

Prediction of a Putative Functional Region in the Human Bax Protein by Computational Analysis

María Sánchez-Aguilar, Laurence A. Marchat and Absalom Zamorano
Institutional Program of Molecular Biomedicine, ENMH-IPN, Guillermo Massieu Helguera # 239 Fracc.
“La Escalera” Ticomán CP 07320 Mexico City, Mexico

Abstract: Structural domains are relevant elements for interactions between Bcl-2 family members to regulate apoptosis and cell survival. Although BH1-BH4 domains can be identified by their sequence and selective occurrence in Bcl-2-family members, structural regions have not been entirely determined yet. The functional residues of Bax, the most representative pro-apoptotic protein of the human Bcl-2 family, remain almost unknown. Here, we identified the human Bax homologues through PSI-Blast analysis. By phylogenetic study, protein sequence multialignment and three-dimensional mapping, we detected the most conserved amino acids in these proteins. Based on these results, we predicted that the human Bax protein has a putative functional region formed by ten relevant residues in BH1 (G103, N106, G108, R109, V111 and A112) and BH2 domains (W151, G157 and W158), as well as outside of them (V95). Interestingly, the structural analysis of this functional region showed that these residues are closely located inside the protein, forming a putative active site. Moreover, this site seems to be protected by the C-terminal end $\alpha 9$ helix that could act as a regulating gate for the access to this region. In addition, the hydrophobic feature of this helix suggests that it could be involved in the insertion into the mitochondrial membrane that is thought to be important for Bcl-2 family members dimerization and activation.

Key words: Bcl-2 family, bioinformatics, structure, apoptosis

INTRODUCTION

Disruption of the balance between pro- and anti-apoptotic signals leads to diseases related to uncontrolled cell proliferation or degeneration.^[1] Bcl-2 family members form a set of proteins classified according to their overexpression effects. Thus, Bcl-2-family is an interesting prototype of a protein-interaction network that induces or inhibits apoptosis. Typically, Bcl-2 proteins have been involved in embryogenesis, tissue repair and immune response.^[2, 3] They all possess at least one of the four Bcl-2 homology domains (BH1 to BH4). For example, pro-apoptotic proteins have the central short BH3 domain, whereas most anti-apoptotic members contain at least BH1 and BH2.^[4, 5] Interestingly, proteins that are the most closely related to Bcl-2, including Bcl-2, Bcl-XL and Bcl-W, have the four BH domains.^[6] Mutagenesis assays have demonstrated that BH1, BH2 and BH3 domains play a relevant role in the homo- and heterodimerization of various Bcl-2 family members.^[7] Regarding the three-dimensional protein structure, the α -helices in BH1, BH2 and BH3 domains form a long hydrophobic cleft to which the amphipathic BH3 domain of another monomer can bind.^[8] On the other hand, BH4 is thought to be important for the structure stabilization of the hydrophobic groove.^[9] Particularly,

it has been suggested an structural explanation for the reported mutations in BH1 (G145A and G145E) and BH2 (W188A) domains of Bcl-2, which eliminate its ability to regulate apoptosis.^[10] However, structural conservation related to specific residues has not been completely described in the entire family.

Bax is a cytoplasmic protein that belongs to the Bcl-2 family. In early steps of apoptosis, Bax is detected in the external membrane of the mitochondria.^[6] Additionally, the homodimerization and heterodimerization with Bcl-2 was mapped in residues 63-73, in the BH3 domain.^[8] Moreover, it has been shown that the interaction of a Bax dimer with the mitochondria induces cell toxicity, suggesting that Bax forms a channel. The three-dimensional structure of a complete Bax monomer has been recently determined. Bax consists in a central hydrophobic helix surrounded by eight amphipathic helices with varying lengths. Furthermore, α -helices 4, 5, 6, 7 and 8 form a groove that is covered by the $\alpha 9$ -helix.^[11, 12]

In the present work, we took advantage of this data to perform a functional characterization of the human Bax protein through a computer-structure approach. By comparison with homologous proteins, we predicted a region with conserved amino acids that are closely located inside the protein. We proposed that this functional region corresponding to BH1 and BH2 domains, could be related to an evolutionary selection

Corresponding Author: Absalom Zamorano, PhD, Mexico City, Tel 5729-6000, ext 55542-62, Fax 55532.

