Supporting Information

Analysis of Hydrophobic Binding Hotspots of Highly Flexible Bcl-xL Protein-Protein Interfaces by Cosolvent Molecular Dynamics Simulation

*Chao-Yie Yang, and Shaomeng Wang**

Departments of Internal Medicine, Pharmacology and Medicinal Chemistry, University of Michigan, 1500 E. Medical Center Drive, Ann Arbor, MI 48109, USA

* Corresponding author: Email: shaomeng@umich.edu; Phone:+1 734 6150362; Fax:+1 734 6479647.

Methods and Calculations

Preparation of molecular dynamic simulation

Crystal structures of one apo-Bcl-xL, three Bcl-xL, and the BH3 domain peptides taken from the Protein Databank[1] were used in this study. The PDB IDs for them are, 1MAZ, 3FDL (with the Bim BH3 peptide), 2BZW (with the Bad BH3 peptide) and 2P1L (with the Bec1 BH3 peptide). Because a loop segment connecting the α1 and the α2 helix (naming see Figure S1) are missing in all available crystal structures, we capped the C-terminus of α1 and the N-terminus of α2 with the N-methyl group (NME) and the acetyl group (ACE), respectively, to mimic the backbone of the missing flexible loop. The sequences of Bcl-xL in 1MAZ, 3FDL and 2P1L are for human while that in 2BZW is for mouse. To make the same comparison between the Bad-bound and Bec1-bound Bcl-xL conformations, we made two mutations at A168S and E193D. The Bcl-xL protein used in our simulations consists of 139 residues (S4—S25-NME, ACE-S83—N197) with the sequence:

SNRELVVDFLSYKLSQKGYSWS-(NME), (ACE)-SEAVKQALREAGDEFEL RYRRAFSDLTSQLHITPGTAYQSFEQVVNELFRDGVNWGRIVAFFSFGGALCVES VDKEMQVLVSRIAAWMATYLNDHLEPWIQENGGWDTFVELYGN.

In preparation of the simulations, we have used the PROPKA[2-3] to determine the protonation state of ionizable groups on Bcl-xL. We found all the titratable groups should be charged according to standard physiological conventions. PMEMD from Amber (version 10)[4] was used for molecular dynamics simulations. The Amber 99SB force field parameters[5] were used for the amino acids.

Pure water solvent simulations

To prepare the topology and coordinate files, counter ions were added to neutralize the charges in Bcl-xL before it was placed in a 13Å octahedral box of water. The TIP3P[6] water model was used. A 3000-step minimization (steps 1-1000 using conjugated gradient followed by 2000 steps steepest decent) was first carried out. After minimization, a 500 ps constant volume and constant temperature (NVT) simulation was performed to raise the temperature of the system to 298K while constraining backbone atoms with a 5 kcal/mol/Å^2 force constant with reference to the crystal structure. A second 200 ps constant pressure and constant temperature (NPT) simulation at 298 K was performed while constraining backbone atoms with a 2 kcal/mol/Å^2 force constant with reference to the crystal structure. The system was then equilibrated for 1 ns at 298K without any constraints. The system is then ready for the 32 ns production run. All the MD simulations were in the isobaric isothermal (NPT, T = 298K and P = 1 atm) ensemble. The SHAKE[7] algorithm was used to fix bonds involving hydrogen. The PME method[8] was used and the non-bonded cutoff distance was set at 10Å. The time step was 2 fs, and neighboring pairs list was updated every 20 steps. A 32 ns MD simulation of the 139 amino acids Bcl-xL required 24 days of four dedicated dual core dual AMD 2.0 GHz processors nodes in our cluster. Advances in computing facility should relieve the limitation of this approach to study larger protein systems and allow investigating dynamical motions proteins at a longer timescale.

