2574 *Biochemistry* **2010,** *49,* 2574–2584
DOI: 10.1021/bi100029k

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Identification of Core Structural Residues in the Sequentially Diverse and Structurally Homologous Bcl-2 Family of Proteins†

Dilraj Lama and Ramasubbu Sankararamakrishnan*

Department of Biological Sciences and Bioengineering, Indian Institute of Technology-Kanpur, Kanpur 208016, India

Received November 4, 2009

ABSTRACT: The Bcl-2 family of proteins regulates the intrinsic pathway of apoptosis and plays a significant role in mitochondrial outer membrane permeabilization. Bcl-2 homologues belonging to both anti- and pro-apoptotic classes have been identified in diverse organisms. While anti-apoptotic Bcl-2 proteins possess up to four BH sequence domains (BH1−BH4), the pro-apoptotic counterparts have either three BH (BH1−BH3) domains or only the BH3 domain. Many anti-apoptotic viral homologues do not seem to have any detectable BH homology regions and exhibit a very low level of sequence identity with other Bcl-2 family members. However, structures determined for several Bcl-2 anti- and pro-apoptotic proteins and their viral homologues show a remarkably conserved helical fold characterized by a central hydrophobic helix surrounded by five or six amphipathic helices. In this study, we have analyzed 16 nonredundant Bcl-2 structures from human, mouse, *Caenorhabditis elegans*, and five different viral species. While the length of the central hydrophobic helix is preserved in all the Bcl-2 structures, variations in length are observed for other helices. We performed multiple-structure alignment of all 16 structures. Eighty structurally equivalent positions, the bulk of them in the helical regions, constituted the ungapped blocks in the structure-based sequence alignment. Analysis of helix bundle geometry indicates that helix−helix packing differed in different Bcl-2 structures. This is presumably to accommodate disparate residue substitutions. Residue properties such as solvent accessibility, conservation of chemical nature, and/or size and involvement in interhelical interactions were analyzed in each position of the ungapped alignment regions. A sequence motif made up of small amino acids has been detected in the central helix that is proposed to be important for helix−helix association. We have found that residues in 22 positions in the helical regions are buried, exhibit conservation in hydrophobicity and/or size, and participate in interhelical interactions in at least 12 of the 16 structures studied. We also found 15 additional positions in which residues exhibit two of the three properties investigated. We suggest that these positions constitute the important structural core in the diverse Bcl-2 family members and could play a significant role in the folding of the protein. Results of our studies have been used in the identification of three putative Bcl-2 homologues from three different viral organisms. This study will help in the genome-wide identification of hitherto unrecognized Bcl-2 family members, especially in viral genomes.

The Bcl-2 family[1] (B cell leukemia/lymphoma 2 family) of proteins is involved in the regulation of an evolutionarily conserved physiological process of cell death termed apoptosis that is critical during development, maintenance of tissue homeostasis, and protection against pathogens (*1*, *2*). Homologues of Bcl-2 proteins are found in mammals, birds, fish, and amphibians and also in invertebrates such as nematodes, fruit flies, and several types of animal viruses (*3*). The protein family comprises both pro- and anti-apoptotic members that can either promote cell death or help in cell survival. Interaction among these opposing members is a crucial event by which the proteins regulate apoptosis (*4*, *5*). Since the first structure of an anti-apoptotic member protein Bcl-X_L was published (*6*), several structures of Bcl-2 proteins have been determined by X-ray crystallography or NMR (*7*). Knowledge of Bcl-2 structures has given us valuable insights into their fold, molecular mechanism, and mode of interaction of pro- and anti-apoptotic members at the molecular level. At the primary structure level, these proteins share homology only in a small stretch of sequence domains called "Bcl-2 homology" (BH) domains (*8*). There are four such domains (BH1−BH4). The anti-apoptotic class of Bcl-2 proteins contains domains BH1−BH4. The pro-apoptotic Bcl-2 proteins are further divided into two classes on the basis of the presence or absence of BH domains. The "multidomain" pro-apoptotic proteins contain domains BH1−BH3. The "BH3 only" class has members that share sequence homology with other Bcl-2 proteins only in the BH3 domain. Moreover, recent structural studies on viral homologues such as Myxoma virus M11L protein (*9*, *10*) and Vaccinia virus A52, B14, K7, N1L, and F1L proteins (*11*−*14*) indicate a similar fold, although no recognizable BH domains were found in their sequences. These studies further highlight the extent of sequential diversity among the Bcl-2 family members. However, in spite of these sequential differences, all anti-apoptotic proteins, all multidomain pro-apoptotic proteins, at least one member of the BH3 only class of pro-apoptotic proteins, and the Bcl-2 homologues of viral proteins exhibit a very similar α-helical fold. The overall topology

† R.S. is the Joy Gill Chair Professor at Indian Institute of Technology-Kanpur.

*To whom correspondence should be addressed. E-mail: rsankar@iitk.ac.in. Phone: +91 512 2594014. Fax: +91 512 2594010.

[1] Abbreviations: Bcl-2 family, B cell leukemia/lymphoma 2 family; BH domain, Bcl-2 homology domain; MUSTANG, Multiple Structural Alignment Algorithm; NMR, nuclear magnetic resonance; rmsd, root-mean-square deviation; SASA, solvent accessible surface area.

 Published on Web 02/08/2010

Table 1: Set of 16 Nonredundant Bcl-2 Family Protein Structures

	protein	source	PDB entry[a]	Uniprot entry[b]	length[c]	ref
			Anti-Apoptotic Proteins			
1	Bcl-X_L	human	1R2D (1.95 Å)	Q07817	142	*24*
2	Bcl-2	human	1G5M	P10415	164	*28*
3	Bcl-w	human	1MK3	Q92843	170	*22*
4	Mcl-1	mouse	1WSX	P97287	157	*29*
5	A1	mouse	2VOH/A (1.9 Å)	Q07440	150	*30*
6	CED-9	*C. elegans*	1OHU/B (2.03 Å)	P41958	168	*31*
			Pro-Apoptotic Proteins			
7	Bax	human	1F16	Q07812	192	*32*
8	Bak	human	2IMS (1.48 Å)	Q16611	163	*33*
9	Bid	human	2BID	P55957	195	*34*
			Bcl-2 Viral Homologues			
10	Bcl-2 homologue	Kaposi Sarcoma virus	1K3K	P90504	146	*35*
11	BHRF1	Epstein-Barr virus	1Q59	P03182	160	*36*
12	M11	γ herpes virus 68	2ABO	P89884	131	*37*
13	M11L	Myxoma virus	2O42/A (2.91 Å)	Q85295	138	*9*
14	N1L	Vaccinia virus	2I39/A (2.2 Å)	P21054	115	*13*
15	A52	Vaccinia virus	2VVW/A (1.9 Å)	Q01220	150	*11*
16	B14	Vaccinia virus	2VVY/A (2.69 Å)	P24772	142	*11*

[a]Four-letter unique Protein Data Bank (*21*) entry. If the structure is determined by X-ray, then the resolution of the structure is given in parentheses. If more than one polypeptide chain is present, then the chain considered for this study is given after the PDB entry. [b]UniProt (*38*) sequence entry. [c]The length of the polypeptide chain for which the structure is determined as reported in the respective PDB entry.

consists of a central hydrophobic helix surrounded by five or six amphipathic helices.

Protein structures are known to be evolutionarily more conserved than their sequences (*15*−*20*). In the case of Bcl-2 family proteins, amino acid sequence comparison shows that there is little sequence similarity between many members, although they adopt a remarkably similar fold. Examples of proteins with diverse sequences assuming the same helix bundle (*19*) or sheet structure (*20*) have been reported for other families of proteins. In a quest to find a rationale for the preservation of structural fold in these families of proteins, knowledge of structures has been used to find some conserved features at the primary sequence level. It must be pointed out that conventional sequence alignment methods will not be useful in these cases.

We can detect important positions in diverse primary sequences having similar three-dimensional folds more efficiently by performing a structural alignment of the homologous proteins. This study aims to find conserved features in a sequentially diverse and structurally similar Bcl-2 family of proteins. A set of 16 distinct Bcl-2 family proteins that have representatives from diverse organisms such as viruses, *Caenorhabditis elegans*, mice, and humans were considered. A multiple-structure alignment was performed, and each structurally equivalent position in the alignment was examined for various features such as chemical nature or size, solvent accessibility, and its involvement in interhelical interactions. This led to the identification of 22 positions in the helical region where all the analyzed features are conserved. We also found additional positions where most of the properties are conserved. On the basis of this analysis, we conclude that these positions could be important for the conserved helical fold observed in the Bcl-2 family of proteins. The results obtained from the structural analysis have been used to identify and confirm three new putative Bcl-2 homologues from three different viral species.

