

Supporting Information

Lee et al. 10.1073/pnas.1100652108

SI Materials and Methods

Phylogenetic Analysis. Bcl-2 multi-BH domain and BH3-only protein sequences of human, mouse (*Mus musculus*), chicken (*Gallus gallus*), frog (*Xenopus tropicalis*), fish (*Danio rerio*), insect (*Drosophila melanogaster*), hydra (*Hydra magnipapillata*), sea anemone (*Nematostella vectensis*), planaria (*Schmidtea mediterranea*), and sponge (*Geodia cydonium* and *Suberites domuncula*) were obtained from Bcl2DB (1) or Swissprot database.

For multi-BH domain proteins, the BH3, BH1, and BH2 domains of the sequences were identified by aligning them to a hidden Markov Model (HMM) using hmalign in the HMMER3 package (2). The HMM was built with BH3, BH1, and BH2 profiles from a seed of Pfam domain (PF00452, 143aa). The alignment was inspected, and for each species and each type of protein, only one major isoform was kept. Phylogenetic analyses were then performed on the conserved core of the alignment (1). Amino acid sites where gaps exist in the alignments were excluded from the calculation. A neighbor-joining (NJ) tree was then built with Poisson distance, and a maximum-likelihood tree was built with JTT distance (data not shown). The trees and bootstrap analyses were performed by using the SeaView package (3).

For BH3-only proteins, we identified the BH3 domain of the sequences by aligning them to a HMM built with the pattern of the BH3 motif signature (entry PS01259, 15aa) from the PROSITE database. The alignment was inspected and manually adjusted based on literature reports (4, 5). A NJ tree was then built with Poisson distance, and a maximum-likelihood (ML) tree was built with JTT distance. The trees and bootstrap analyses were also performed by using the SeaView package (3) and displayed with FigTree software (<http://tree.bio.ed.ac.uk/software/figtree/>).

Cell-Killing Assays. Retroviral expression constructs were made by using the pMIG vector (MSCV-IRES-GFP; GFP sequence is that of EGFP) as described (6, 7), or the modified vectors pMIH or pMICH in which the GFP cassette is replaced by one encoding hygromycin B resistance or mCherry, respectively. Viruses were produced and used to infect SV40 large T-antigen mouse embryonic fibroblasts as described (8). Colony formation assays in which the effect of expression of a single construct (in pMIG) was

examined as described in Czabotar et al. (9). In assays examining the prosurvival function of sjA, wild-type or *bax^{-/-}bak^{-/-}* MEFs in which sjA was stably expressed (from a pMIH construct) were superinfected with alternative constructs in pMIG, then GFP^{+ve} cells sorted and colonies scored as described (9). In the tripartite assays, sjA stably expressing cells were coinfecting with retroviruses expressing both sjB or vector (in pMIG) and a BH3-only protein (in pMICH). For these assays, GFP^{+ve}/mCherry^{+ve} cells were sorted and colonies scored.

Crystallization and Structure Determination. Crystallization trials were performed by using recombinant sjAΔC41 protein in complex with a human Bak BH3 peptide (PSSTMGQVGRQLAIIGDD-INRRYDSE) mixed in a ratio of 1:1.3 and concentrated to 10 mg/mL. Well-diffracting crystals were grown in hanging drops at 22 °C with a reservoir solution of 1 M trisodium citrate and 0.1 M CHES at pH 9.0. Before flash freezing in liquid N₂, crystals were equilibrated into reservoir solution containing 15% (vol/vol) ethylene glycol as cryoprotectant. Because no structural solution could be obtained by molecular replacement using previously solved structures of prosurvival proteins in complex with BH3 peptides as search models, the crystals were derivatized by soaking with 0.5 mM ethylmercuryphosphate for 72 h. A two-wavelength MAD dataset (inflection and high energy remote of L-III edge) for the Hg derivative was collected (Australian Synchrotron MX2 beamline), in addition to a higher-resolution native dataset. Data were integrated by using XDS (10) and scaled using the program SCALA (part of the CCP4 package) (11). The data were scaled together, and anomalous and dispersive differences were extracted by using SHELXC (12). SHELXD (12) was used to locate and refine two Hg sites per asymmetric unit (one per monomer). Initial phasing was carried out in SHELXE (12) and density-modification procedures in SHELXE were used to extend the phases to the resolution limit of the native dataset. The resulting electron-density map was readily interpretable. Several rounds of building in COOT (13) and refinement in REFMAC5 (14) and PHENIX (15) led to the final model. The X-ray data collection and refinement statistics are summarized in Table S1.