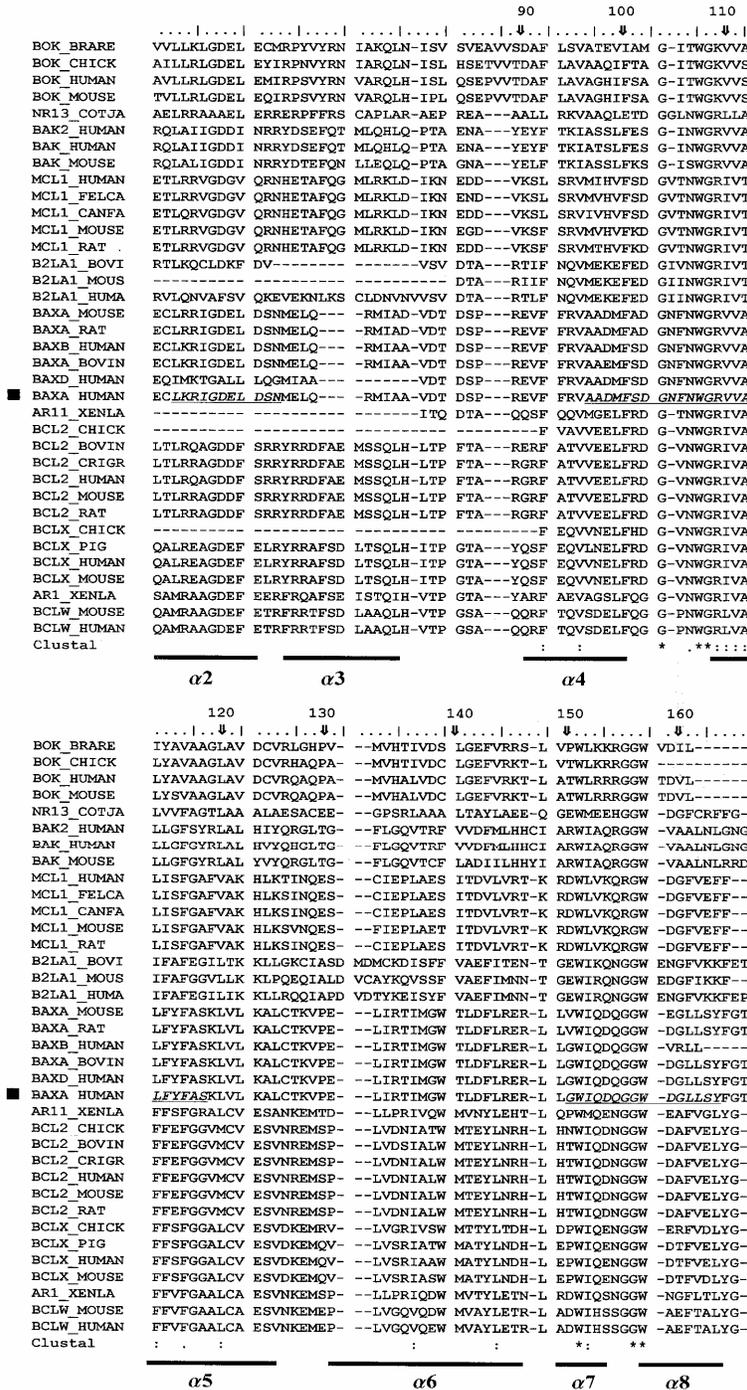


Fig 2: Multialignment of human Bax homologs using the ClustalW program and the substitution BLOSUM62 matrix. Numbers at the top represent the amino acid position related to the first methionine in the human Bax protein indicated by a black square. BH domains are underlined and predicted α -helices are shown by a continuous line. Conserved (star) and substituted (point) residues are indicated.