MATERIALS AND METHODS

Bcl-2 Homologues Selected for the Study. As of October 2009, a total of ~70 protein structures belonging to the Bcl-2 family are deposited in the Protein Data Bank (PDB) (*21*). This set is redundant since it contains the same protein structures determined using two different experimental methods (X-ray and NMR) (*6*), the same protein determined by different research groups (*22*, *23*), mutants of the same protein (*24*), or the same protein in apo and complex forms (*25*−*27*). From this pool of Bcl-2 structures, we selected a set of 16 distinct nonredundant proteins, and they were grouped into anti-apoptotic, pro-apoptotic, and Bcl-2 viral homologues as shown in Table 1. The residues in the determined structures which were either histidine tags or due to cloning artifacts were removed. Structures that had selenomethionine as the modified residues were reconstructed as standard methionine residues. For structures that had multiple chains, only one chain was selected for the study as mentioned in Table 1.

Analysis of Bcl-2 Structures. For each Bcl-2 structure, the helix bundle fold was characterized by analyzing the interhelical angles and distances. The helix−helix crossing angle was determined as the angle between the two helix axes. The helix axis for each helix in a Bcl-2 structure was determined using the rotational least-squares method (*39*, *40*). The interhelical distance for each pair of interacting helices was calculated as the shortest line segment connecting each finite helix axis (*41*). For each residue in a Bcl-2 structure, the solvent accessible surface area (SASA) (*42*) was calculated using Vadar, version 1.4 (*43*). A probe radius of 1.4 Å and van der Waals radii of atoms from the work of Chothia (*44*) were used in SASA calculations. Residues involved in interhelical atomic contacts were identified by application of the following criterion. Residues from two different helices are said to interact if there is at least one pair of heavy atoms, one from each residue whose distance is less than the sum of their van

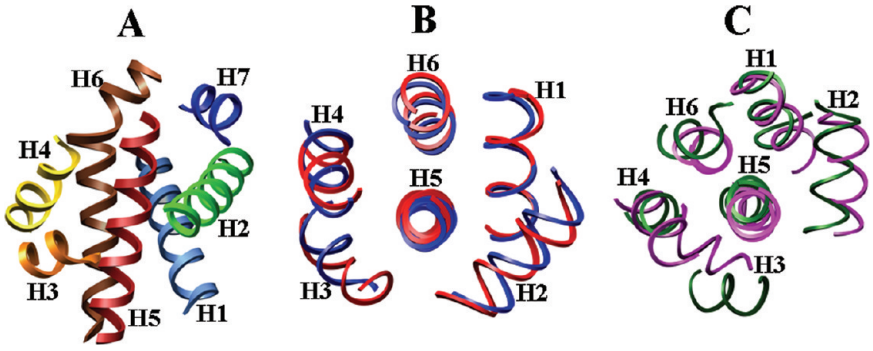

FIGURE 1: (A) Seven helical segments of the Bcl-X_L protein (PDB entry 1R2D), a representative example of the Bcl-2 structural fold. All seven helices are shown in different colors, and this coloring scheme is followed in subsequent figures. For the sake of clarity, loop segments are not displayed. The structure is viewed approximately perpendicular to the helix axis of central helix H5. In panels B and C, pairs of Bcl-2 structures showing the smallest [PDB entries 1R2D (blue) and 1G5M (red)] and the largest [PDB entries 2BID (magenta) and 2O42 (dark green)] rmsd are superposed. The structures are viewed down the helix axis of central hydrophobic helix H5. Only helical segments H1−H6 corresponding to the ungapped alignment blocks (see Figure 3) are shown. All molecular plots were created using the UCSF Chimera package (*69*).

der Waals radii + 0.6 Å. The same criterion was used in earlier studies to find interhelical interactions (*44*−*46*).

A structure-based multiple-sequence alignment for the 16 selected Bcl-2 structures was produced using MUSTANG, version 0.3 (Multiple Structural Alignment Algorithm) (*47*). Although several structural alignment programs are available, we have chosen MUSTANG to conduct the multiple-structure alignment since it has been shown to be more reliable in generating quality alignment on very distantly related protein data sets. Residues in each aligned position, their properties, and the extent of conservation across the diverse Bcl-2 proteins were analyzed to determine common features that could explain the structural conservation of the Bcl-2 helical fold.

RESULTS

Variations in the Lengths of Different Helices in the Bcl-2 Helical Fold. It is necessary to scrutinize all Bcl-2 structures under uniform criteria. As a first step, we have analyzed the Bcl-2 helical fold from diverse family members. The variations observed in the length of different helices were determined via application of a somewhat generous and at the same time uniform φ, ψ criterion as follows. A residue is considered to adopt a helical conformation if its backbone dihedral angles φ and ψ satisfy the criterion (*48*)

$$-140° \leq \varphi \leq -30° \text{ and } -90° \leq \psi \leq 45°$$

A polypeptide segment is considered to be a helix if a continuous stretch of at least four residues is found to be in the helical conformation. The core of the Bcl-2 fold consists of seven major helices as shown in Figure 1A, in which the central hydrophobic helix (H5) is surrounded by other amphipathic helices. Additional C-terminal helices present in some of the Bcl-2 members were not considered for this analysis. Similarly, helix H7 was excluded because of different levels of C-terminal truncation. We have examined all 16 Bcl-2 structures and found that the average lengths (±standard deviation) of helices H1−H6 are 16.8 ± 3.7, 17.5 ± 2.8, 9.8 ± 3.0, 14.5 ± 3.1, 20.1 ± 1.8, and 22.3 ± 5.8 residues, respectively. It is obvious that the central hydrophobic helix (H5) and the succeeding helix (H6) are the longest helices while helix H3 is the shortest. However, the lengths of H5 and H6 show the smallest and largest variations, respectively, among the diverse Bcl-2 family members. The length of H6 is more than 25 residues in all six anti-apoptotic Bcl-2 family members, and it is only ≤15 residues long in some of the viral homologues.

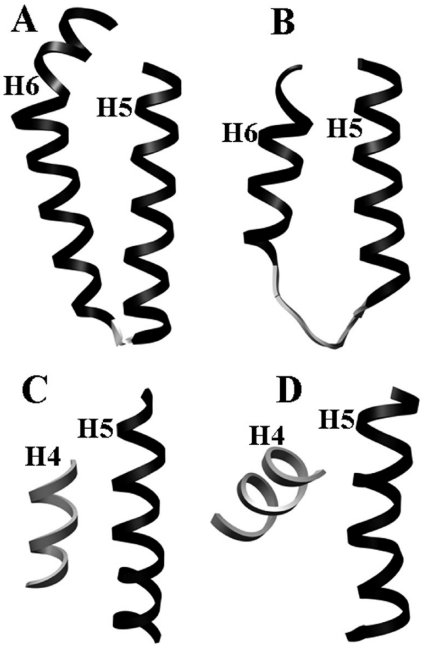

FIGURE 2: (A and B) H5−H6 helix pairs from two Bcl-2 homologues (PDB entries 1G5M and 2I39, respectively) to illustrate the variations in the lengths of helix H6. (C and D) Packing between H4 and H5 helix pairs from two Bcl-2 homologues (PDB entries 1MK3 and 2BID, respectively) is significantly different. The ω angles in the two cases are 177.4° and 71.4°, respectively.

A representative example for each case is shown in panels A and B of Figure 2. The region comprising short helix H3 has been shown to exhibit flexibility in both experimental (*24*) and simulation studies (*49*).

Structure-Based Sequence Alignment of the 16 Bcl-2 Structures. The PDB files of the structures were given as the input for MUSTANG to generate a multiple-structure alignment. Individual helical segments H1−H7 were found to align well across all 16 structures. However, some manual adjustments were required to bring some reported signatures in the BH1−BH3 sequence domains in some of the structures. This includes "GD" and "GW" signatures in helix H2 and the beginning of H7, respectively (*7*). The final structure-based sequence alignment for helical regions is shown in Figure 3 (For details of the entire alignment, see Figure S1 of the Supporting Information). The ungapped alignment for all 16 structures is found in 80 positions, of which 72 lie in blocks of helical regions (Figure 3).

