 2008; 3:20–30. [PubMed: 18289121]
13. 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; 435:677–681. [PubMed: 15902208]
14. Nguyen M, Marcellus RC, Roulston A, Watson M, Serfass L, Murthy Madiraju SR, Goulet D, Viallet J, Belec L, Billot X, Acoca S, Purisima E, Wiegmans A, Cluse L, Johnstone RW, Beauparlant P, Shore GC. Small molecule obatoclax (GX15-070) antagonizes MCL-1 and overcomes MCL-1-mediated resistance to apoptosis. Proc Natl Acad Sci U S A. 2007; 104:19512–19517. [PubMed: 18040043]

15. Ku B, Liang C, Jung JU, Oh BH. Evidence that inhibition of BAX activation by BCL-2 involves its tight and preferential interaction with the BH3 domain of BAX. Cell Res. 2011; 21:627–641. [PubMed: 21060336]
16. Lee EF, Czabotar PE, Yang H, Sleebs BE, Lessene G, Colman PM, Smith BJ, Fairlie WD. Conformational changes in Bcl-2 pro-survival proteins determine their capacity to bind ligands. J Biol Chem. 2009; 284:30508–30517. [PubMed: 19726685]
17. Lee EF, Czabotar PE, Smith BJ, Deshayes K, Zobel K, Colman PM, Fairlie WD. Crystal structure of ABT-737 complexed with Bcl-xL: implications for selectivity of antagonists of the Bcl-2 family. Cell Death Differ. 2007; 14:1711–1713. [PubMed: 17572662]
18. Boersma MD, Sadowsky JD, Tomita YA, Gellman SH. Hydrophile scanning as a complement to alanine scanning for exploring and manipulating protein-protein recognition: application to the Bim BH3 domain. Protein Sci. 2008; 17:1232–1240. [PubMed: 18467496]
19. Lee EF, Czabotar PE, van Delft MF, Michalak EM, Boyle MJ, Willis SN, Puthalakath H, Bouillet P, Colman PM, Huang DC, Fairlie WD. A novel BH3 ligand that selectively targets Mcl-1 reveals that apoptosis can proceed without Mcl-1 degradation. J Cell Biol. 2008; 180:341–355. [PubMed: 18209102]
20. Day CL, Smits C, Fan FC, Lee EF, Fairlie WD, Hinds MG. Structure of the BH3 domains from the p53-inducible BH3-only proteins Noxa and Puma in complex with Mcl-1. J Mol Biol. 2008; 380:958–971. [PubMed: 18589438]
21. Dutta S, Gulla S, Chen TS, Fire E, Grant RA, Keating AE. Determinants of BH3 binding specificity for Mcl-1 versus Bcl-xL. J Mol Biol. 2010; 398:747–762. [PubMed: 20363230]
22. Debartolo J, Dutta S, Reich L, Keating AE. Predictive Bcl-2 Family Binding Models Rooted in Experiment or Structure. J Mol Biol. 2012
23. Lee EF, Fedorova A, Zobel K, Boyle MJ, Yang H, Perugini MA, Colman PM, Huang DC, Deshayes K, Fairlie WD. Novel Bcl-2 homology-3 domain-like sequences identified from screening randomized peptide libraries for inhibitors of the pro-survival Bcl-2 proteins. J Biol Chem. 2009; 284:31315–31326. [PubMed: 19748896]
24. Kalyoncu S, Keskin O, Gursoy A. Interaction prediction and classification of PDZ domains. BMC Bioinformatics. 2010; 11:357. [PubMed: 20591147]
25. Ferraro E, Via A, Ausiello G, Helmer-Citterich M. A neural strategy for the inference of SH3 domain-peptide interaction specificity. BMC Bioinformatics. 2005; 6(Suppl 4):S13. [PubMed: 16351739]
26. Obenauer JC, Yaffe MB. Computational prediction of protein-protein interactions. Methods Mol Biol. 2004; 261:445–468. [PubMed: 15064475]
27. Liao WW, Arthur JW. Predicting peptide binding to Major Histocompatibility Complex molecules. Autoimmun Rev. 2011; 10:469–473. [PubMed: 21333759]
28. Kaufmann K, Shen N, Mizoue L, Meiler J. A physical model for PDZ-domain/peptide interactions. J Mol Model. 2011; 17:315–324. [PubMed: 20461427]
29. Hou T, Xu Z, Zhang W, McLaughlin WA, Case DA, Xu Y, Wang W. Characterization of domain-peptide interaction interface: a generic structure-based model to decipher the binding specificity of SH3 domains. Mol Cell Proteomics. 2009; 8:639–649. [PubMed: 19023120]
30. Chaudhury S, Gray JJ. Identification of structural mechanisms of HIV-1 protease specificity using computational peptide docking: implications for drug resistance. Structure. 2009; 17:1636–1648. [PubMed: 20004167]
31. Grigoryan G, Keating AE. Structure-based prediction of bZIP partnering specificity. J Mol Biol. 2006; 355:1125–1142. [PubMed: 16359704]
32. Apgar JR, Hahn S, Grigoryan G, Keating AE. Cluster expansion models for flexible-backbone protein energetics. J Comput Chem. 2009; 30:2402–2413. [PubMed: 19360809]
33. Smith CA, Kortemme T. Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. J Mol Biol. 2010; 402:460–474. [PubMed: 20654621]
34. King CA, Bradley P. Structure-based prediction of protein-peptide specificity in Rosetta. Proteins. 2010; 78:3437–3449. [PubMed: 20954182]