RESULTS

PSI-BLAST analysis against all Swiss-Prot database led to the detection of 35 homologous proteins that showed low e-values (4e-44 to e-101) and high identity (23-97%) with the human Bax protein. These proteins are from various eukaryotic organisms through evolutionary scale, including mammals (*Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Bos taurus*, *Canis familiaris*, *Felis silvestris catus*, *Sus scrofa* and *Cricetulus griseus*), fishes (*Brachydanio rerio*), amphibians (*Xenopus laevis*) and birds (*Gallus gallus* and *Coturnix coturnix japonica*). All of them are Bcl-2 family members; several proteins belong to the subgroup of pro-apoptotic proteins (Bok, Bak, Bax and MCL1) whereas others are Bcl-2-like survival factors (Bcl-2, BclX, BclW, NR13, Xr1, Xr11 and B2LA1) (Table 1).

The unrooted phylogenetic tree of protein sequences listed in Table 1, including the human Bax protein, was constructed by applying the Neighbor-Joining method that is a simplified version of the minimum evolution method and is considered as the adequate algorithm for the characterization of a small set of sequences. Phylogenetic inference of human Bax homologues showed a clear separation between pro-(Bok, Bak, Bax and MCL1) and anti-apoptotic (Bcl-2, BclX, BclW, NR13, Xr1, Xr11 and B2LA1) factors (Fig 1). The human Bax protein was evolutionary related to bovine, mouse and rat homologues that all emerged from a common Operational Taxonomic Unit (OTU). Moreover, the Bax OTU was as compact as the Bcl-2 OTU, exhibiting the shortest branches of the phylogenetic tree. A similar observation could be made for MCL1 proteins that form a compact UTO closely related to the Bax cluster. However, the other members of the Bcl-2 family, including Bok, Bak, B2LA1, BclW and BclX, emerged from longer branches, suggesting that they are more divergent (Fig 1).

In order to carry out an extensive comparative study of the amino acids sequence of the human Bax protein and its homologues, we aligned the whole data set using the ClustalW algorithm. Visualization of the multialignment allowed the identification of the two conserved BH1 and BH2 domains in regions that were predicted as α -helices in the human Bax protein^[21]: i) BH1 domain (α 4 and α 5 helices): (F/L)-X₂-(V/I)-X₇-G-X₁₋₂-(N/S/T)-W-G-(R/K)-(L/I/V)-(L/V)-(A/T/S)-(F/I/L)-X₂-(F/V)-X₃-(L/M/V), and ii) BH2 domain (α 7 and α 8 helices): W-(I/M/L)-X₄-GW, where X is any amino acid. Additionally, we identified a conserved region in the α 6 helix: (V/I/L)-X₇-(L/I/M/V). A detailed examination of identical and substituted amino acids in the multialignment showed that conserved residues in all these proteins were Glycine (G103, G108 and G157) and Tryptophan (W107, W151 and W158),

which are both non polar amino acids (the amino acid position is related to the human Bax protein) (Fig 2). Interestingly, the majority of substituted residues also involved non polar amino acids, such as: (F/L)92, (V/I)95 (L/I/V)110, (L/V)111, (F/I/L)113, (F/V)116, (L/M/V)120, (V/I/L)136, (L/I/M/V)144 and (I/M/L)152. In contrast, the (N/S/T) substitution at position 106 between α 4 and α 5 helices and the (R/K) substitution at position 109 in the amino terminus of the α 5 helix involve polar residues, whereas (A/T/S)112 substitution in the BH1 domain involves residues with distinct biochemical properties. On the other hand, the BH3 domain and other regions without known function appeared as more variable domains, since no conserved residue was identified using ClustalW (Fig 2).