Cosolvent mapping simulations and analyses

An equilibrated cosolvent box (20% v/v isopropanol in water) was provided by Dr. Barril. For MD simulations, the structure of Bcl-xL was first neutralized with counter

ions and placed in a 13Å octahedral box of the cosolvent system. After a 3000-step minimization, a series of equilibration protocols were used as follows:

1. 50 ps, NVT, temperature changes from 0 to 298 K, all heavy atoms of proteins were constrained with a 5 kcal/mol/Å^2 harmonic force constant.
2. 50 ps, NPT, temperature changes from 298 to 350 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
3. 50 ps, NPT, temperature changes from 400 to 450 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
4. 50 ps, NPT, temperature changes from 450 to 500 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
5. 50 ps, NPT, temperature changes from 500 to 550 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
6. 100 ps, NPT, temperature kept at 550 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
7. 50 ps, NPT, temperature changes from 550 to 425 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
8. 50 ps, NPT, temperature changes from 425 to 298 K, all heavy atoms of proteins were constrained with a 1 kcal/mol/Å^2 harmonic force constant.
9. 1 ns, NPT, temperature kept at 298K.

The Procheck program[9-10] was used to examine the backbone dihedral angles of the protein conformations after step 8. The corresponding Ramachandran plots for the apo- and three holo-Bcl-xL conformations were shown in Figure S6. More than 89 % of the amino acids in Bcl-xL have dihedral angles in favored regions except Ser23 and Trp24

(S20 and W21 in Figure S6) in the apo-Bcl-xL. Both amino acids are at the end of NME-cap and located at the back side of the binding site. The deviations of their backbone dihedral angles from favored regions may be attributed to the unresolved loop segment in the crystal structures which were not included in our simulations. A production simulation of 32 ns at 298K was performed. Snapshots of the whole system were taken at intervals of 2 ps.

The hotspot analyses are similar to those reported by Barril et. al.[11] but with our own implementation.[12] They are as follows. Conformations of Bcl-xL from the 32 ns simulation were aligned first along the α5 helix and the cosolvents were imaged into an octahedral box. Evenly spaced 0.5Å grids covering the entire Bcl-xL molecule were created and the counts of the probe atoms in isopropanol (terminal carbon or hydroxyl oxygen atoms) occupying each grid point were calculated. These calculations were performed using the ptraj program from Amber suite. After normalizing the counts of the probe atoms occupying each grid point (N_p), an empirical formula:

$$\Delta G^{CM} = - kT \log (N_p/N_0), \quad (S1)$$

where k is the Boltzman constant, T is room temperature and N_0 is the counts of the probe atoms occupying any grid point (expected occupancy) in 20% v/v cosolvent box without a protein. The N_0 values for terminal carbon and hydroxyl oxygen atoms were provided by Dr. Barril. Based on the values of ΔG, only the grids calculated to be lower than -0.83 kcal/mol were kept for the next analysis. A search procedure was used to find the grid point with the lowest ΔG value and grid points within 1.4 Å of this grid point were removed. The same procedure continued until all grid points were visited. Results of this procedure give *pseudo* atoms in a space. A cluster analysis based on a Depth First

Search (DFS) algorithm with a criterion of two *pseudo* atoms being connected within 2.5 Å was implemented in a C++ program to generate chemical graphs. In analogy to the graph theory, pseudo atoms are vertices and bonds are edges. These chemical graphs form the bases of hotspots probes and were reported in Figure 3 and 4.

Conformational free energy calculations

Bcl-xL conformations were extracted every 100 ps from the simulations in either water or cosolvent. The conformational free energy of Bcl-xL was calculated using the MM-PBSA method with the equation,

$$\Delta G^{MM\text{-}PBSA} = E_{MM} + G_{solv} - TS, \qquad \text{(S2)}$$

where E_{MM} is the molecular mechanics energy, G_{solv} is the solvation free energy and S is the entropy for the protein. The PBradii was set to mbondi2 in the topology file for the solvation free energy calculation. In the normal mode calculations, a distance-dependent dielectric constant $\varepsilon = 4r$ was used, the maximum cycle was set to 60,000, and the convergence tolerance was 0.0002 kcal $mol^{-1}Å^{-1}$.