35. London N, Lamphear C, Houghland J, Fierke C, Schueler-Furman O. Identification of a novel class of Farnesylation targets by structure-based modeling of binding specificity. PLoS Comput Biol. 2011; 7:e1002170. [PubMed: 21998565]
36. Fu X, Apgar JR, Keating AE. Modeling backbone flexibility to achieve sequence diversity: the design of novel alpha-helical ligands for Bcl-xL. J Mol Biol. 2007; 371:1099–1117. [PubMed: 17597151]
37. Mandell DJ, Kortemme T. Backbone flexibility in computational protein design. Curr Opin Biotechnol. 2009; 20:420–428. [PubMed: 19709874]
38. London N, Movshovitz-Attias D, Schueler-Furman O. The structural basis of peptide-protein binding strategies. Structure. 2010; 18:188–199. [PubMed: 20159464]
39. Vanhee P, Reumers J, Stricher F, Baeten L, Serrano L, Schymkowitz J, Rousseau F. PepX: a structural database of non-redundant protein-peptide complexes. Nucleic Acids Res. 2010; 38:D545–D551. [PubMed: 19880386]
40. Vanhee P, Stricher F, Baeten L, Verschueren E, Lenaerts T, Serrano L, Rousseau F, Schymkowitz J. Protein-peptide interactions adopt the same structural motifs as monomeric protein folds. Structure. 2009; 17:1128–1136. [PubMed: 19679090]
41. Stein A, Ceol A, Aloy P. 3did: identification and classification of domain-based interactions of known three-dimensional structure. Nucleic Acids Res. 2011; 39:D718–D723. [PubMed: 20965963]
42. Stein A, Aloy P. Contextual specificity in peptide-mediated protein interactions. PLoS One. 2008; 3:e2524. [PubMed: 18596940]
43. Raveh B, London N, Schueler-Furman O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins. 2010; 78:2029–2040. [PubMed: 20455260]
44. Raveh B, London N, Zimmerman L, Schueler-Furman O. Rosetta FlexPepDock ab-initio: Simultaneous Folding, Docking and Refinement of Peptides onto Their Receptors. PLoS One. 2011; 6:e18934. [PubMed: 21572516]
45. London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O. Rosetta FlexPepDock web server--high resolution modeling of peptide-protein interactions. Nucleic Acids Res. 2011
46. Antes I. DynaDock: A new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility. Proteins. 2010; 78:1084–1104. [PubMed: 20017216]
47. Niv MY, Weinstein H. A flexible docking procedure for the exploration of peptide binding selectivity to known structures and homology models of PDZ domains. J Am Chem Soc. 2005; 127:14072–14079. [PubMed: 16201829]
48. Arun Prasad P, Gautham N. A new peptide docking strategy using a mean field technique with mutually orthogonal Latin square sampling. J Comput Aided Mol Des. 2008; 22:815–829. [PubMed: 18465087]
49. Schon O, Friedler A, Bycroft M, Freund SM, Fersht AR. Molecular mechanism of the interaction between MDM2 and p53. J Mol Biol. 2002; 323:491–501. [PubMed: 12381304]
50. Grigoryan G, Reinke AW, Keating AE. Design of protein-interaction specificity gives selective bZIP-binding peptides. Nature. 2009; 458:859–864. [PubMed: 19370028]
51. Jochim AL, Arora PS. Assessment of helical interfaces in protein-protein interactions. Mol Biosyst. 2009; 5:924–926. [PubMed: 19668855]
52. London N, Raveh B, Movshovitz-Attias D, Schueler-Furman O. Can self-inhibitory peptides be derived from the interfaces of globular protein-protein interactions? Proteins. 2010; 78:3140–3149. [PubMed: 20607702]
53. Henchey LK, Jochim AL, Arora PS. Contemporary strategies for the stabilization of peptides in the alpha-helical conformation. Curr Opin Chem Biol. 2008; 12:692–697. [PubMed: 18793750]
54. 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; 6:595–601. [PubMed: 20562877]
55. Schreiber G, Keating AE. Protein binding specificity versus promiscuity. Curr Opin Struct Biol. 2011; 21:50–61. [PubMed: 21071205]
56. Hilpert K, Winkler DF, Hancock RE. Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. Nat Protoc. 2007; 2:1333–1349. [PubMed: 17545971]

57. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins. 1999; 35:133–152. [PubMed: 10223287]
58. Yanover C, Bradley P. Extensive protein and DNA backbone sampling improves structure-based specificity prediction for C2H2 zinc fingers. Nucleic Acids Res. 2011; 39:4564–4576. [PubMed: 21343182]
59. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000; 28:235–242. [PubMed: 10592235]
60. Chen L, Willis SN, Wei A, Smith BJ, Fletcher JI, Hinds MG, Colman PM, Day CL, Adams JM, Huang DC. Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. Mol Cell. 2005; 17:393–403. [PubMed: 15694340]
61. Rautureau GJ, Day CL, Hinds MG. Intrinsically disordered proteins in bcl-2 regulated apoptosis. Int J Mol Sci. 2010; 11:1808–1824. [PubMed: 20480043]
62. Petros AM, Nettesheim DG, Wang Y, Olejniczak ET, Meadows RP, Mack J, Swift K, Matayoshi ED, Zhang H, Thompson CB, Fesik SW. Rationale for Bcl-xL/Bad peptide complex formation from structure, mutagenesis, and biophysical studies. Protein Sci. 2000; 9:2528–2534. [PubMed: 11206074]
63. Munoz V, Serrano L. Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. J Mol Biol. 1995; 245:275–296. [PubMed: 7844817]
64. Warr MR, Acoca S, Liu Z, Germain M, Watson M, Blanchette M, Wing SS, Shore GC. BH3-ligand regulates access of MCL-1 to its E3 ligase. FEBS Lett. 2005; 579:5603–5608. [PubMed: 16213503]
65. Lee DH, Ha JH, Kim Y, Bae KH, Park JY, Choi WS, Yoon HS, Park SG, Park BC, Yi GS, Chi SW. Interaction of a putative BH3 domain of clusterin with anti-apoptotic Bcl-2 family proteins as revealed by NMR spectroscopy. Biochem Biophys Res Commun. 2011; 408:541–547. [PubMed: 21527247]
66. Kim N, Yoo JC, Han JY, Hwang EM, Kim YS, Jeong EY, Sun CH, Yi GS, Roh GS, Kim HJ, Kang SS, Cho GJ, Park JY, Choi WS. Human nuclear clusterin mediates apoptosis by interacting with Bcl-XL through C-terminal coiled coil domain. J Cell Physiol. 2012; 227:1157–1167. [PubMed: 21567405]
67. Kvansakul M, Wei AH, Fletcher JI, Willis SN, Chen L, Roberts AW, Huang DC, Colman PM. Structural basis for apoptosis inhibition by Epstein-Barr virus BHRF1. PLoS Pathog. 2010; 6:e1001236. [PubMed: 21203485]
68. Flanagan AM, Letai A. BH3 domains define selective inhibitory interactions with BHRF-1 and KSHV BCL-2. Cell Death Differ. 2008; 15:580–588. [PubMed: 18084238]
69. Banadyga L, Lam SC, Okamoto T, Kvansakul M, Huang DC, Barry M. Deerpox virus encodes an inhibitor of apoptosis that regulates Bak and Bax. J Virol. 2011; 85:1922–1934. [PubMed: 21159883]
70. Pena JC, Thompson CB, Recant W, Vokes EE, Rudin CM. Bcl-xL and Bcl-2 expression in squamous cell carcinoma of the head and neck. Cancer. 1999; 85:164–170. [PubMed: 9921989]
71. van Delft MF, Huang DC. How the Bcl-2 family of proteins interact to regulate apoptosis. Cell Res. 2006; 16:203–213. [PubMed: 16474435]