The evaluation of amino acid changes in Bax homologues was also studied through the ConSurf protocol expressing the conservation score for each residue in a nine grades scale. Results showed that ten amino acids were in the ninth grade, indicating that they are the most conserved residues. Three of them were conserved in all the proteins (G108, W151 and W158) whereas the others were substituted residues including (I/V)95, (G/K)103, (E/N/S/T)106, (K/P/R)109, (I/V)111, (A/S/T)112 and (G/Y)157 (the amino acid position is related to the human Bax protein). Notably, only two substitutions involved non-polar amino acids (Table 2). In accordance with results shown in Figure 2, all these residues are located in conserved BH1 and BH2 domains. In contrast, the BH3 domain and other regions with unknown function did not present any conserved residue.

Table 2: Detection of conserved and substituted amino acids in human Bax homologues through ConSurf analysis

Identical Residue	Substituted Residue
G108	(I / V) 95*
W151	(G / K) 103
W158	(E / N / S / T) 106
	(K / P / R) 109
	(I / V) 111*
	(A / S / T) 112
	(G / Y) 157

Numbers are relative to the amino acid position in the human Bax protein.

*Non polar residue.

To gain insights into the relative position of conserved residues, we analyzed the three-dimensional structure of the human Bax protein using a nine grades colors scale to represent the conservation scores of each amino acid as implemented in ConSurf server. We

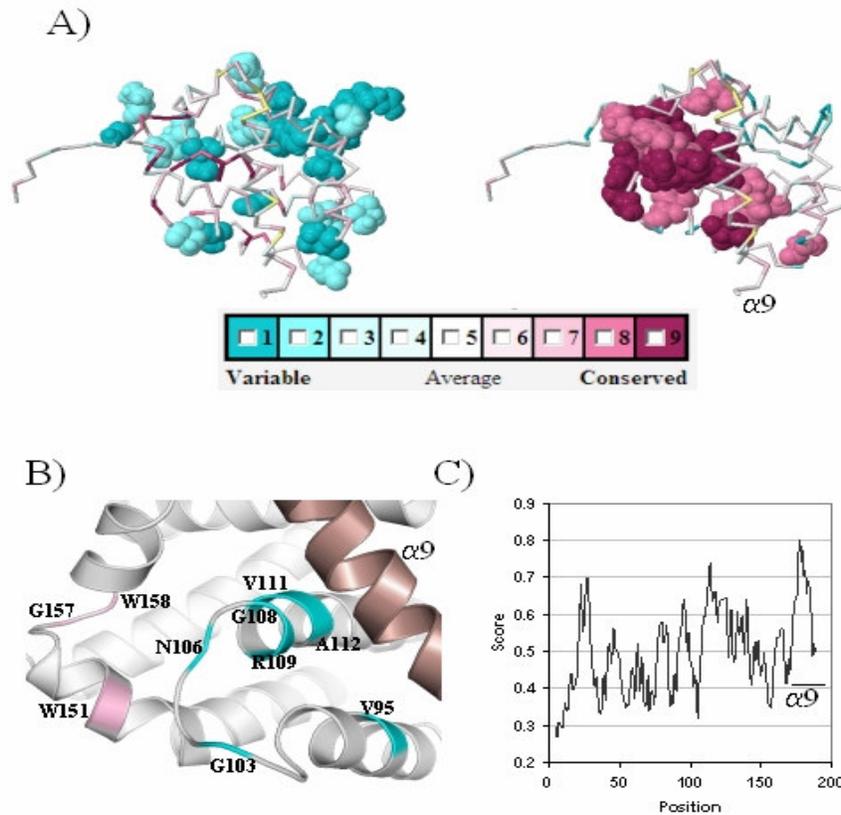


Fig 3: Three-dimensional mapping of the putative active site of the human Bax protein. A) Localization of the conserved amino acids. The human Bax structure was obtained from the PDB file (PDB:1F16)^[9] and the most and less conserved residues were indicated according to the nine grade colored scale shown at the bottom. B) Magnification of the putative functional region of the human Bax protein. C) Hydrophobic profile of the human Bax protein.

