JPP JOURNAL OF PROTEINS AND PROTEOMICS
3(3), December 2012, pp. 187-196

Research Article

IN SILICO CHARACTERIZATION OF BOVINE *(BOS TAURUS)* ANTIAPOPTOTIC PROTEINS

V. G. Vidhya[1], Akhilesh Upgade[2], Anusha Bhaskar[2]* and Dipanjana Deb[1]

[1]*Department of Biotechnology, Faculty of Science and Humanities, SRM University, Kattankulathur Chennai – 603203, Tamil Nadu, India*
[2]*Department of Biotechnology, CRD, PRIST University, Vallam, Thanjavur - 613403, Tamil Nadu, India*

Abstract: Anti-apoptotic proteins, found in a variety of species, protect organisms from cell death and shows great diversity in structure. In this study, a total of 15 anti-apoptotic proteins of *Bos taurus* retrieved from Swiss-Prot database were analysed and characterized using *in silico* tools. Primary structure analysis showed that most of the proteins are hydrophobic in nature due to the high content of non-polar residues. The presence of cysteines in Q01314, Q95108, P63243, Q5EAC7, Q27ID4 and Q2KJE0 infer that these proteins may form disulphide (SS) bonds, which are regarded as a positive factor for stability. The computed aliphatic index infers that these anti-apoptotic proteins may be stable over a wide range of temperature. Secondary structure analysis shows that most of the proteins would display predominant α-helical structure and rest would have mixed secondary structure. The very high coiled structural content of P35445 (69.18%) and Q5EAC7 (52.23%) are due to the rich content of more flexible glycine and hydrophobic proline amino acids. SOSUI server predicts one transmembrane region in P35445 and Q05688 and two transmembrane regions in Q1RMX3 amongst *Bos taurus* antiapoptotic proteins. The predicted transmembrane regions were visualized and analysed using helical wheel plots. The presence of disulphide (SS) bonds in Q01314, Q95108, P62243, Q5EAC7, Q27ID4 and Q2KJE0 were predicted and also were identified from the three-dimensional structure. The disulphide bonds identified from the three-dimensional structure might be correct as the evaluation parameters are within the acceptable limits for the modelled 3D structures. These structures will provide a good foundation for functional analysis of experimentally derived crystal structures.

Keywords: Anti-apoptotic proteins; computational analysis; disulphide bridges; homology modeling; proteomics tools

Introduction

Apoptosis, a crucial biological process, plays an essential role in regulating development, homeostasis, and immune defence by clearing redundant or abnormal cells in organisms. A delicate balance between pro-apoptotic and anti-apoptotic mechanisms determines whether a cell death signal can activate the execution of the apoptotic program. In this balance, pro-apoptotic proteins promote apoptosis and anti-apoptotic proteins inhibit apoptosis (Schimmer, 2004).

As members of the anti-apoptotic family of proteins, inhibitors of apoptosis proteins (IAPs) can inhibit the downstream components of the caspase activation pathways in the regulation of apoptosis and play important roles in regulating the progress of apoptosis in many species (Nachmias, 2004).

IAP family members are characterized by the BIR domain, the name of which derives from the original discovery of these apoptosis suppressors in the genome of baculoviruses (Crook *et al.*, 1993). The BIR domains consist of approximately 70

Corresponding Author: **Anusha Bhaskar**
E-mail: dranushaparthiban@gmail.com

Received: September 11, 2012
Accepted: December 26, 2012
Published: December 30, 2012

amino acids that contain the characteristic sequence CX2CX16HX6C. With both hydrophobic and hydrophilic residues on its surface, the BIR core is theoretically capable of supporting protein-protein interactions. There are three subtypes of BIR domain, BIR1, BIR2, and BIR3, classified by their evolutionary relationship in phylogenies (Huang *et al.*, 2000). The progress of apoptosis is regulated in an orderly way by a series of signal cascades under certain circumstances. Three main factors, IAP, IAP antagonist, and caspase, are involved in regulating this progress. Above all, the accepted regulatory model is that IAP can suppress cellular apoptosis through the inhibition of caspases. Therefore, these IAPs also called as anti-apoptotic proteins are essential for cell survival, which requires the active inhibition of apoptosis, by inhibiting the expression of pro-apoptotic factors as well as promoting the expression of anti-apoptotic factors (Fan *et al.*, 2005).