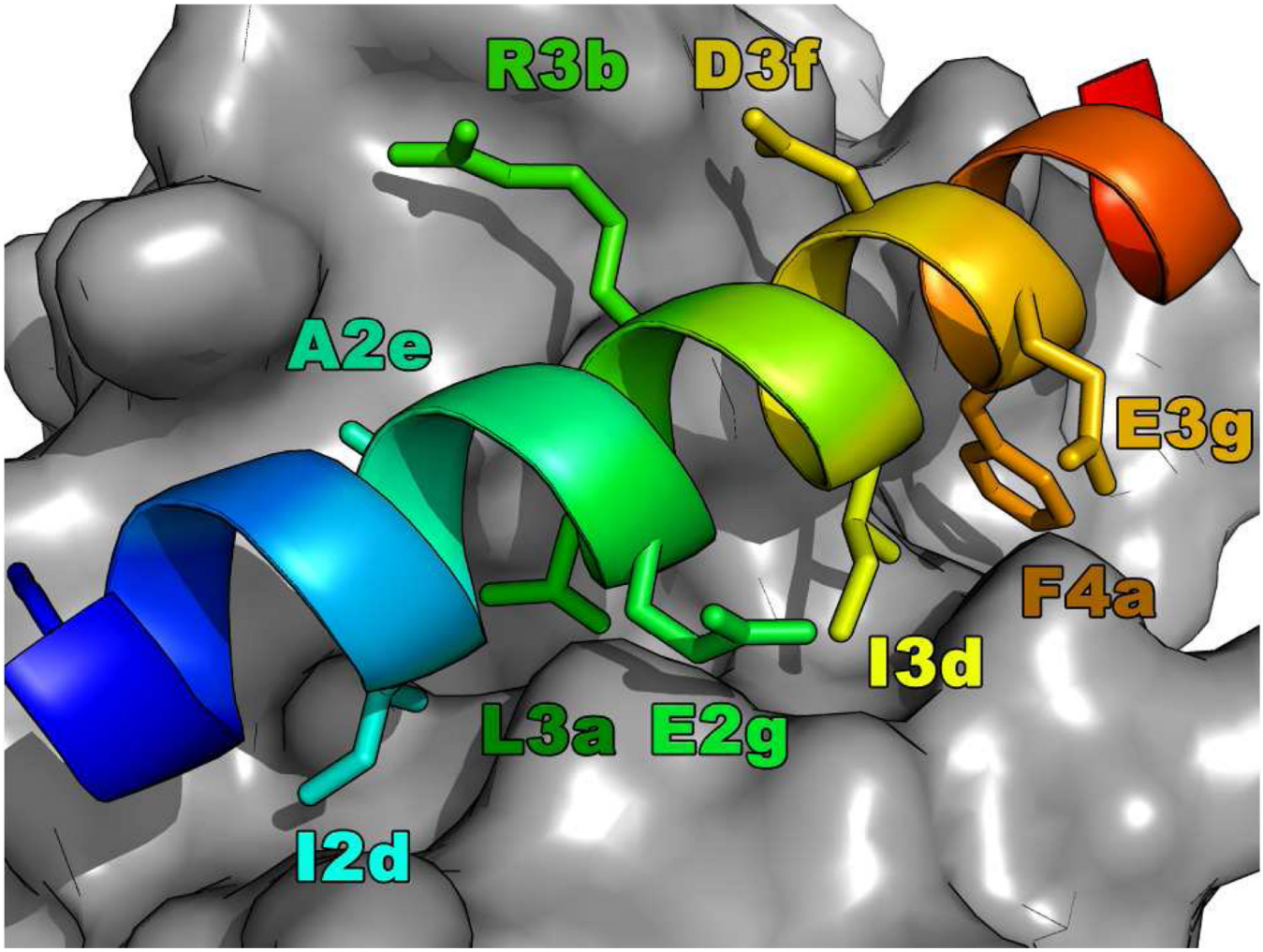

Fig. 1. BIM peptide in complex with Mcl1

The amphiphilic BIM BH3-only helical peptide (cartoon representation, rainbow coloring) binds Mcl1 (surface representation, gray) in a hydrophobic groove. 10 peptide interface positions that were mutated in the different datasets used in this study are indicated in stick representation, accompanied with their position nomenclature.