focused on the most (eighth and ninth grades) and less (first and second grades) conserved positions (Fig 3A). The most conserved amino acids, including the ten residues described above (Table 2), are spatially close in the central part of the protein. Remarkably, this conserved region seems to be protected by the $\alpha 9$ helix. In contrast, the more variable positions showed a dispersed distribution at the external part of Bax. In the ConSurf methodology, it was assumed that the conservation grade of an amino acid is indicative of its evolution rate. Then, sites that are evolutionary conserved are thought to have evolved slowly indicating their structural and functional importance, while those that change more quickly over evolutionary time are more variable and represent less relevant amino acids. This categorization takes various parameters in consideration, including the phylogenetic

relation of sequences, the amino acids similarity and the empirical Bayesian algorithm.^[13]

To extensively examine this central region of Bax, we measured the distance between the ten most conserved amino acids. The minimal, maximal and mean distance were 3.82 Å (G108-R109), 25.77 Å (V95-G157) and 12.9 Å, respectively (Table 3). The proximity of these residues confirms that they could be involved in the formation of an active site inside the human Bax protein (Fig 3B). Finally, in order to evaluate the functional relationship between this putative active site and the $\alpha 9$ helix that seems to protect it (Fig 3A), hydrophobic profile of the complete Bax protein was determined (Fig 3C). Particularly, $\alpha 9$ -helix is presented as the most hydrophobic region. This suggests that its interaction with mitochondrial membrane could regulate the active site exposure.

Table 3: Distance in Angstrom (Å) between the conserved residues forming the putative active site in the human Bax protein

	V95	G103	N106	G108	R109	V111	A112	W151	G157	W158
V95	0	15.68	16.16	13.41	10.71	14.16	10.91	20.31	25.77	23.92
G103	-	0	7.94	12.10	10.94	16.36	15.84	13.21	20.54	20.33
N106	-	-	0	6.17	6.48	9.92	10.87	10.70	14.91	13.75
G108	-	-	-	0	3.82	5.48	5.45	14.90	17.43	15.00
R109	-	-	-	-	0	5.80	5.13	12.97	16.84	14.75
V111	-	-	-	-	-	0	3.83	15.05	15.59	12.47
A112	-	-	-	-	-	-	0	16.95	18.92	16.03
W151	-	-	-	-	-	-	-	0	9.05	10.44
G157	-	-	-	-	-	-	-	-	0	3.83
W158	-	-	-	-	-	-	-	-	-	0

Numbers in bold represent the minimal and maximal distances

DISCUSSION

In silico analysis of genomes have confirmed that protein domains are fundamental units through evolution because they duplicate and mix in different ways. Besides, several domain families seem to be versatile while others are restricted to specific organisms or kingdoms.^[22] The increasing number of structures resolved through X rays or nuclear and magnetic resonance (RMN) suggests that protein structures are more conserved through evolution than amino acid sequences.^[23] Thus, several families are constituted by proteins sharing only 5% identity, while related members of other families exhibit 50% structure identity, mainly in the central part of the proteins.^[24] Moreover, in some cases, it is difficult to predict the presence of a given domain whose size and location are not well defined in the family members.^[22] Therefore, evolutionary analysis of protein families as well as the knowledge of any data about a specific gene or its product, such as genome localization, molecular organization, structure or expression will contribute to better understand the function of a specific protein. Then, this information can be used as a guide for further experimental studies and practical applications.