Figure S1 Sequence alignment of the BH3 domain of BH3 proteins that bind with the Bcl-2 family of proteins. Key hydrophobic residues are denoted h1-4 and p1 is the conserved polar residue. These five residues contribute predominantly to the binding between the BH3 domain peptides and Bcl-2 proteins

Protein								h1				h2			h3		p1		h4						
mBIM		D	L	R	P	E	I	R	I	A	Q	E	L	R	R	I	G	D	E	F	N	E	T	Y	T	R	R
Bim	81-106	D	M	R	P	E	I	W	I	A	Q	E	L	R	R	I	G	D	E	F	N	A	Y	Y	A	R	R
Puma	130-155	E	E	Q	W	A	R	E	I	G	A	Q	L	R	R	M	A	D	D	L	N	A	Q	Y	E	R	R
mBmf	128-151		H	R	A	E	V	Q	I	A	R	K	L	Q	C	I	A	D	Q	F	H	R	L	H	T	Q	
Bad	103-128	N	L	W	A	A	Q	R	Y	G	R	E	L	R	R	M	S	D	E	F	V	D	S	F	K	K	G
Bik	51-75		M	E	G	S	D	A	L	A	L	R	L	A	G	I	G	D	E	M	D	V	S	L	R	A	P
Hrk	26-51	R	S	S	A	A	Q	L	T	A	A	R	L	K	A	I	G	D	E	L	H	Q	R	T	M	W	R
Bid	81-104			D	I	I	R	N	I	A	R	H	L	A	Q	V	G	D	S	M	D	R	S	I	P	P	G
Noxa	18-43	P	A	E	L	E	V	E	C	A	T	Q	L	R	R	F	G	D	K	L	N	F	R	Q	K	L	L
Beclin-1	105-130	G	S	G	T	M	E	N	L	S	R	R	L	K	V	T	G	D	L	F	D	I	M	S	G	Q	T

Figure S2. Number convention of the helices in Bcl-xL.

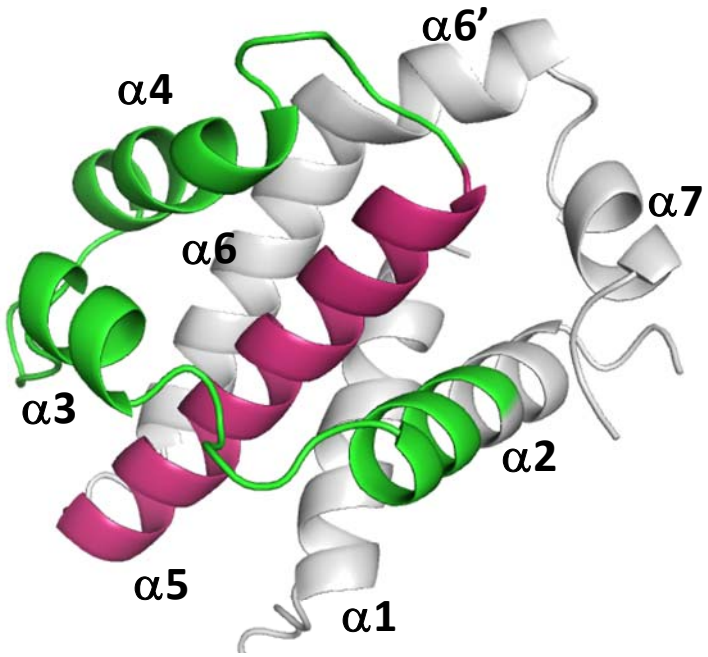


Figure S3. Calculated backbone RMSD of the Bcl-xL conformations simulated in the water (black) and the cosolvent (red). The reference conformation is the crystal structure of apo Bcl-xL. Alignment of the α5 helix (seq. W137-K157) was undertaken before calculating the RMSD of the α5 helix (B) and the α3 plus α4 (seq. A93-W137) (A). (C) is the RMSD of all amino acids after alignment. The total free energy (ΔG) and the accessible surface area (ASA) calculated using the MM-PBSA method are in (D) and (E). The units used are kcal/mol.