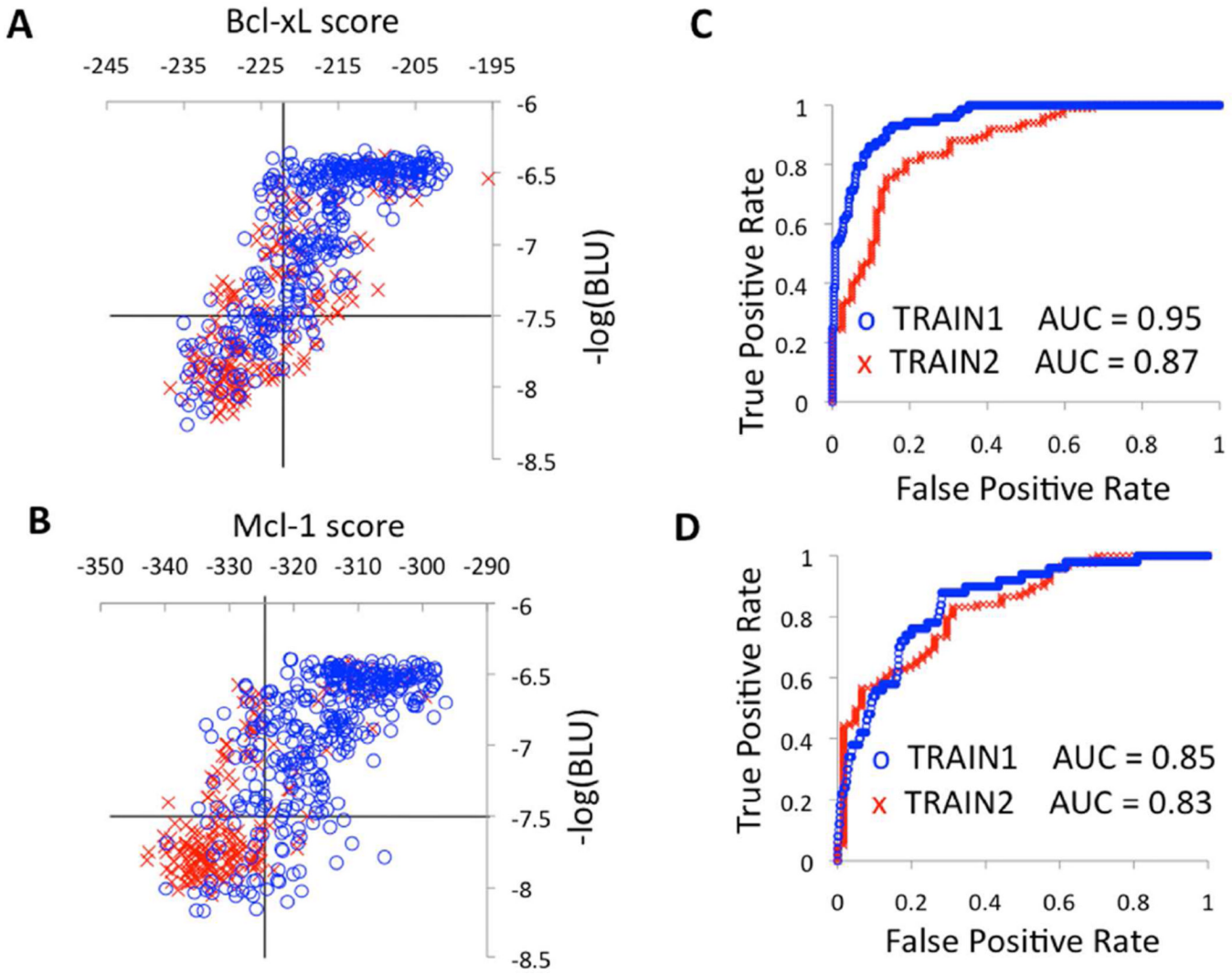

Fig. 2. Predicted binding values for BIM-derived peptides to Bcl-xL and Mcl-1 are in good agreement with experimental results from SPOT arrays

A.+B. Correlation of calculated score (X-axis) to the experimental peptide binding signal (Y-axis; −log(SPOT intensity) measured in BLU) for (A) Bcl-xL and (B) Mcl-1, for TRAIN1 (blue circles) and TRAIN2 probed with 1μM receptor (red crosses) (see Materials and Methods for description; the experimental data for TRAIN2 are also shown in Fig. 5A). The −7.5 −log (BLU) threshold (see Fig. S1), and the selected score thresholds are indicated with lines. **C.+D**. ROC curves for the discrimination of binders and non binders (threshold for binding defined as −log(BLU)=−7.5) for (C) Bcl-xL and (D) Mcl-1. Black circles indicate the selected thresholds. The high Area Under the Curve (AUC) values demonstrate good discrimination between BH3 Bim-derived peptides that bind / do not bind to different Bcl-2 binding proteins.