1. Aouacheria A, Brunet F, Gouy M (2005) Phylogenomics of life-or-death switches in multicellular animals: Bcl-2, BH3-Only, and Bnip families of apoptotic regulators. *Mol Biol Evol* 22:2395–2416.
2. Eddy SR (2009) A new generation of homology search tools based on probabilistic inference. *Genome Inform* 23:205–211.
3. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221–224.
4. Bouillet P, Strasser A (2002) BH3-only proteins - evolutionarily conserved proapoptotic Bcl-2 family members essential for initiating programmed cell death. *J Cell Sci* 115: 1567–1574.
5. Lasi M, et al. (2010) The molecular cell death machinery in the simple cnidarian Hydra includes an expanded caspase family and pro- and anti-apoptotic Bcl-2 proteins. *Cell Res* 20:812–825.
6. Chen L, et al. (2005) Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Mol Cell* 17:393–403.
7. Van Parijs L, Refaeli Y, Abbas AK, Baltimore D (1999) Autoimmunity as a consequence of retrovirus-mediated expression of C-FLIP in lymphocytes. *Immunity*, 11:763–770, and retraction (2009) 30:612.
8. Willis SN, et al. (2005) Proapoptotic Bak is sequestered by Mcl-1 and Bcl-xL, but not Bcl-2, until displaced by BH3-only proteins. *Genes Dev* 19:1294–1305.
9. Czabotar P.E., et al. (2011) Mutations to Bax beyond the BH3 domain disrupts interactions with pro-survival proteins and promotes apoptosis. *J Biol Chem* 286:7123–7131.
10. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
11. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62:72–82.
12. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64:112–122.
13. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
14. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
15. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58:1948–1954.

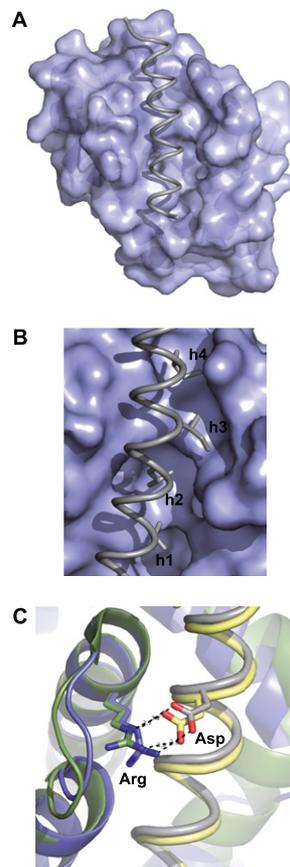


Fig. S5. BH3 peptides bind sjA similarly to the manner in which they bind mammalian prosurvival proteins. (A) Crystal structure of sjA in complex with BakBH3 peptide (PDB ID code 3QBR). The BakBH3 peptide (gray) forms a helix that binds into the canonical hydrophobic groove on the surface of sjA (blue). (B) The four BH3 domain-defining hydrophobic residues (h1–h4) project into hydrophobic pockets on the surface of sjA. (C) The conserved Asp on BakBH3 (gray) forms an electrostatic interaction with the conserved arginine in the BH1 domain of sjA (blue), similar to that seen in the BimBH3:Bcl-x_L complex (yellow:green; PDB ID code 3FDL).

Table S1. X-ray data collection and refinement statistics

Crystal	Native	HgHR	HgIP
Data collection			
Space group	I23	I23	I23
Cell dimensions			
<i>a</i> , Å	148.63	149.18	148.67
Oscillation range, degrees	120	200	180
Wavelength, Å	1.009234	0.991874	1.009234
Resolution, Å	2.6 (2.74–2.6)	3.4 (3.58–3.40)	3.4 (3.58–3.40)
Total observations	247340 (35418)	187311 (27509)	165848 (24419)
Unique	16951 (2461)	7735 (1113)	7670 (1101)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.069 (0.946)	0.159 (0.832)	0.155 (0.746)
<i>I</i> / <i>σ</i>	33.4 (3.2)	27.5 (5.2)	24.9 (5.2)
Completeness, %	99.9 (100)	99.9 (100)	99.9 (100)
Anomalous completeness	—	99.8 (100)	99.8 (100)
Multiplicity	14.6 (14.4)	24.2 (24.7)	21.6 (22.2)
Anomalous multiplicity	—	12.6 (12.6)	11.1 (11.2)
Refinement			
Resolution, Å	29.15–2.60 (2.68–2.60)		
No. of reflections	16303 (1095)		
<i>R</i> _{work}	0.171 (0.216)		
<i>R</i> _{free}	0.234 (0.307)		
No. of atoms			
Protein	2941		
Water	146		
Other solvent	23		
rms deviations			
Bond lengths, Å	0.009		
Bond angles, °	1.127		
Chiral, Å ³	0.072		

Values in parentheses are for highest-resolution shell.