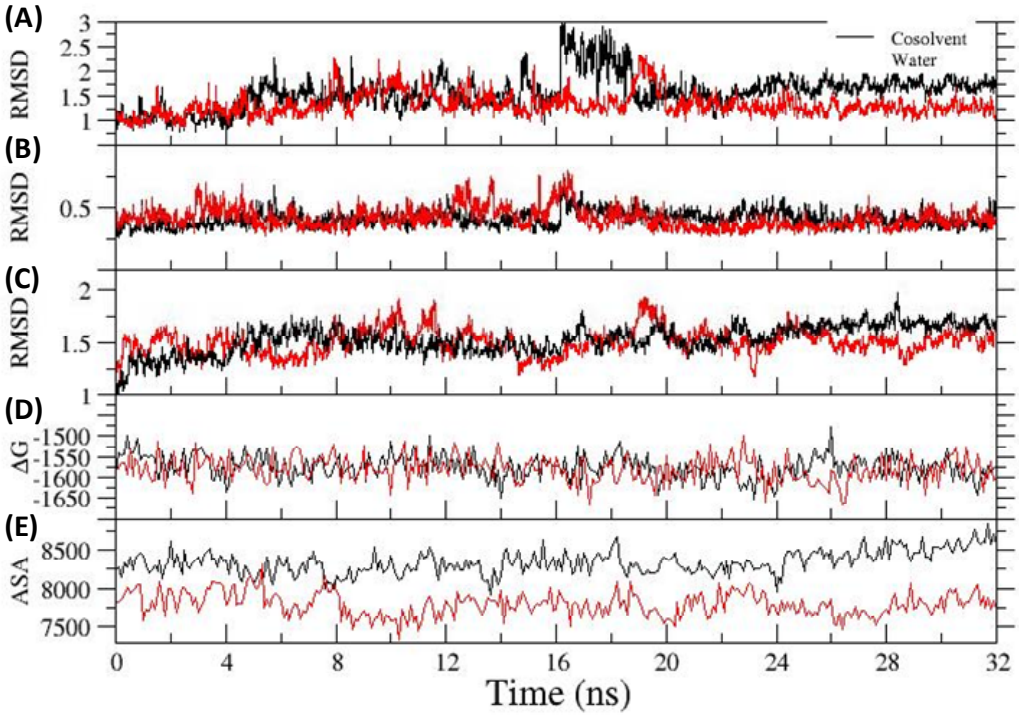


Figure S4. (A) Comparison of the conformation of the crystal structure (grey) and snapshot of Bcl-xL at 32 ns in pure water (cyan). The surface representation of Bcl-xL (B) at 32 ns in pure water simulations; (C) at 16 ns and (D) at 32 ns in the cosolvent simulation. Bim BH3 peptide was aligned and residues interacting with Bcl-xL were shown by the green stick model.