In the present work, we performed a functional characterization of the human Bax protein through a computer-structure approach, mainly based on the ConSurf protocol.^[13, 25] As expected, the human Bax protein is closely related to other Bcl-2 family members, clustering with pro-apoptotic factors while anti-apoptotic proteins form a separated group in the

phylogenetic tree. A similar observation was reported when the complete Bcl-2 family was phylogenetically characterized.^[26]

Although results obtained from the ConSurf protocol were slightly different from those obtained using the ClustalW algorithm, both strategies identified the BH1 and BH2 domains as the most conserved part of Bax homologues. Moreover, they were in agreement with a previous work in which the evolutionary conservation of BH1 and BH2 domains were used to suggest that conserved residues G108 and W151 are involved in the formation of an active site in Bcl-2 structural homologues, including the human Bax protein.^[10] Remarkably, *in silico* predictions and experimental results also demonstrated that G108 and W151 mutation disrupts the pore forming ability of the Bcl-2 protein^[27] and completely abrogates its heterodimerization and consequently death repressor activity in IL-3 deprivation, gamma irradiation and glucocorticoid-induced apoptosis.^[4, 28, 29] Recent studies by ConSurf analysis with dataset of proteins whose three-dimensional structure was known showed that regions with evolutionarily conserved residues were often functionally important.^[30-33] However, since the ConSurf protocol does not allow to quantify the global conservation of a region formed by conserved residues, it is only considered as a qualitative method.^[34] In the present work, we extended the list of the relevant residues in the active site of the Bax protein, demonstrating the importance of V95, G103, N106, V111, A112, G157 and W158 amino acids, in addition to G108, R109 and W151 residues previously described.^[10] In fact, the mean distance between these

ten amino acids is 12.9 Å, which is close to the 10.56 Å measured between only three residues.^[10] The high conservation of these amino acids in the Bcl-2 family members shows that BH1 and BH2 domains have evolved slowly due to their functional importance for Bcl-2 proteins dimerization and activation. On the other hand, BH3 domain that is also necessary for the dimerization, does not correspond to a conserved region. Finally, the hydrophobic feature of the α 9-helix at the C-terminal end of Bax protein suggests that it could be a transmembranal domain. Based on our results, we proposed an hypothetical working model in which the α 9-helix could be a mobile element responsible for the interaction with phospholipids of the mitochondrial membrane, allowing the exposure of the internal active site and the activation of the Bax protein.

ACKNOWLEDGEMENTS

This project was supported by SIP-IPN (20060798) and COFAA-IPN, Mexico.

REFERENCES

1. Barr, P.J. and L.D. Tomei, 1994. Apoptosis and its role in human disease. *Biotechnology*, 12(5): 487-93.
2. Bouillet, P., et al., 1999. Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis and to preclude autoimmunity. *Science*, 286(5445): 1735-8.
3. Lindsten, T., et al., 2000. The combined functions of proapoptotic Bcl-2 family members Bak and Bax are essential for normal development of multiple tissues. *Mol Cell*, 6(6): 1389-99.
4. Borner, C., et al., 1994. Dissection of functional domains in Bcl-2 alpha by site-directed mutagenesis. *Biochem Cell Biol*, 72(11-12): 463-9.
5. Gross, A., et al., 1998. Enforced dimerization of Bax results in its translocation, mitochondrial dysfunction and apoptosis. *EMBO J*, 17(14): 3878-85.
6. Adams, J.M. and C.S. Cory, 1998. The Bcl-2 protein family: arbiters of cell survival. *Science*, 281(5381): 1322-6.
7. Sattler, M., et al., 1997. Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. *Science*, 275(5302): 983-6.
8. Zha, H., et al., 1996. Proapoptotic protein Bax heterodimerizes with Bcl-2 and homodimerizes with Bax via a novel domain (BH3) distinct from BH1 and BH2. *J Biol Chem*, 271(13): 7440-4.
9. Bauchier-Hayes, L., L. Lartigue and D.D. Newmeyer, 2005. Mitochondria: pharmacological manipulation of cell death. *J Clin Invest*, 115(10): 2640-7.
10. Gurudutta, G.U., et al., 2005. Structural conservation of residues in BH1 and BH2 domains of Bcl-2 family proteins. *FEBS Letter*, 579(17): 3503-7.
11. Suzuki, M., R.J. Youle, and N. Tjandra, 2000. Structure of Bax: coregulation of dimer formation and intracellular localization. *Cell*, 103(4): 645-54.
12. Mancinelli, F., et al., 2006. Molecular Dynamics simulation and automated docking of the proapoptotic Bax protein its complex with a peptide designed from the Bax-binding domain of anti-apoptotic Ku70. *J Cell Biochem*, 99(1): 305-18.
13. Landau, M., et al., 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acid Res*, 33(Web Server Issue): W299-W302.
14. Altschul, S.F., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25(17): 3389-402.
15. Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 22(22): 4673-80.
16. Saitou, N. and M. Nei, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*, 4(4): 406-25.
17. Mayrose, I., et al., 2004. Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol Biol Evol*, 21(9): 1781-91.
18. Jones, D.T., W.R. Taylor and J.M. Thornton, 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*, 8(3): 275-82.
19. Kyte, J. and R.F. Doolittle, 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 157(1): 105-32.
20. Gasteiger, E., et al., 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: *The Proteomics Protocols Handbook* (ed. JM. Walker) pp 571-607. Humana Press.
21. Petros, A.M., E.T. Olejniczak and S.W. Fesik, 2004. Structural biology of the Bcl-2 family of proteins. *Biochim et Biophys Acta*, 1644(2-3): 83-94.