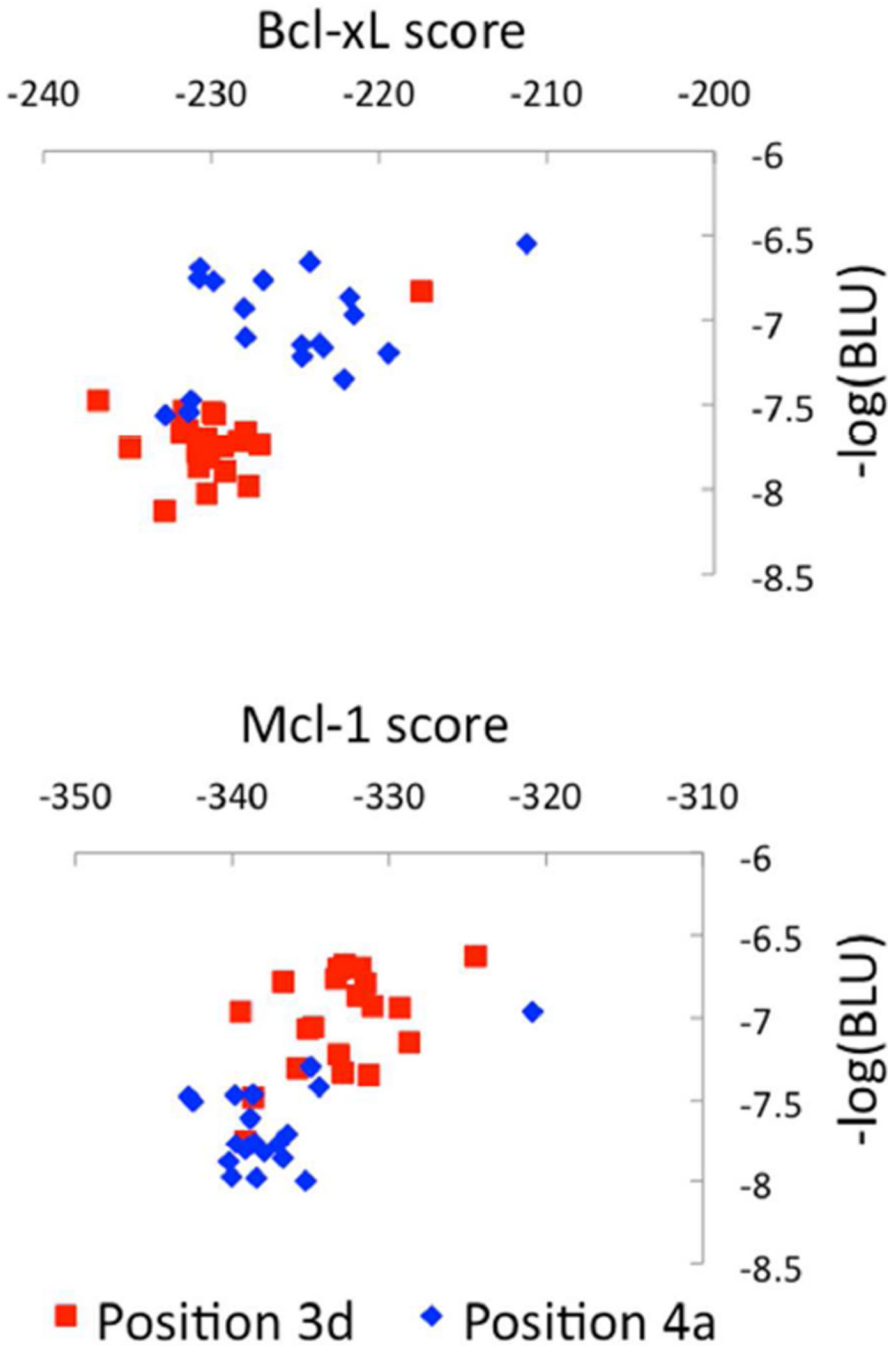

Fig. 3. Binding specificity and promiscuity at positions 3d and 4a

Binding score estimation for peptide binding to Bcl-xL (top) and Mcl-1 (bottom) recapitulates distinct specificity-determining residues identified using SPOT experiments [21]. Bcl-xL binding peptides show a restricted range of amino acids at position 4a (blue diamonds), while many residues are tolerated at position 3d (red squares). Mcl-1 binding peptides in contrast show specific preferences for position 3d, but little discrimination at position 4a.