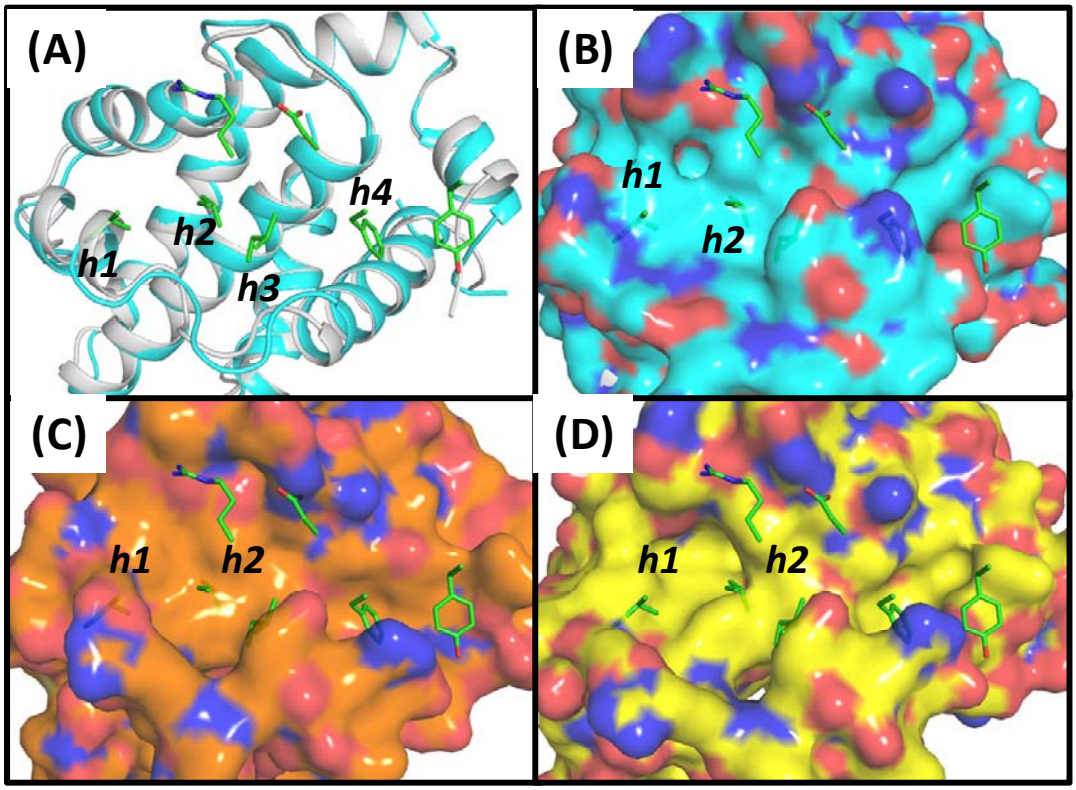


Figure S5. Accessible surface area differences (ΔASA) of hydrophobic residues (including Tyr) around the binding site of Bcl-xL. (A) Between the Bim-(blue), Bad-(red), Bec1- (green) bound and Apo Bcl-xL crystal structures. (B-C) Between conformations obtained from 32 ns cosolvent and aqueous simulations. Positive ΔASA corresponds to exposure whereas negative buried. α2-α5 helices are labeled on top.