22. Orengo, C.A. and J.M. Thornton, 2005. Protein families and their evolution—a structural perspective. *Annu Rev Biochem*, 74: 867-900.
23. Chotia, C. and A.M. Lesk, 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J*, 5(4): 823-6.
24. Redfern, O., C. Bennet and C.A. Orengo, 2003. *Bioinformatics Encyclopedia*. R. Apweiler, New York, Wiley.
25. Glaser, F., et al., 2003. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. *Bioinformatics*, 19(1): 163-4.
26. Lanave, C., M. Santamaria, and C. Saccone, 2004. Comparative genomics: the evolutionary history of the Bcl-2 family. *Gene*, 333: 71-9.
27. Baffy, G., et al., 1993. Apoptosis induced by withdrawal of interleukin-3 (IL-3) from an IL-3-dependent hematopoietic cell line is associated with repartitioning of intracellular calcium and is blocked by enforced Bcl-2 oncoprotein production. *Biol Chem*, 268(9): 6511-9.
28. Yin, X.M., Z.N. Oltvai and S.J. Korsmeyer, 1994. BH1 and BH2 domains of Bcl-2 are required for inhibition of apoptosis and heterodimerization with Bax. *Nature*, 369(6478): 321-3.
29. Diaz, J.L., et al., 1997. A common binding site mediates heterodimerization and homodimerization of Bcl-2 family members. *J Biol Chem*, 272 (17): 11350-5.
30. Aloy, P., et al., 2001. Automated structure-based prediction of functional sites in proteins: applications to assessing the validity of inheriting protein function from homology in genome annotation and to protein docking. *J Mol Biol*, 311(2): 395-408.
31. Landgraf, R., I. Xenarios, and D. Eisenberg, 2001. Three-dimensional cluster analysis identifies interfaces and functional residue clusters in proteins. *J Mol Biol*, 307(5): 1487-502.
32. Valdar, W.S. and J.M. Thornton, 2001. Conservation helps to identify biologically relevant crystal contacts. *J Mol Biol*, 313(2): 399-416.
33. Lichtarge, O. and M.E. Sowa, 2002. Evolutionary predictions of binding surfaces and interactions. *Curr Opin Struct Biol*, 12(1): 21-7.
34. Ebner, S., N. Sharon and N. Ben-Tal, 2003. Evolutionary analysis reveals collective properties and specificity in the C-type lectin and lectin-like domain superfamily. *Proteins*, 53(1): 44-55.