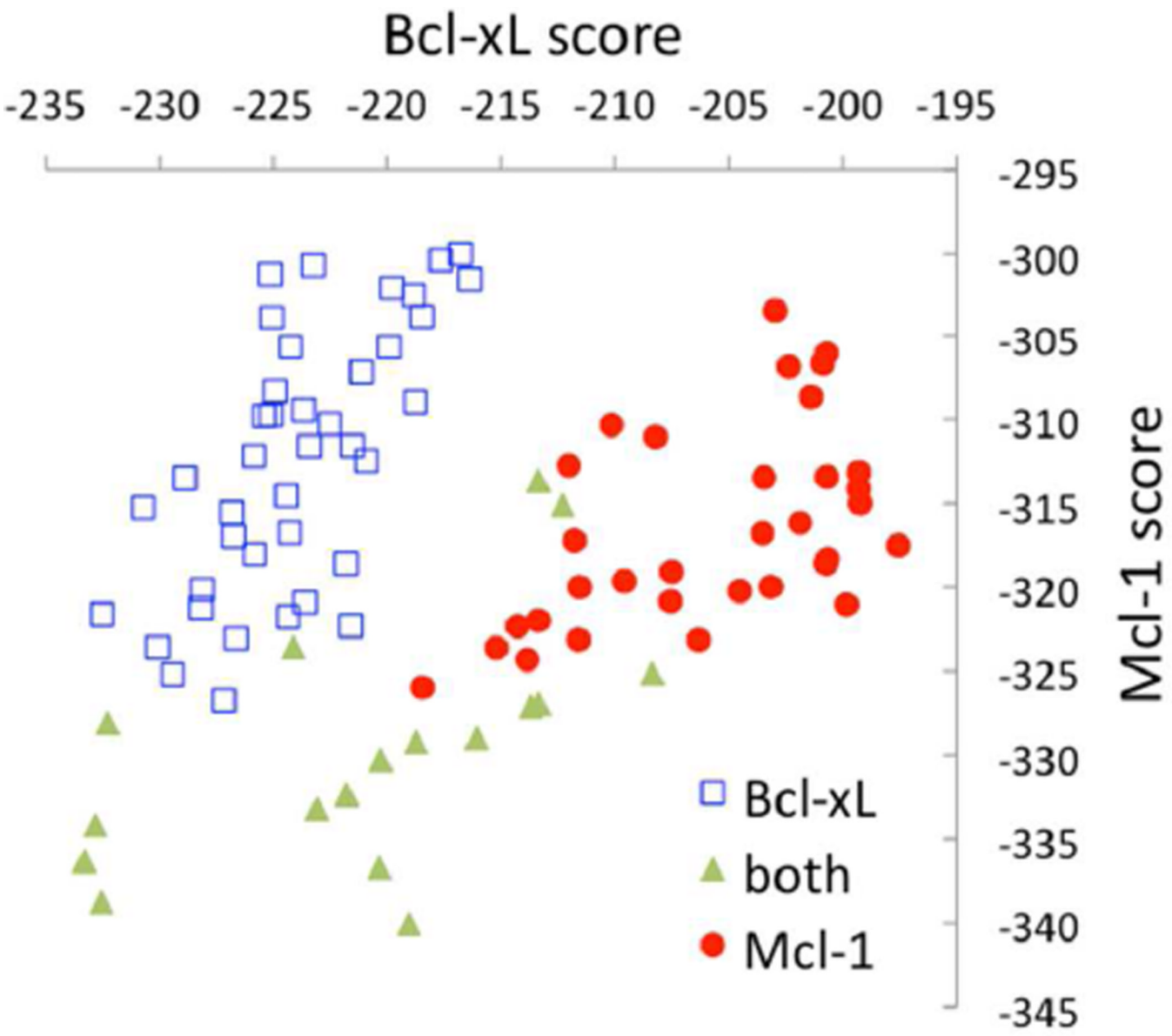

Fig. 4. Specificity prediction for yeast display-derived peptides

FlexPepBind is able to discriminate between Bcl-xL binders/non-binders (Bcl-xL specific binders all obtain good scores for binding to Bcl-xL, but worse scores for binding to Mcl-1; blue squares in upper left quadrant), as well as detect sequences that bind both proteins (peptides that bind to both proteins mostly obtain good scores for both; green triangles in lower left quadrant). However, it performs badly in detecting the Mcl-1 specific peptides, which all obtain unfavorable scores both for Bcl-xL and Mcl-1 binding (red circles in upper right quadrant–should be located in lower right quadrant).