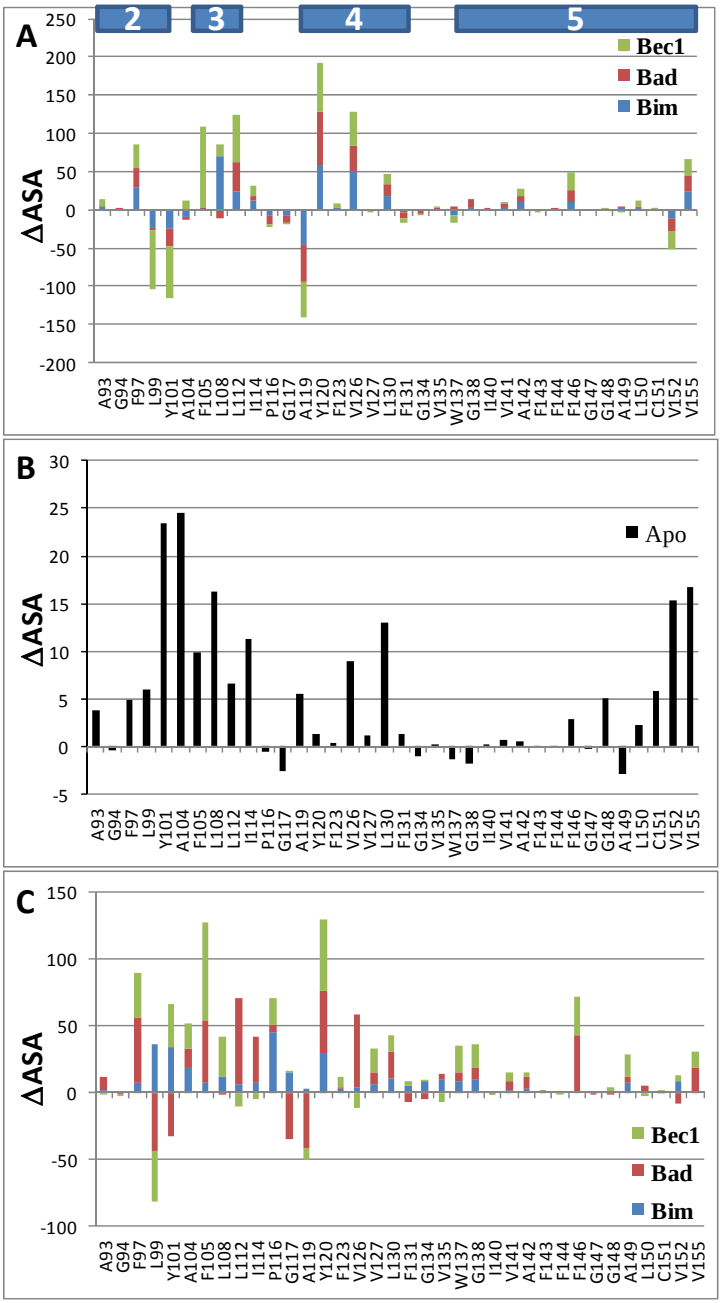
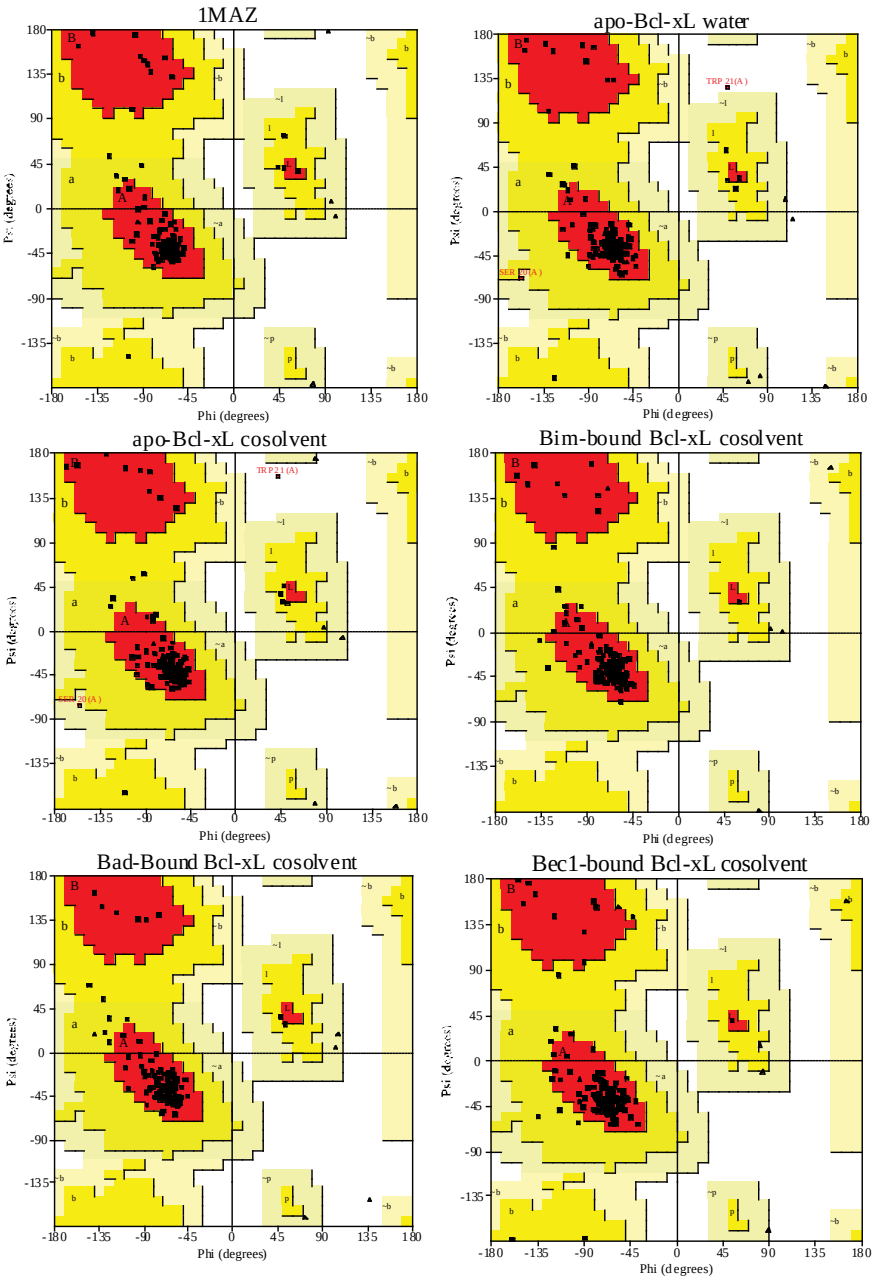


Figure S6. Ramachandran plots using the apo-Bcl-xL in the crystal structure, the apo-Bcl-xL conformation in the water environment, the apo-Bcl-xL and three holo-Bcl-xL conformations in the cosolvent environment after step 8 in the simulation preparation.