Numerous structure and function studies of anti-apoptotic proteins have been reported experimentally from time to time while computational studies are much more limited. Many IAP family members have been identified in diverse species ranging from viruses to mammals. However, studies on bovine (*Bos taurus*) anti-apoptotic proteins have not yet been reported so far. So, the effort has been taken to study the physicochemical and structural properties of anti-apoptotic proteins in *Bos taurus* as it is an important model organism for biomedical research and development, in addition to being agriculturally important (Liu *et al.*, 2009).

Today, computational tools provide researchers a cost effective way to understand physicochemical and the structural properties of a protein for the successful design of many biological experiments with in a short range of time. Several physicochemical properties of a protein such as molecular weight, grand average hydropathy (GRAVY), aliphatic index (AI), extinction coefficient (EC), isolelectric point (pI), instability index (II) etc., can be computed along with their functional characterization. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function, physical and chemical properties.

Sequence analysis and physicochemical characterization of proteins using bio-computation tools have been done by many researchers and reported. However, physicochemical characterization of anti-apoptotic proteins has not been done so far. Therefore in this study, we had focussed on the *in silico* characterization and homology modeling of 15 anti-apoptotic proteins of bovine (*Bos taurus*).

Materials and Computational Methods

Anti-apoptotic protein sequences - Anti-apoptotic protein sequences were retrieved from the manually curated public protein database SwissProt (Arnold *et al.*, 2006). Swiss-Prot was scanned for the key word anti-apoptosis. The sequences were retrieved in FASTA format and used for further analysis.

Amino acid composition - The amino acid compositions of 15 anti-apoptotic protein sequences of *Bos taurus* were computed using the tool CLC free workbench (CLC bio, 2006).

Primary structure analysis - Percentages of hydrophobic and hydrophilic residues were calculated from the primary structure analysis.

Physico-chemical parameters - The physico-chemical parameters such as theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient (Gill and Von Hippel, 1989), half-life (Bachmair *et al.*, 1986; Gonda *et al.*, 1989; Tobias *et al.*, 1991; Ciechanover and Schwartz, 1989), instability index (Guruprasad *et al.*, 1990), aliphatic index (Ikai, 1980) and grand average hydropathy (GRAVY) (Kyte and Doolittle, 1982) were computed using Expasy's ProtParam (*http://us.expasy.org/tools/protparam.html*) prediction server.

Secondary structure prediction - The tools SOPM, SOPMA (Comb *et al.*, 2000) and SSCP (Secondary Structural Content Prediction) server (Eisenhaber *et al.*, 1996) were used for the secondary structure prediction.

Identification of transmembrane region - The SOSUI server (Hirokawa *et al.*, 1998) performed

the identification of transmembrane regions. The predicted transmembrane helices were visualized and analysed using helical wheel plots generated by the program Pepwheel (Ramachandran and Sasiskharan, 1968) included in the EMBOSS 2.7 suite.

Presence of SS bonds - The presence of -SS- bond and their bonding patterns were predicted by CYS_REC (CYS-REC, 2006) and RASMOL server. CYS_REC (*http://linux1. softberry.com/berry.phtml*) identified the position of a cysteine, total number of cystiene presented along with the most probable –SS- bond pairs in the protein sequences. The latter tool Rasmol involves the identification of –SS- bonds using the 3D structure of a protein.

Homology modeling and validation - The modeling of 3D structure of 6 anti-apoptotic proteins was performed by Swiss-model (Arnold *et al.*, 2006). Homology modeling of these six proteins was done by using a template structure from PDB (*http://www.pdb.org/pdb/ home/home.do*) through BLASTP search (*http:// blast.ncbi.nlm.nih.gov/blast.cgi*). The modelled 3D structures were evaluated using the online server Rampage (Lovell *et al.*, 2002), ProQ (Protein quality server) (Cristobal *et al.*, 2001) and ProSA. The structure validation of anti-apoptotic proteins was performed by online PROCHECK (Laskowski *et al.*, 1996) and What IF (Vriend, 1990) server.