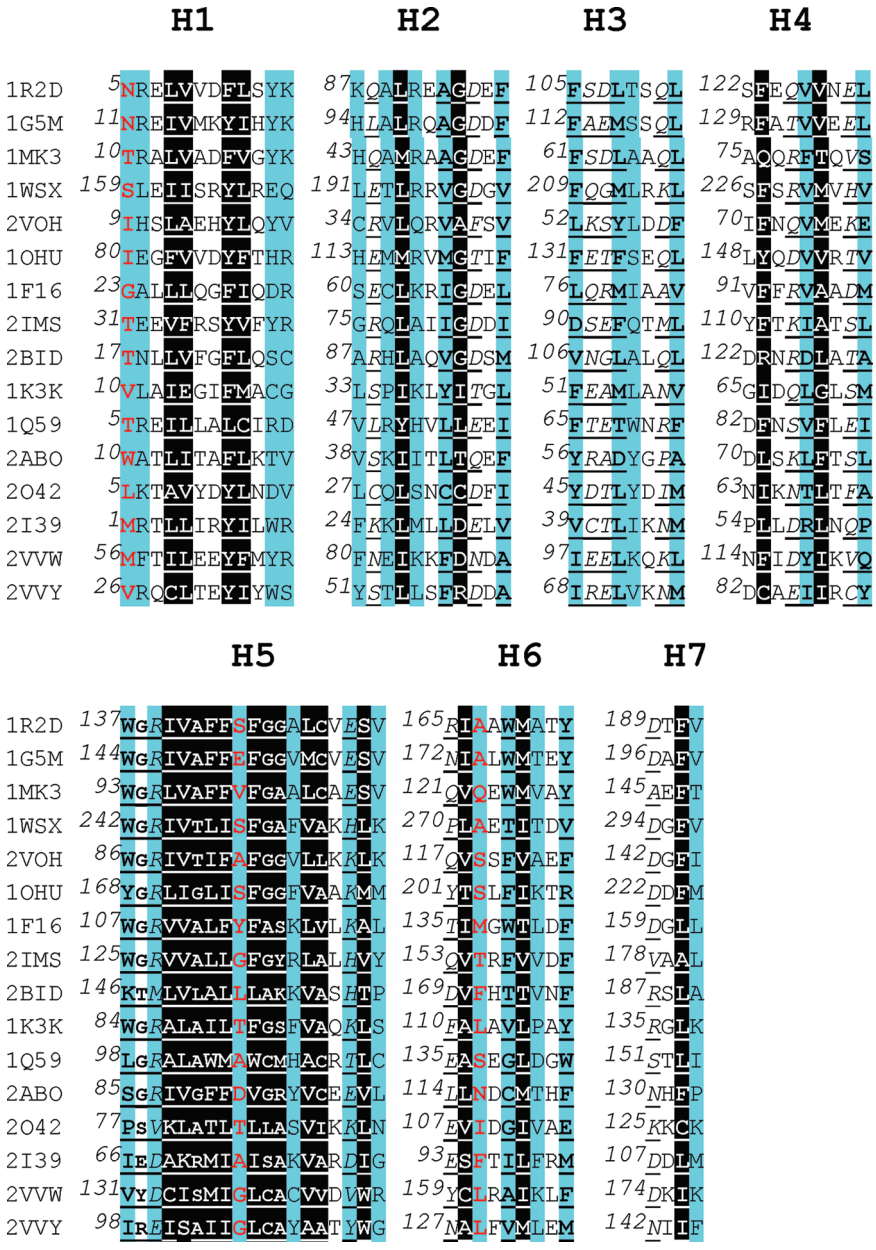

FIGURE 3: Ungapped alignment regions from different helical segments produced from structure-based sequence alignment. MUSTANG (*47*) was used for multiple-structure alignment. Segments belonging to different helices are indicated. The following notations are used to describe the properties: red for buried, bold and underlined for hydrophobic, italic and underlined for hydrophilic, cyan background for interhelical interactions, and black background for all three properties conserved. Small residues are shown in reduced font. The properties are displayed only if residues in the structurally equivalent positions in at least 12 of 16 structures exhibit the properties. For each helical segment, the residue number of the first residue is provided as in their respective UniProt entry (see Table 1).

These 80 residues constitute ~70% of the total sequence length for which structural information is available in the small anti-apoptotic Bcl-2 homologue Vaccinia virus N1L protein (PDB entry 2I39). In the case of the large pro-apoptotic protein BID (PDB entry 2BID), this ungapped alignment region covers ~41% of the polypeptide segment in the protein structure. It should be mentioned that the long loop connecting helices H1 and H2 is absent in the viral Bcl-2 homologues. In the ungapped blocks of alignments, residues at each position assumed an α-helical conformation in at least eight structures, and at least four such positions were observed consecutively to form helical regions. The rest of the analysis will focus on these ungapped alignment blocks and more specifically on the core helical regions.

Sequence and Structure Variations in the Ungapped Regions. We have performed pairwise sequence alignments of all 16 Bcl-2 sequences by considering only the 80 positions that make up the blocks of ungapped alignments. We have analyzed the percent sequence identity of all 120 possible pairwise sequence alignments (Table 2 and Table S1 of the Supporting Information). Only eight pairs exhibit a level of sequence identity

Table 2: Pairwise Sequence and Structural Comparisons of All Possible 120 Bcl-2 Homologue Pairs in the Ungapped Alignment Regions[a]

level of sequence identity (%)	>40	30−40	20−30	<20
no. of sequence pairs	3	5	27	85
rmsd (Å)	<2	2−3	3−4	>4
no. of structural pairs	5	23	62	30

[a]The ungapped alignment region was obtained from multiple-structure alignment of 16 nonredundant Bcl-2 homologue structures. For details, see Figure 3 and Figure S1 of the Supporting Information.

Table 3: Analysis of Helix Bundle Geometry in the Bcl-2 Helical Fold Calculated for All 16 Structures of Bcl-2 Homologues

interacting helix pair[a]	⟨ω⟩ (±standard deviation)[a] (deg)	⟨d⟩ (±standard deviation)[b] (Å)
H5−H6	164.0 ± 7.1	9.8 ± 0.8
H5−H7	84.8 ± 12.3	10.4 ± 1.6
H2−H3	87.1 ± 14.4	11.7 ± 1.6
H1−H6	133.7 ± 14.9	9.3 ± 1.3
H6−H7	96.7 ± 15.2	15.8 ± 2.2
H3−H5	58.2 ± 18.0	10.3 ± 1.5
H1−H7	47.0 ± 18.0	13.1 ± 2.3
H1−H5	53.9 ± 22.6	9.4 ± 1.6
H2−H7	66.3 ± 24.5	10.6 ± 1.6
H2−H5	95.4 ± 25.0	9.7 ± 1.2
H1−H2	127.9 ± 25.2	9.7 ± 1.1
H3−H4	112.7 ± 27.4	11.4 ± 2.5
H4−H6	43.5 ± 30.5	10.4 ± 1.3
H4−H5	138.8 ± 34.5	9.9 ± 1.1

[a]Average and standard deviation of helix−helix crossing angles calculated for the interacting helix pairs. The data are presented according to the increasing order of standard deviation observed in ω. [b]Average and standard deviation of interhelical distances calculated for the interacting helix pairs.

of >30%, and 27 sequence pairs exhibit a level of sequence identity between 20 and 30% which could be described as the so-called twilight zone region (*50*). More than 80 of 120 pairs show less than 20% sequence identity, and in the case of some of the Bcl-2 protein pairs, the level of sequence identity falls to as low as 5%. This clearly reveals the extent of diversity in Bcl-2 members. The low level of sequence identity observed among the majority of Bcl-2 members even in the structurally homologous regions will make the task of identifying potential remote Bcl-2 homologues a difficult and challenging one.

To quantify the common core helical fold and to evaluate the extent of structural homology of Bcl-2 family proteins, we determined the root-mean-square deviation (rmsd) for all possible structural pairs from the 16 Bcl-2 homologue structures using MOLMOL (*51*). Only the backbone atoms (Cα, N, and C′) in the structurally conserved ungapped alignment regions were considered for this purpose (Figure 3). With the 120 possible pairs, the rmsd values varied from 1.5 to 5.3 Å (Table S1 of the Supporting Information). The Bcl-2 homologues Bcl-X_L (PDB entry 1R2D) and Bcl-2 (PDB entry 1G5M) are structurally the most similar with a rmsd of 1.5 Å. The pro-apoptotic Bid (PDB entry 2BID) and the anti-apoptotic viral homologue M11L from Myxoma virus (PDB entry 2O42) form the most diverse pair. Overall, more than 90 of 120 pairs analyzed have rmsd values of >3 Å (Table 2). This suggests that although the Bcl-2 family proteins have an overall common helical fold, differences in the packing of helices could have contributed to the higher rmsd observed in some of the Bcl-2 structural pairs. This is clear from the superposition of the most [PDB entries 1R2D and 1G5M (Figure 1B)] and least [PDB entries 2BID and 2O42 (Figure 1C)] similar Bcl-2 structural pairs.

Characterization of the Bcl-2 Helix Bundle Fold. We have characterized the Bcl-2 helix bundle geometry by determining the helix−helix crossing angles (ω) and interhelix distances (d) of all the 14 interacting helix pairs (Table 3). It is clear that some helix pairs exhibit a larger variation in these parameters compared to others. Interactions between central hydrophobic helix H5 and its succeeding helix H6 are the most stable ones as revealed by the standard deviation of ω (164 ± 7.1°) and d (9.8 ± 0.8 Å). However, the interhelix angle shows the largest variation with H5 and its preceding helix H4 (138.8 ± 34.5°), although the change observed in the distance between these two interacting helices is small. Other helix pairs involving helices H3 and H7 exhibit a large variation in interhelix distance. Helix H3 has been suggested to undergo conformational changes upon interacting with its partner Bcl-2 protein (*27*). Helix H7 is truncated at different levels in many Bcl-2 structures. BH3-containing helix H2 and C-terminal helix H7 both adopt a perpendicular orientation with respect to central hydrophobic helix H5. Two other helix pairs, H2 and H3 and H6 and H7, also assume a perpendicular orientation with respect to each other. This analysis clearly shows that helical packing in the core varies considerably across the different Bcl-2 structures. A representative example of the H4−H5 interacting helix pair showing the largest difference in ω angle (177.4° in PDB entry 1MK3 and 71.4° in PDB entry 2BID) is shown in panels C and D of Figure 2.