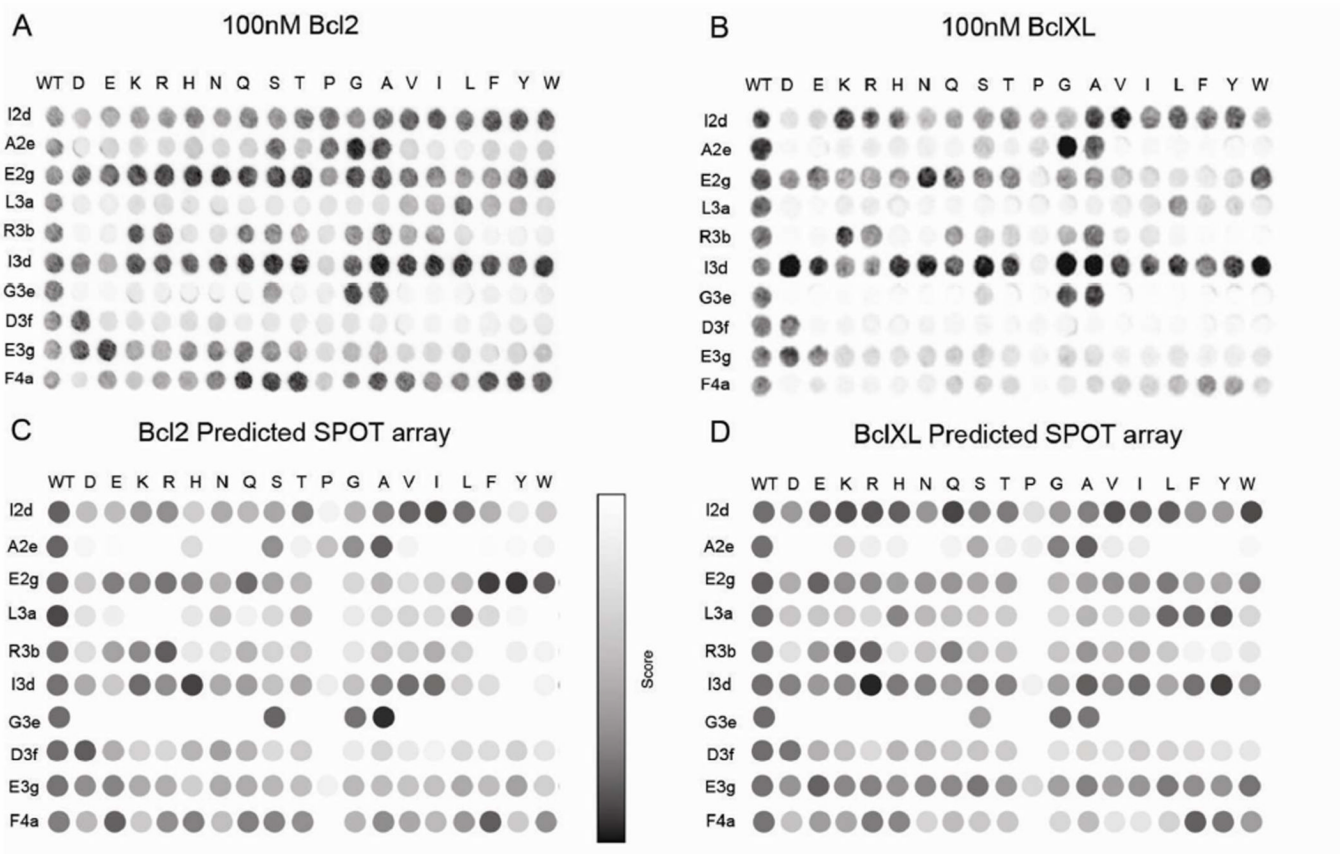

Fig. 5. Computational and experimental characterization of Bcl-xL and Bcl-2 BIM-derived peptide-binding specificities

Sequences of 180 point mutants of the BIM BH3 peptide (TRAIN2) were evaluated for binding to Bcl-2 (left panels) and Bcl-xL (right panels). Experimental and predicted SPOT arrays are shown in the upper and lower panels, respectively. The predicted SPOT intensities are similar to the experimental SPOT arrays (r=0.72 for Bcl-xL and 0.67 for Bcl-2). The experimental specificity profiles of Bcl-xL and Bcl-2 are very similar (r=0.88; see Text for comparison). The array image in Fig. 5A was taken from [21].

Table 1

Pearson correlation between predicted and experimental binding values of BIM-derived peptides to Bcl-xL and Mcl-1[a]

	BIM BH3-derived peptide sets[b]		
	TRAIN1 (n=360; 100nM)	**TRAIN2 (n=181; 1μM)**	**TRAIN2 (n=181; 100nM)**
Bcl-xL	r=0.80	r=0.73	r=0.69
Mcl-1	r=0.67	r=0.71	r=0.65

[a]Experimental values were taken from SPOT array experiments [21]. Binding predictions were performed using the weighted score of the peptide (Wsc), according to the calibrated protocol described in the text (see Materials and Methods and Results for more details, and Supplementary Table S1 for alternative scoring schemes).

[b]TRAIN1: Combinatorial library of mutations at 5 positions. TRAIN2: 180 point mutations at 10 positions [21]. TRAIN2 array was probed by either 1μM or 100nM of the receptor.

Table 2

Identification of binders and non-binders in the training sets. Binders and non-binders were defined using a threshold of −log(BLU) = −7.5.

		TRAIN1 (100nM)					TRAIN2 (1μM)				
	Threshold[a]	**AUC**[b]	**Bind**[c]	**NON**[d]	**TPR**	**FPR**	**AUC**	**Bind**	**NON**	**TPR**	**FPR**
Bcl-xL	−223	0.95	73	287	86%	10%	0.89	121	79	90%	32%
Mcl-1	−325	0.85	50	310	54%	10%	0.85	139	61	96%	59%

[a]Score threshold set to achieve 10% FPR rate on TRAIN1 (see Fig. 2)

[b]Area Under the ROC plot

[c/d]Number of ‘binding’/’non-binding’ peptides in the set (according to a −7.5 −log(BLU) threshold)

Table 3

Computational recapitulation of distinct binding preferences of Bcl-2 family proteins to different BH3 substrates.

	Peptide binding affinity		
	<100 nM[a]	**<10 μM**	**>100 μM**
Bcl-xL	**Bim, Bmf, Bad, Puma, Hrk, Bid, Bik**		**Noxa**
Predicted rank[b]	1, 2, 3, 4, 5, 6, 8		7
Bcl-2	**Bim, Bmf, Puma, Bad**	**Hrk, Bid, Bik**	**Noxa**
Predicted rank	1, 2, 4, 6	3, 5, 7	8
Mcl-1	**Bim, Noxa, Puma**	**Hrk, Bmf, Bik, Bid**	**Bad**
Predicted rank	1, 2, 3	4, 5, 6, 7	8

[a] IC50 values of Bcl-2 like protein with BH3-only derived peptide (as reported in [60]).

[b] Predicted rank of binding by FlexPepBind among the biological substrates Bim, Bid, Bad, Bik, Noxa, Puma, Hrk & Bmf.