Results

Table 1 shows the 15 anti-apoptotic protein sequences of *Bos taurus* (bovine) retrieved from SWISS-PROT. The primary structure analysis was done and different parameters computed using ExPasy ProtParam tool and tabulated (Tables 2, 3, 4). The results suggest that proteins from bovine are mostly hydrophobic and their hydrophobic nature is due to the presence of high content of non-polar residues. The presence of 21 cysteine residues in Q01314, Q95108, P63243, Q5EAC7, Q27ID4 and Q2KJE0 indicate the presence of disulphide bonds in corresponding anti-apoptotic proteins (Table 5).

The average molecular weight of anti apoptotic proteins calculated is 46263.57dalton (Table 4). The isoelectric point (pI) is the pH at which the surface of protein is covered with charge but net charge of the protein is zero. At pI proteins are stable and compact. The computed pI value of P17870, P35445, P61285, Q01314, O02703, Q05688, Q9NOP9, Q5EAC7, Q1RMX3 and Q2KJE0 which have pI < 7 indicates that these anti-apoptotic proteins are acidic and pI of P62894, Q95108, P63243, Q02399, Q27ID4 are >7 indicating the basic nature of corresponding anti-apoptotic proteins. The computed isoelectric point (pI) will be useful for developing buffer systems for purification of the recombinant proteins by isoelectric focussing method.

Although Expasy's Protparam computes the extinction coefficient (EC) for a range of (276, 278, 279, 280 and 282 nm) wavelength, 280nm is favoured because proteins absorb strongly at this wavelength while other substances commonly in protein solutions do not interfere. Extinction coefficient of anti apoptotic proteins at 280 nm range from 11585 to 109135 M^{-1} cm^{-1} with respect

Table 1
Anti-apoptotic proteins of *Bos taurus* retrieved from Swiss Prot

S.No	*Accession No*	*Gene name*	*Protein Name*
1	P17870	ARRB1	Beta-arrestin-1
2	P62894	CYC	Cytochrome c
3	P61285	DYL1	Dynein light chain 1, cytoplasmic
4	P35445	COMP	Cartilage oligomeric matrix protein
5	Q01314	AKT1	RAC-alpha serine/ threonine-protein kinase
6	Q95108	THIOM	Thioredoxin, mitochondrial
7	O02703	BAX	Apoptosis regulator BAX
8	Q9N0P9	PIM1	Serine/threonine-protein kinase pim-1
9	P63243	GBLP	Guanine nucleotide-binding protein subunit
10	Q05688	IGF1R	Insulin-like growth factor 1 receptor
11	Q02399	CDK5	Cyclin-dependent kinase 5
12	Q5EAC7	CPIN1	Anamorsin
13	Q1RMX3	B2CL2	Bcl-2-like protein 2
14	Q27ID4	TERT	Telomerase reverse transcriptase
15	Q2KJE0	TAXB1	Tax1-binding protein 1 homolog

Table 2
Amino acid composition (in %) of anti-apoptotic proteins computed using CLC free Work Bench tool