Helices H2, H4, and H7 along with central hydrophobic helix H5 together form the hydrophobic groove that is characteristic of the Bcl-2 fold. Characterization of individual Bcl-2 homologue structures indicates that this groove exhibits certain features specific to a particular Bcl-2 homologue (*32*, *33*, *52*). Because this groove plays a crucial role in the binding of pro- and anti-apoptotic proteins, we wanted to see whether there is any specific pattern that distinguishes the hydrophobic grooves between different groups of Bcl-2 members as classified in Table 1. The interhelical distances among helices H2, H4, and H5 have been used for this purpose (helix H7 is not included because of different levels of truncation in different Bcl-2 structures). We have determined the sum of the interhelical distances between three helix pairs, namely, H2−H5, H4−H5, and H2−H4. This parameter can give an indication of whether these helices have resulted in a narrow or broad hydrophobic groove. Our analysis indicates that the sum of interhelical distances between the three helix pairs in five of six anti-apoptotic Bcl-2 homologues ranges from 35 to 38 Å (only mouse A1 has a broader hydrophobic groove with a sum distance of ∼42 Å). Viral Bcl-2 homologues exhibit a similar range of 30−40 Å. In contrast to anti-apoptotic Bcl-2 proteins, pro-apoptotic proteins Bax, Bak, and Bid result in a greater distance ranging from 41 to 45 Å. Such a broad groove has been noticed in the individual structures of pro-apoptotic Bcl-2 structures (*32*, *33*, *52*). Surface plots of Bcl-2 representative examples from anti-apoptotic, pro-apoptotic, and viral homologues are depicted in Figure 4. This analysis clearly indicates that the hydrophobic groove in anti-apoptotic homologues is somewhat narrower than that in their pro-apoptotic counterparts, and this feature could be important for the tight binding of the BH3 helix of pro-apoptotic partners. One should also note that only three pro-apoptotic structures have been used in this analysis. The significance of a broader hydrophobic groove in pro-apoptotic proteins and its biological significance should be investigated further when more pro-apoptotic structures become available. However, the primary aim of this study is to elucidate the common features that give rise to a characteristic Bcl-2 fold among the seemingly diverse sequences in different classes of Bcl-2 proteins. Hence, we have analyzed different residue properties at the structurally equivalent positions among all 16 Bcl-2 homologues to determine the core residues that are responsible for this unique fold.

Residue Properties at Each Position in the Ungapped Alignment Region. Pairwise sequence alignment clearly showed the poor sequence identity in a majority of the Bcl-2 sequence pairs. Although the Bcl-2 homologues adopt a similar fold, the

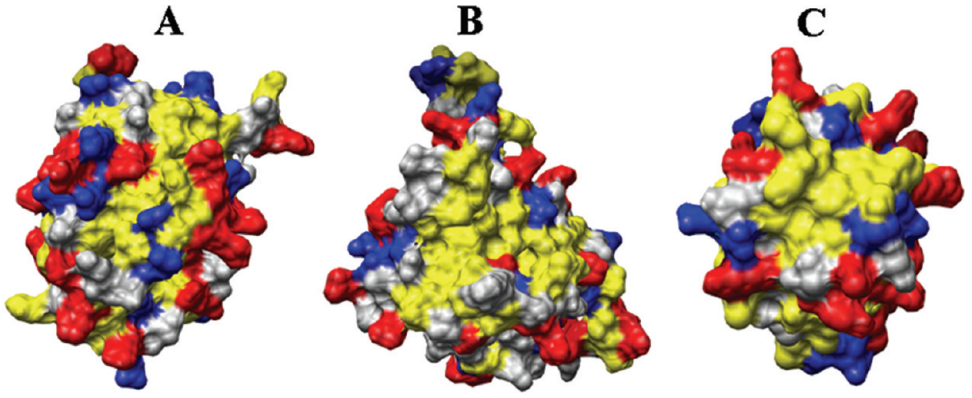

FIGURE 4: Surface representation of (A) an anti-apoptotic protein Bcl-2 (PDB entry 1G5M), (B) a pro-apoptotic protein Bid (PDB entry 2BID), and (C) a viral homologue from Vaccinia virus (PDB entry 2VVY, chain A). Molecular surface plots were created using the UCSF Chimera package (*69*). The hydrophobic residues (Ala, Val, Ile, Leu, Met, Cys, Gly, Phe, Trp, and Tyr) are colored yellow. Asp and Glu are colored red. Lys, Arg, and His are colored blue. All other residues are colored gray. The sums of pairs of interhelical distances involving helices H2, H4, and H5 are 37.6, 45.0, and 33.8 Å, respectively, for Bcl-2, Bid, and the Vaccinia virus Bcl-2 homologue, respectively.

rmsd analysis indicated differences in packing of helices, and this observation is obviously due to substitutions of residues at different positions. Hence, the intriguing question is what makes the Bcl-2 family of proteins adopt the same fold despite a high level of divergence at the sequence level. There must be some conservation at some of the structurally equivalent positions in the ungapped regions. To answer this question, we have analyzed certain properties at each position, including accessible surface area, size and chemical conservation, and the tendency to participate in interhelical interactions for each residue in all 16 Bcl-2 structures.

Solvent Inaccessible Residues in the Bcl-2 Helical Fold. The interior of the proteins is usually tightly packed and is not accessible to the solvent. They generally constitute the core of a protein, and a change in these residues is expected to disturb the residue−residue packing. Hence, a higher degree of conservation is expected for those residues that are buried (*53*, *54*). To investigate this property, we have calculated the solvent accessible surface area (SASA) for all the residues in each Bcl-2 structure as described in Materials and Methods. The average solvent accessible value was calculated for the residues in each of the 80 structurally equivalent positions in the ungapped alignment block from all 16 structures. A particular position belonging to the ungapped alignment region is considered to be buried if the average accessible surface area of the residues in that position is ≤ 20 Å^2 (*19*). In total, we found that residues belonging to 27 positions can be considered as buried and 25 of them occur in helical regions. Buried residues are colored red or have a black background in Figure 3. The maximum number of 12 buried positions occurs in helix H5. This is the central hydrophobic helix surrounded by other amphipathic helices. The next helix with a larger number of five buried residues is observed for helix H1. Helices H2, H4, H6, and H7 together have eight positions that are solvent inaccessible. Only two residues outside the structurally conserved helical regions have solvent accessible areas of < 20 Å^2. Using our criterion, none of the residues in helix H3 can be considered buried. The polar character of three of eight positions is maintained in this helix in diverse Bcl-2 structures (see below) which could probably explain its preference to be exposed to the solvent.

Residue Conservation. Absolute residue conservation is not observed in any of the 80 positions in the ungapped alignment block. However, it is possible that the chemical nature or size of the residues could be conserved at certain important positions. First, we broadly classified the amino acids as either hydrophobic (A, V, I, L, M, C, G, F, W, and Y) or hydrophilic (K, R, H, D, E, P, S, T, N, and Q) and examined each of the 80 columns. The same classification was used in deriving a hydrophobicity scale based on 511 protein structures (*55*). In total, 37 positions show conservation in hydrophobic character, of which 33 occur in the helical regions (bold and underlined in Figure 3). In 12 positions, conservation of the hydrophobic nature of residues is observed in all 16 structures. In other positions, the number of structures retaining the hydrophobic character varies from 12 to 15. Not surprisingly, the hydrophobic nature is conserved in a maximum number of 13 residues in central helix H5. This is followed by helices H1, H2, H4, and H6, and each has four residues showing conservation of hydrophobic character. Only four residues in the nonhelical regions maintain their hydrophobic character.

In contrast to the number of structurally equivalent positions retaining their hydrophobic character, only 14 positions exhibit conservation in maintaining the hydrophilic nature of residues (italic and underlined in Figure 3), and 11 of them lie in the helical regions. They seem to be equally distributed in four of the seven Bcl-2 helices, with helix H3 having the maximum number of three hydrophilic positions. It should be pointed out that helix H3 did not have any residues that can be considered buried. Even in hydrophobic helix H5, two positions (one in the beginning and the other in the end of the helix) can be considered hydrophilic in nature. According to our classification, only helix H1 does not have any position that can be considered polar.