Conformations	1MAZ	apo-Bcl-xL (w)	apo-Bcl-xL (c)	Bim-bound (c)	Bad-bound(c)	Bec1-bound (c)
Residues in most favored regions	94.50%	89.30%	90.20%	93.40%	92.60%	90.20%
Residues in allowed regions	5.50%	9.00%	8.20%	6.60%	7.40%	9.80%
Residues in generous allowed regions		0.80%	0.80%			
Residues in disallowed regions		0.80%	0.80%			

Figure S7. Backbone RMSD of Bim-, Bad-, Bec1-bound Bcl-xL simulated in the water (black line) and the cosolvent (red) environment. The crystal structures of the peptide-bound Bcl-xL were used as the reference structures in each case when calculating backbone RMSD.

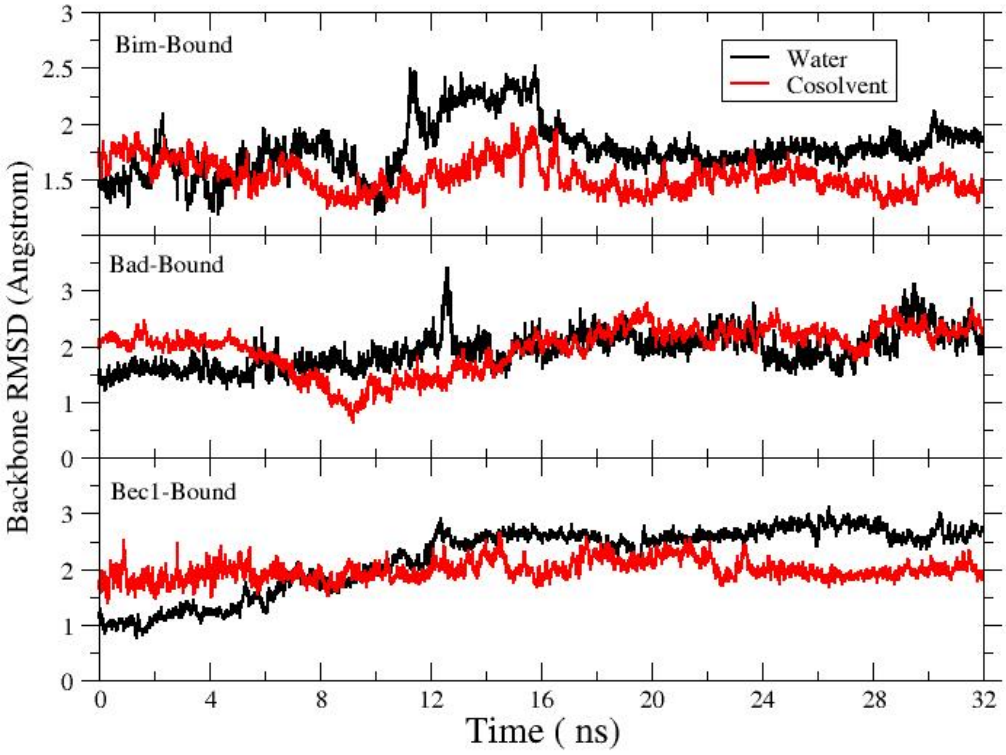


Table S1. Resolutions of the crystal structures used in this work.

PDB ID	Resolution (Å)
1MAZ	2.2
3FDL	1.78
2BZW	2.3
2P1L	2.5
2YXJ	2.2
3INQ	2

References:

(1) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucl. Acids Res.* **2000**, *28*, 235.

(2) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704.

(3) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins* **2008**, *73*, 765.

(4) Case, D. A.; Darden, T. A.; T.E. Cheatham, I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; B.Wang; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; K.F.Wong; Paesani, F.; Vanicek, J.; X.Wu; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A.; AMBER10, University of California, San Francisco: 2008.

(5) Wang, J.; Cieplak, P.; Peter A. Kollman *Journal of Computational Chemistry* **2000**, *21*, 1049.

(6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *The Journal of Chemical Physics* **1983**, *79*, 926.

(7) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Computational Physics* **1977**, *23*, 327.

(8) Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993**, *98*, 10089.

(9) Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J Biomol NMR* **1996**, *8*, 477.

(10) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *Journal of Applied Crystallography* **1993**, *26*, 283.

(11) Seco, J.; Luque, F. J.; Barril, X. *Journal of Medicinal Chemistry* **2009**, *52*, 2363.

(12) Yang, C.-Y.; Wang, S. *ACS Medicinal Chemistry Letters* **2010**, *1*, 125.