We have also examined the conservation of amino acids on the basis of their size. We have examined all 80 positions in which small and weakly polar residues (G, A, C, T, and S) are group conserved. We found five positions, all of them in central hydrophobic helix H5, in which the size of the residues seems to be conserved (shown in smaller font in Figure 3). Such a group-based conservation gives rise to a (G/S/T)xxx(A/T/G/S)xxxx(G/A/C/S)(G/A/S)xx(C/A) sequence motif. It has been shown that motifs such as GxxxG drive helix−helix association in transmembrane proteins (*56*, *57*). In fact, in at least two families of transmembrane helical bundle proteins, small and weakly polar residues have been shown to be group-conserved at the helix−helix interface (*58*, *59*). It has been shown that mutagenesis of the positions at which small and weakly polar residues at the helix−helix interface are group conserved affects the expression

and activity of the mutant proteins (*60*, *61*). Hence, a high degree of conservation of small and weakly polar residues as a group could be attributed to the role of these positions in tight helix−helix association. Helix H5 is the central helix around which the other helices pack. The sequence motif consisting of small residues would thus help in efficient packing of the surrounding helices with this helix.

Conservation of Interhelical Interactions. We have also determined the involvement of each residue in interhelical interactions in all Bcl-2 structures using the criterion described in Materials and Methods. In the ungapped alignment region, we determined the number of residues in each structurally equivalent position participating in helix−helix interactions. Since contacts between the helices are likely to play a significant role in the preservation of the Bcl-2 fold, conservation of interhelical interactions in structurally equivalent positions across the diverse Bcl-2 proteins will be a major factor in the Bcl-2 family of proteins. All 14 interacting helix pairs were considered for this purpose (see Table 3 for the possible interacting helix pairs). Except in two positions, residues in all 19 positions in helix H5 are involved in interhelical contacts in at least 12 of the 16 structures (cyan or black background). Helices H1 and H2 are the other two helices showing the maximum number of residues participating in interactions with other helices. In total, 53 structurally equivalent positions in the ungapped regions have residues that participate in interhelical interactions in at least 12 of the 16 structures considered in this study. Forty-eight of these positions lie in the helical regions, and 41 of them interact with a specific helix or helices (Table 4A) in at least 12 structures. A maximum number of six residues from central helix H5 interacts with helix H6. Residues from some of the positions belonging to the helical regions of H3 (L108 and L112; residue numbers correspond to those of PDB entry 1R2D), H4 (L130), H5 (A142 and V155), H6 (I166), and H7 (V192) show interhelical interactions in at least 12 structures, but their interacting helices are not always the same in all the structures and hence are not included in Table 4A. We have also looked at specific residue pairs that have been shown to have a stable interaction in molecular dynamics simulations (*49*). We found that 10 such interactions, a majority of them involving residues from central helix H5, are observed in more than 50% of the structures analyzed in this study (Table 4B).

Table 4: (A) Residues in Structurally Equivalent Positions from One Helix Involved in Interhelix Interactions with Another Specific Helix

helix	residues[a] (interacting helix)[b]
H1	N5 (H7), R6 (H6), **L8** (H2), **V9** (H6), **F12** (H2, H5), **L13** (H5, H6), Y15 (H2), K16 (H5)
H2	K87 (H1), A89 (H7), **L90** (H1, H7), R91 (H1), A93 (H5), **G94** (H5), F97 (H5)
H3	F105 (H5), T109 (H5)
H4	**F123** (H6), V126 (H3), **V127** (H5, H6)
H5	W137 (H7), R139 (H4), **I140** (H6), **V141** (H2), **F143** (H4, H6), **F144** (H1, H2, H6), S145 (H2), **F146** (H3, H4), **G147** (H6), **G148** (H1), A149 (H3), **L150** (H3), **C151** (H1, H6), E153 (H3), **S154** (H6)
H6	A167 (H1), W169 (H4), **M170** (H1, H5), A171 (H1), Y173 (H4)
H7	**F191** (H2, H5)

(B) Specific Residue Pairs Which Are Shown To Exhibit Stable Interactions in Molecular Dynamics Simulations (*49*) and Are Observed in More Than 50% of the Structures in the Present Data Set

interacting helices	interacting residues[c]
H1−H5	V9...F144
	F12...F144
	L13...C151
	K16...V152
H4−H5	V127...F143
	L130...F143
	L130...R139
H4−H6	F123...W169
	V127...Y173
H5−H6	F143...M170

[a]The residues correspond to the sequence reported in PDB entry 1R2D. For structurally equivalent residues in other Bcl-2 homologues, see Figure 3. Residues shown in bold have all three properties (buried, chemical nature and/or size, and involvement in interhelical interactions) conserved. [b]Specific helix with which the residues from a given helix interact in at least 12 of 16 structures. [c]The residues correspond to the sequence reported for the Bcl-X_L structure in PDB entry 1R2D. For structurally equivalent residues in other Bcl-2 homologues, see Figure 3.

DISCUSSION

In our data set, the longest and shortest polypeptide chains for which structures are determined consisted of 195 (PDB entry 2BID) and 115 (PDB entry 2I39) residues, respectively. Multiple-structure alignment of all 16 Bcl-2 homologue structures using MUSTANG (*47*) resulted in 374 columns. No gaps were introduced in 80 positions, indicating that structurally equivalent residues were observed in all 16 structures and a majority of 90% of these positions lie in the seven helical regions. We have focused our analysis mainly on the structurally equivalent positions in the helical regions and studied properties such as solvent accessible surface area, chemical nature, size, and their involvement in interhelical interactions. Of 72 structurally equivalent positions in the helical regions, 25 were considered buried, and the hydrophobic or hydrophilic nature of 44 positions was conserved. In five positions in central hydrophobic helix H5, small and weakly polar residues were group conserved. Residues in larger numbers (48 positions) in the helical regions were found to take part in interhelical interactions. We examined the number of positions that are simultaneously solvent inaccessible, show conservation in chemical nature or size, and participate in interhelical interactions. In total, 22 positions in the helical regions exhibit all of these properties (Table 5; residues with a black background in Figure 3), and central helix H5 alone has 11 such positions. Among the 11 remaining positions, helix H1 has four positions and helices H2, H4, and H6 have two positions each. The residues in the conserved positions in helices H1, H2, H4, and H6 are separated by three to four residues. Outside the core helical regions, only two positions (V163 and G187; PDB entry 1R2D numbering) just before and after helix H6 possess all the properties that were analyzed. We have shown these positions in two very distantly related Bcl-2 proteins, Bcl-X_L and N1L, in Figure 5. Bcl-X_L is from humans, and its anti-apoptotic homologue, N1L, is from Vaccinia virus; they show just 9% sequence identity in the ungapped regions. N1L does not possess any recognizable BH sequence domains. Structural superposition shows that the rmsd between the two structures in the ungapped alignment regions is 3.6 Å. Our systematic analysis of residue properties has suggested the 22 positions showing conservation in solvent accessibility, chemical nature and/or size, and interhelical interactions in these disparate Bcl-2 homologues. Interestingly, most of these positions occur in helices H1 and H4−H6 at the opposite side of the hydrophobic groove. We propose that these 22 positions could constitute the core nucleus that is important for maintaining the Bcl-2 helical fold. If we slightly relax the criteria

for conservation, then there are another 15 positions in the helical regions where residues in these positions possess two of the three properties (Table 5). It should be noted that many of these additional residue positions are observed in the region where the hydrophobic groove is located.

Table 5: Residues in Structurally Equivalent Positions Having at Least Two of the Three Analyzed Properties

properties	helix	residue[a]
buried, hydrophobic, and interhelical interactions	H1	L8, V9, F12, L13
	H2	L90, G94
	H4	F123, V127
	H5	I140, V141, **A142**, F143, F144, F146, **G147**, **G148**, L150, **C151**, S154
	H6	I166, M170
	H7	F191
buried and interhelical interactions	H1	N5
	H5	S145
	H6	A167
hydrophobic and interhelical interactions	H2	A93, F97
	H3	F105, L108, L112
	H4	V126, L130
	H5	W137
	H6	W169, Y173
hydrophilic and interhelical interactions	H5	R139, E153

[a]Residues correspond to the sequence of the Bcl-X_L protein from PDB entry 1R2D. For structurally equivalent positions in other Bcl-2 homologues, see Figure 3. Residues shown in bold are those residues in which small size is conserved. The properties are conserved in at least 12 of 16 structures analyzed in this study.

The Bcl-2 family of proteins has joined the group of proteins that are sequentially diverse yet exhibit remarkable similarity at the structural level. Other examples in this group include the cytokines and cytochromes that form four-helix bundles (*19*), hemoglobins from diverse organisms (*18*), restriction endonucleases (*62*), copper-binding cupredoxins involved in electron transport (*20*), and members of the major intrinsic protein superfamily (*59*, *63*). The list of such examples is growing rapidly as new structures from distantly related family members emerge. It is not straightforward to find the sequence−structure relationship in these families of proteins. In this study, our approach is similar to that used in cytokines, cytochromes, hemoglobins, and cupredoxins by Chothia, Lesk, and co-workers (*18*−*20*). Sequence diversity implies a large number of different substitutions, and such disparate substitutions are accommodated by changes in the helix packing geometry as demonstrated in our analysis, but still they seem to maintain the overall helix bundle fold. There could be slight variation in the number of identified core residue positions if a stricter cutoff in solvent accessible surface area is used to identify solvent inaccessible residues or a different hydrophobicity scale is selected to define the chemical nature of residues. However, we believe the choice of different options to define these properties will not dramatically change the conserved positions.

To exploit the results reported in this analysis, we searched the NCBI nonredundant sequence database (http://www.ncbi.nlm.

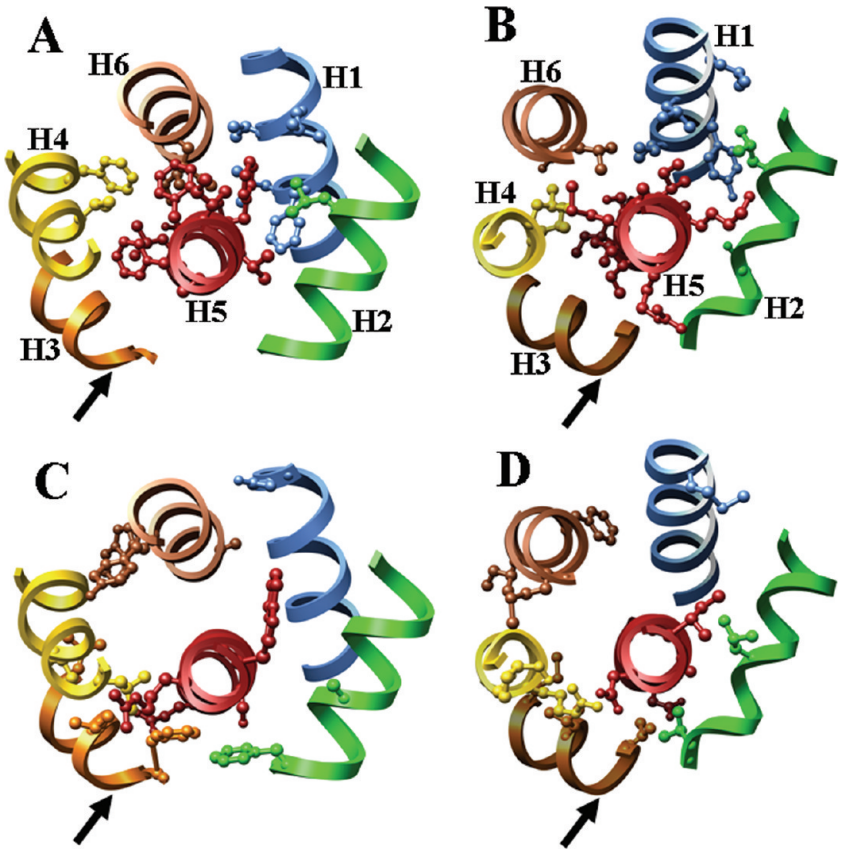

FIGURE 5: Residues in structurally equivalent positions with conserved properties are shown in ball-and-stick representations for two diverse Bcl-2 homologues, namely, Bcl-X_L from humans (A and C) (PDB entry 1R2D) and N1L from Vaccinia virus (B and D) (PDB entry 2I39). In panels A and B, labels of individual helices are shown. The position of the hydrophobic cleft is indicated with a black arrow. Panels A and B show residue positions that conserve all three properties (buried, chemical nature and/or size, and involvement in interhelical interactions) are shown. Panels C and D display the residue positions that have two of the three properties. See also Table 5 for details about individual residues.

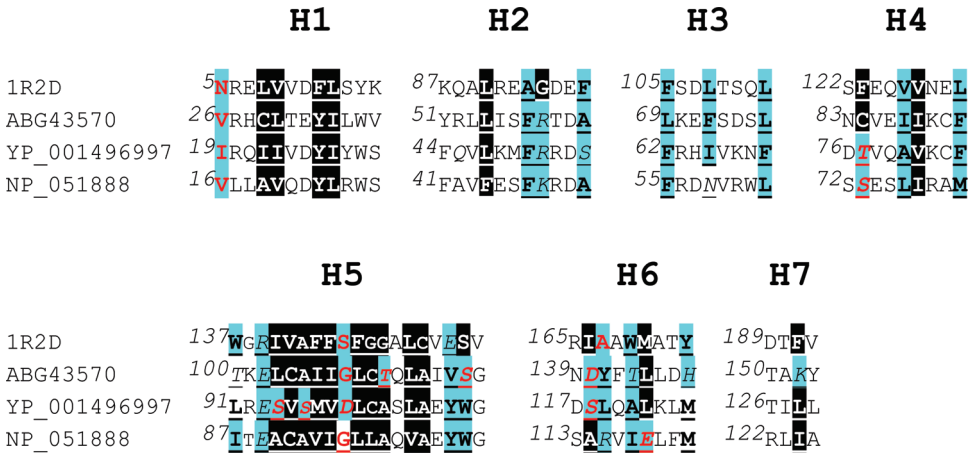

FIGURE 6: Structure-based sequence alignment corresponding to the 72 core residue positions for the putative Bcl-2 homologues from the three viral organisms. Different helical segments are indicated. For an explanation of the different notations in the alignment, see Figure 3. For the sake of comparison, the sequence of Bcl-X_L is also shown (PDB entry 1R2D). The accession numbers of the Bcl-2 homologues from *V. virus*, *T. virus*, and *R. fibroma virus* are ABG43570, YP_001496997, and NP_051888, respectively. Residue properties of only the 22 core residue positions and the 15 additional positions are displayed.

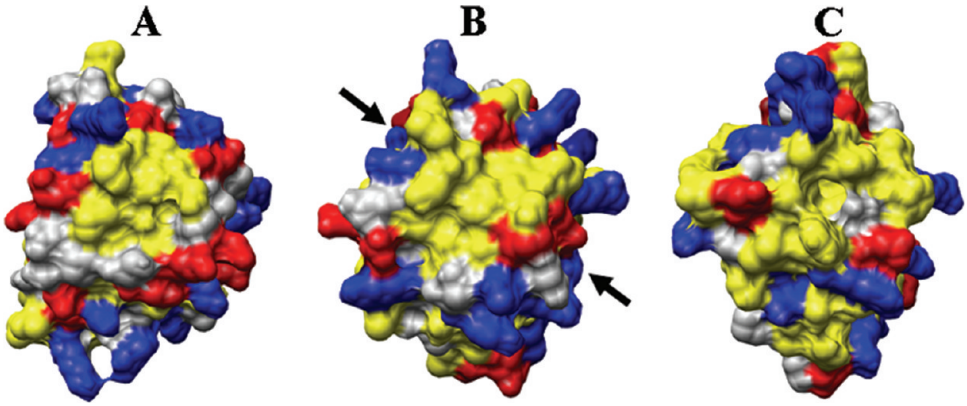

FIGURE 7: Surface representation of the three putative Bcl-2 viral homologues from (A) *V. virus*, (B) *T. virus*, and (C) *R. fibroma virus*. Molecular surface plots were created using the UCSF Chimera package (*69*). The hydrophobic residues (Ala, Val, Ile, Leu, Met, Cys, Gly, Phe, Trp, and Tyr) are colored yellow. Asp and Glu are colored red. Lys, Arg, and His are colored blue. All other residues are colored gray. A pronounced hydrophobic groove is clearly present at the center in the putative Bcl-2 homologues. Even the distributions of charged residues seem to be conserved at least in the top left and bottom right sides of the groove and are indicated by arrows.

nih.gov/) using PSI-BLAST (*64*) by considering the amino acid sequence corresponding to the 72 structurally conserved positions (Figure 3) as a query sequence from each of the viral Bcl-2 homologues (Table 1). Each query sequence was subjected to at least three rounds of iterations of PSI-BLAST as implemented in the FASTA suite of packages (http://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml). We identified three sequences that are annotated as hypothetical protein or whose function is not clear, and their accession numbers in NCBI are ABG43570, YP_001496997, and NP_051888. They are from *Variola virus* (smallpox virus), *Tanapox virus*, and *Rabbit fibroma virus*, respectively. These viral organisms have been implicated in various diseases, and their genome sequences are known (*65*−*67*). The reported *E* values ranged from 0.004 to 7.5, and the levels of sequence identity with the 16 Bcl-2 homologues considered for this study varied from 10.7 to 38.2%. The highest level of sequence identity for all the three sequences was found to be with the Bcl-2 homologue of Vaccinia virus (PDB entry 2VVY). We used this structure as a template to model the structure of the three potential Bcl-2 homologues identified in the PSI-BLAST search using MODELER (http://salilab.org/modeller/) (*68*) (version 9v7). The modeling protocol for building the structure was similar to that used previously in our laboratory to successfully model the sequentially diverse group of proteins belonging to the superfamily of major intrinsic proteins (MIPS) (*59*). The models thus generated were used for further structural analysis. For each model, the solvent accessibility of individual residues and interhelical interactions were determined as described in Materials and Methods. Structure-based sequence alignment of the three potential Bcl-2 homologues is shown in Figure 6 along with a representative Bcl-2 homologue. This region corresponds to the ungapped alignment block produced for the 16 Bcl-2 homologues (Figure 3). Of the 22 core structural positions, 16−18 positions have all three properties conserved in the three potential Bcl-2 homologues. Four of five residues are either small or weakly polar residues in the central hydrophobic helix. Among the 15 additional positions, 11 residues have two of the three conserved properties. We have also plotted the molecular surface of the modeled structures (Figure 7). The hydrophobic groove is clearly present in all the potential Bcl-2 homologues, and a closer examination reveals a similar distribution of charged residues present on the outer rim of the hydrophobic groove especially at the upper left and lower right of the groove. Identification of these viral proteins as putative Bcl-2

homologues in this study is a major step toward designing further experiments that can unambiguously demonstrate Bcl-2 anti-apototic activity.

It is likely that there are many such anti-apoptotic Bcl-2 homologues in several viral organisms that are yet to be identified. These viral Bcl-2 proteins could be potentially good drug targets, and identification of these homologues will be of utmost importance. The core residue positions identified in this study could be used to detect remote Bcl-2 homologues from viral genomes and genomes of other organisms. Experimentally discovered or computationally identified new Bcl-2 homologues can be verified for the existence of these structurally equivalent positions. Another 15 positions in which two of three properties are conserved could be used as additional constraints in the identification of new Bcl-2 family members. Satisfaction of these additional constraints will increase the level of confidence in the prediction of distantly related Bcl-2 family members as demonstrated in the identification of three new putative Bcl-2 homologues in this study. The interhelical interactions due to the residues from these core residue positions could be important in the folding of Bcl-2 homologues. Mutation experiments in these positions can be conducted to investigate the role of individual residues in the folding pathway that leads to the unique Bcl-2 helical fold.

ACKNOWLEDGMENT

We are grateful to Professor Arthur Lesk and his co-workers for MUSTANG. We thank Dr. Marc Lensink for his program which can calculate helix axis. Mr. Ravi Kumar Verma is greatly acknowledged for his help in modeling the putative viral Bcl-2 homologs. We thank all our lab members for the useful discussions.

SUPPORTING INFORMATION AVAILABLE

Information about pairwise sequence identity and root-mean-square deviation obtained for all possible Bcl-2 homologue pairs in the ungapped alignment region (Table S1) and the entire structure-based sequence alignment for the 16 Bcl-2 homologues (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Meier, P., Finch, A., and Evan, G. (2000) Apoptosis in development. *Nature 407*, 796–801.
2. Youle, R. J., and Strasser, A. (2008) The BCL-2 protein family: Opposing activities that mediate cell death. *Nat. Rev. Mol. Cell Biol. 9*, 47–59.
3. Reed, J. C. (2000) Mechanisms of apoptosis. *Am. J. Pathol. 157*, 1415–1430.
4. Kelekar, A., and Thompson, C. B. (1998) Bcl-2-family proteins: The role of the BH3 domain in apoptosis. *Trends Cell Biol. 8*, 324–330.
5. Adams, J. M., and Cory, S. (2001) Life-or-death decisions by the Bcl-2 protein family. *Trends Biochem. Sci. 26*, 61–66.
6. Muchmore, S. W., Sattler, M., Liang, H., Meadows, R. P., Harlan, J. E., Yoon, H. S., Nettesheim, D. G., Chang, B. S., Thompson, C. B., Wong, S.-L., Ng, S.-C., and Fesik, S. F. (1996) X-ray and NMR structure of human Bcl-X_L, an inhibitor of programmed cell death. *Nature 381*, 335–341.
7. Petros, A. M., Olejniczak, E. T., and Fesik, S. W. (2004) Structural biology of the Bcl-2 family of proteins. *Biochim. Biophys. Acta 1644*, 83–94.
8. Hardwick, J. M., and Youle, R. J. (2009) Snapshot: Bcl-2 proteins. *Cell 138*, 404–405.
9. Douglas, A. E., Corbett, K. D., Berger, J. M., McFadden, G., and Handel, T. M. (2007) Structure of M11L: A myxoma virus structural homolog of the apoptosis inhibitor, Bcl-2. *Protein Sci. 16*, 695–703.
10. Kvansakul, M., van Delft, M. F., Lee, E. F., Gulbis, J. M., Fairlie, W. D., Huang, D. C. S., and Colman, P. M. (2007) A structural viral mimic of prosurvival Bcl-2: A pivotal role for sequestering proapoptotic Bax and Bak. *Mol. Cell 25*, 933–942.
11. Graham, S. C., Bahar, M. W., Cooray, S., Chen, R. A.-J., Whalen, D. M., Abrescia, N. G. A., Alderton, D., Owens, R. J., Stuart, D. I., Smith, G. L., and Grimes, J. M. (2008) Vaccinia virus proteins A52 and B14 share a Bcl-2-like fold but have evolved to inhibit NF-κB rather than apoptosis. *PLoS Pathog. 4*, e1000128.
12. Kalverda, A. P., Thompson, G. S., Vogel, A., Schroder, M., Bowie, A. G., Khan, A. R., and Homans, S. W. (2009) Poxvirus K7 protein adopts a Bcl-2 fold: Biochemical mapping of its interactions with human DEAD box RNA helicase DDX3. *J. Mol. Biol. 385*, 843–853.
13. Aoyagi, M., Zhai, D., Jin, C., Aleshin, A. E., Stec, B., Reed, J. C., and Liddington, R. C. (2007) Vaccinia virus N1L protein resembles a B cell lymphoma-2 (Bcl-2) family protein. *Protein Sci. 16*, 118–124.
14. Kvansakul, M., Yang, H., Fairlie, W. D., Czabotar, P. E., Fischer, S. F., Perugini, M. A., Huang, D. C. S., and Colman, P. M. (2008) Vaccinia virus anti-apoptotic F1L is a novel Bcl-2-like domain-swapped dimer that binds a highly selective subset of BH3-containing death ligands. *Cell Death Differ. 15*, 1564–1571.
15. Rost, B. (1997) Protein structures sustain evolutionary drift. *Folding Des. 2*, S19–S24.
16. Nagano, N., Orengo, C. A., and Thornton, J. M. (2002) One fold with many functions: The evolutionary relationship between TIM barrel families on their sequences, structures and functions. *J. Mol. Biol. 321*, 741–765.
17. Lecomte, J. T. J., Vuletich, D. A., and Lesk, A. M. (2005) Structural divergence and distant relationships in proteins: Evolution of the globins. *Curr. Opin. Struct. Biol. 15*, 290–301.
18. Lesk, A. M., and Chothia, C. (1980) How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins. *J. Mol. Biol. 136*, 225–270.
19. Hill, E. E., Morea, V., and Chothia, C. (2002) Sequence conservation in families whose members have little or no sequence similarity: The four-helical cytokines and cytochromes. *J. Mol. Biol. 322*, 205–233.
20. Gough, J., and Chothia, C. (2004) The linked conservation of structure and function in a family of high diversity: The monomeric cupredoxins. *Structure 12*, 917–925.
21. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235–242.
22. Denisov, A. Y., Madiraju, M. S. R., Chen, G., Khadir, A., Beauparlant, P., Attardo, G., Shore, G. C., and Gehring, K. (2003) Solution structure of human BCL-w. Modulation of ligand binding by the C-terminal helix. *J. Biol. Chem. 278*, 21124–21128.
23. Hinds, M. G., Lackmann, M., Skea, G. L., Harrison, P. J., Huang, D. C. S., and Day, C. L. (2003) The structure of Bcl-w reveals a role for the C-terminal residues in modulating biological activity. *EMBO J. 22*, 1497–1507.
24. Manion, M. K., O'Neil, J. W., Giedt, C. D., Kim, K. M., Zhang, K. Y. Z., and Hockenbery, D. M. (2004) Bcl-X_L mutations supress cellular sensitivity to antimycin A. *J. Biol. Chem. 279*, 2159–2165.
25. Sattler, M., Liang, H., Nettesheim, D. G., Meadows, R. P., Harlan, J. E., Eberstadt, M., Yoon, H. S., Shuker, S. B., Chang, B. S., Minn, A. J., Thompson, C. B., and Fesik, S. F. (1997) Structure of Bcl-x_L-Bak peptide complex: Recognition between regulators of apoptosis. *Science 275*, 983–986.
26. Petros, A. M., Nettesheim, D. G., Wang, Y., Olejniczak, E. T., Meadows, R. P., Mack, J., Swift, K., Matayoshi, E. D., Zhang, H., Thompson, C. B., and Fesik, S. F. (2000) Rationale for Bcl-x_L/Bad peptide complex formation from structure, mutagenesis, and biophysical studies. *Protein Sci. 9*, 2528–2534.
27. Liu, X., Dai, S., Zhu, Y., Marrack, P., and Kappler, J. W. (2003) The structure of a Bcl-x_L/Bim fragment complex: Implications for Bim function. *Immunity 19*, 341–352.
28. Petros, A. M., Medek, A., Nettesheim, D. G., Kim, D. H., Yoon, H. S., Swift, K., Matayoshi, E. D., Oltersdorf, T., and Fesik, S. W. (2001) Solution structure of the antiapoptotic protein bcl-2. *Proc. Natl. Acad. Sci. U.S.A. 98*, 3012–3017.
29. Day, C. L., Chen, L., Richardson, S. J., Harrison, P. J., Huang, D. C. S., and Hinds, M. G. (2005) Solution structure of prosurvival Mcl-1 and characterization of its binding by proapoptotic BH3-only ligands. *J. Biol. Chem. 280*, 4738–4744.
30. Smits, C., Czabotar, P. E., Hinds, M. G., and Day, C. L. (2008) Structural plasticity underpins promiscuous binding of the prosurvival protein A1. *Structure 16*, 818–829.
31. Woo, J.-S., Jung, J.-S., Ha, N.-C., Shin, J., Kim, K.-H., Lee, W., and Oh, B.-H. (2003) Unique structural features of a BCL-2 family protein

CED-9 and biophysical characterization of CED-9/EGL-1 interactions. *Cell Death Differ*. *10*, 1310–1319.
32. Suzuki, M., Youle, R. J., and Tjandra, N. (2000) Structure of Bax: Coregulation of dimer formation and intracellular localization. *Cell 103*, 645–654.
33. Moldoveanu, T., Liu, Q., Tocilj, A., Watson, M., Shore, G., and Gehring, K. (2006) The X-ray structure of a Bak homodimer reveals an inhibitory zinc binding site. *Mol. Cell 24*, 677–688.
34. Chou, J. J., Li, H., Salvesen, G. S., Yuan, J., and Wagner, G. (1999) Solution structure of BID, an intracellular amplifier of apoptotic signalling. *Cell 96*, 615–624.
35. Huang, Q., Petros, A. M., Virgin, H. W., Fesik, S. W., and Olejniczak, E. T. (2002) Solution structure of a Bcl-2 homolog from Kaposi sarcoma virus. *Proc. Natl. Acad. Sci. U.S.A. 99*, 3428–3433.
36. Huang, Q., Petros, A. M., Virgin, H. W., Fesik, S. W., and Olejniczak, E. T. (2003) Solution structure of the BHRF1 protein from Epstein-Barr virus, a homolog of human Bcl-2. *J. Mol. Biol. 332*, 1123–1130.
37. Loh, J., Huang, Q., Petros, A. M., Nettesheim, D. G., van Dyk, L. F., Labrada, L., Speck, S. H., Levine, B., Olejniczak, E. T., and Virgin, H. W. (2005) A surface groove essential for viral Bcl-2 function during chronic infection in vivo. *PLoS Pathog. 1*, e10.
38. UniProt Consortium The Universal Protein Resource (UniProt) 2009. *Nucleic Acids Res. 37*, D169–D174.
39. Lensink, M. F., Christiaens, B., Vandekerckhove, J., Prochiantz, A., and Rosseneu, M. (2005) Penetratin-membrane association: W48/R52/W56 shield the peptide from the aqueous phase. *Biophys. J. 88*, 939–952.
40. Christopher, J. A., Swanson, R., and Baldwin, T. O. (1996) Algorithms for finding the axes of a helix: Fast rotational and parametric least-squares method. *Comput. Chem. 20*, 339–345.
41. Lee, S., and Chirikjian, G. S. (2004) Interhelical angle and distance preferences in globular proteins. *Biophys. J. 86*, 1105–1117.
42. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol. 55*, 379–400.
43. Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D., and Wishart, D. S. (2003) VADAR: A web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res. 31*, 3316–3319.
44. Chothia, C. (1975) Structural invariants in protein folding. *Nature 254*, 304–308.
45. Ghosh, T. S., Chaitanya, S. K., and Sankararamakrishnan, R. (2009) End-to-end and end-to-middle interhelical interactions: New classes of interacting helix pairs in protein structures. *Acta Crystallogr. D65*, 1032–1041.
46. Bowie, J. U. (1997) Helix packing in membrane proteins. *J. Mol. Biol. 272*, 780–789.
47. Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J., and Lesk, A. M. (2006) MUSTANG: A multiple structural alignment algorithm. *Proteins: Struct., Funct., Bioinf. 64*, 559–574.
48. Eswar, N., and Ramakrishnan, C. (2000) Deterministic features of side-chain main-chain hydrogen bonds in globular protein structures. *Protein Eng. 13*, 227–238.
49. Lama, D., and Sankararamakrishnan, R. (2008) Anti-apoptotic Bcl-X_L protein in complex with BH3 peptides of pro-apoptotic Bak, Bad, and Bim proteins: Comparative molecular dynamics simulations. *Proteins: Struct., Funct., Bioinf. 73*, 492–514.
50. Rost, B. (1999) Twilight zone of protein sequence alignments. *Protein Eng. 12*, 85–94.
51. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51–55.
52. McDonnell, J. M., Fushman, D., Milliman, C. L., Korsmeyer, S. J., and Cowburn, D. (1999) Solution structure of the proapoptotic molecule BID: A structural basis for apoptotic agonists and antagonists. *Cell 96*, 625–634.
53. Hubbard, T. J. P., and Blundell, T. L. (1987) Comparison of solvent-inaccessible cores of homologous proteins: Definitions useful for protein modelling. *Protein Eng. 1*, 159–171.
54. Gong, S. S., Worth, C. L., Bickerton, G. R. J., Lee, S., Tanramluk, D., and Blundell, T. L. (2009) Structural and functional restraints in the evolution of protein families and superfamilies. *Biochem. Soc. Trans. 37*, 727–733.
55. Jesior, J.-C. (2000) Hydrophilic framework in proteins? *J. Protein Chem. 19*, 93–103.
56. Senes, A., Gerstein, M., and Engelman, D. M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: The GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J. Mol. Biol. 296*, 921–936.
57. Eilers, M., Patel, A. B., Liu, W., and Smith, S. O. (2002) Comparison of helix interactions in membrane and soluble α-bundle proteins. *Biophys. J. 82*, 2720–2736.
58. Liu, W., Eilers, M., Patel, A. B., and Smith, S. O. (2004) Helix packing moments reveal diversity and conservation in membrane protein structure. *J. Mol. Biol. 337*, 713–729.
59. Bansal, A., and Sankararamakrishnan, R. (2007) Homology modeling of major intrinsic proteins in rice, maize and *Arabidopsis*: Comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. *BMC Struct. Biol. 7*, 27.
60. Chelikani, P., Hornak, V., Eilers, M., Reeves, P. J., Smith, S. O., RajBhandary, U. L., and Khorana, H. G. (2007) Role of group-conserved residues in the helical core of β2-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A. 104*, 7027–7032.
61. Liu, K., Kozono, D., Kato, Y., Agre, P., Hazama, A., and Yasui, M. (2005) Conversion of aquaporin 6 from an anion channel to a water-selective channel by a single amino acid substitution. *Proc. Natl. Acad. Sci. U.S.A. 102*, 2192–2197.
62. Pawlak, S. D., Radlinska, M., Chmiel, A. A., Bujnicki, J. M., and Skowronek, K. J. (2005) Inference of relationships in the 'twilight zone' of homology using a combination of bioinformatics and site-directed mutagenesis: A case study of restriction endonucleases BspI and PvuII. *Nucleic Acids Res. 33*, 661–671.
63. Gonen, T., and Walz, T. (2006) The structure of aquaporins. *Q. Rev. Biophys. 39*, 361–396.
64. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389–3402.
65. Esposito, J. J., Sammons, S. A., Frace, A. M., Osborne, J. D., Olsen-Rasmussen, M., Zhang, M., Govil, D., Damon, I. K., Kline, R., Laker, M., Li, Y., Smith, G. L., Meyer, H., LeDuc, J. W., and Wohlhueter, R. M. (2006) Genome sequence diversity and clues to the evolution of Variola (Smallpox) virus. *Science 313*, 807–812.
66. Nazarian, S. H., Barrett, J. W., Frace, A. M., Olsen-Rasmussen, M., Khristova, M., Shaban, M., Neering, S., Li, Y., Damon, I. K., Esposito, J. J., Essani, K., and McFadden, G. (2007) Comparative genetic analysis of genomic DNA sequences of two human isolates of *Tanapox virus*. *Virus Res. 129*, 11–25.
67. Willer, D. O., McFadden, G., and Evans, D. H. (1999) The complete genome sequence of Shope (Rabbit) Fibroma virus. *Virology 264*, 319–343.
68. Sali, A., and Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779–815.
69. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.