Amino acids	*P17 870*	*P62 894*	*P61 285*	*P35 445*	*Q01 314*	*Q95 108*	*O02 703*	*Q9N 0p9*	*P63 243*	*Q05 688*	*Q02 399*	*Q5E AC7*	*Q1R MX3*	*Q27 ID4*	*Q2K JE0*
Ala	5.5	5.8	7.9	6.1	5.2	6.5	6.2	4.8	4.7	6.1	5.8	9.0	15.5	10.8	5.9
Cys	1.7	1.9	3.4	6.1	1.2	1.9	1.0	1.9	2.5	1.3	2.7	3.2	1.0	2.8	1.8
Asp	7.4	2.9	5.6	13.2	5.8	11.2	5.7	6.7	6.6	4.5	7.2	6.8	4.1	2.9	5.9
Glu	8.6	8.7	7.9	4.2	10.4	5.6	6.2	7.7	2.8	9.4	5.8	7.7	6.7	3.6	11.8
Phe	3.8	3.8	5.6	3.7	5.6	4.7	5.2	4.5	2.5	4.0	5.1	2.3	6.2	4.2	3.5
Gly	5.5	13.5	4.5	9.6	6.0	6.5	11.5	7.0	8.5	6.5	5.5	6.5	10.9	8.3	3.1
His	2.4	2.9	4.5	1.5	2.5	2.8	0.0	4.2	2.5	1.8	2.4	1.6	1.6	2.7	2.9
Ile	3.6	5.8	7.9	2.3	4.2	4,7	4.7	5.8	6.0	4.1	3.8	4.2	0.5	1.3	3.1
Lys	8.4	17.3	11.2	3.4	7.3	9.3	4.7	3.8	5.4	4.1	7.9	8.7	2.1	2.8	10.2
Leu	10	5.8	4.5	4.6	8.8	7.5	11.5	12.1	8.9	8.5	13.7	12.3	8.8	13.6	9.5
Met	1.2	1.9	3.4	1.4	3.3	9.3	4.2	1.3	1.3	4.3	1.4	1.3	2.1	0.9	1.7
Asn	4.1	4.8	4.5	5.8	2.7	2.8	1.6	1.9	4.4	5.4	4.5	2.9	1.6	1.9	4.7
Pro	7.2	3.8	1.1	6.4	4.6	4.7	3.6	6.1	3.2	6.7	6.5	4.8	4.1	7.7	4.4
Gln	2.2	2.9	4.5	6.5	3.5	4.7	5.2	3.5	3.8	2.6	3.1	3.2	4.7	4.5	6.6
Arg	6.0	1.9	2.2	5.8	6.5	1.9	5.7	7.0	4.4	6.2	5.5	2.9	6,7	12.1	3.4
Ser	4.1	1.0	4.5	4.2	5.0	1.9	5.7	7.3	9.5	7.3	4.5	12.3	4.7	5.7	6.1
Thr	6.5	7.7	4.5	5.0	6.5	4.7	7.3	2.6	9.8	4.5	4.1	2.9	5.7	3.7	6.2
Val	8.9	2.9	5.6	7.3	5.6	15.0	5.7	7.2	7.0	7.5	6.2	5.8	8.3	7.3	6.0
Trp	0.0	1.0	1.1	1.4	1.5	0.9	3.1	2.2	4.1	1.3	1.0	0.6	2.6	1.2	0.9
Tyr	3.1	3.8	5.6	1.4	3.8	0.9	1.0	2.2	1.9	4.0	3.4	1.0	2.1	2.1	2.7

Table 3
Content of hydrophilic and hydrophobic residues

Accession No	*Percentage of Hydrophobic Residues*	*Percentage of Hydrophilic Residues*	*Net Hydrophobic Residues Content*
P17870	45.6	21.53	High
P62894	39.4	22.11	High
P61285	41.57	26.96	High
P35445	42.79	29.34	High
Q01314	44.79	22.7	High
Q95108	50	25.30	High
O02703	55.72	21.87	Very high
Q9N0P9	51.11	19.48	Very high
P63243	46.37	31.86	High
Q05688	48.9	24.68	High
Q02399	48.97	22.26	High
Q5EAC7	46.77	25.48	High
Q1RMX3	59.06	19.68	Very high
Q27ID4	54.66	20.71	Very high
Q2KJE0	37.69	28.15	High

to the concentration of Cys, Trp and Tyr (Table 4). The high EC value of Q05688 and Q27ID4 indicates the presence of high concentration of Cys, Trp and Tyr. The computed EC values will help in the quantitative study of protein-protein and protein-ligand interactions in solution. The bio computed half – life of the antiapoptotic protein Q05688 is 1.4 h while for the rest of the antiapoptotic proteins is 30 h.

A protein whose instability index is smaller than 40 is predicted as stable, and a value above 40 predicts that the protein may be unstable (Guruprasad *et al.*, 1990). The instability index available at Expasy's Protparm classifies P35445, P61285, Q05688, Q9N0P9, Q5EAC7 and Q2KJE0 antiapoptotic proteins as unstable (Instability index >40) and others as stable (Instability index <40) (Table 4).

The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains (A, V, I and L) is regarded

Table 4
Parameters computed using Expasy's ProtParam tool

Accession No	*Sequence Length*	*Mol. Wt*	*pI*	*-R*	*+R*	*EC*	*II*	*AI*	*GRAVY*
P17870	418	47131.7	5.9	67	60	19745	36.95	84.35	-0.520
P62894	104	11572.3	9.52	12	20	11585	10.45	59.13	-0.866
P61285	89	10365.8	6.89	12	12	13075	42.74	72.36	-0.448
P35445	736	80469.0	4.36	128	68	72650	41.36	54.42	-0.745
Q01314	480	55748.4	5.64	78	66	65695	36.29	71.90	-0.585
Q95108	166	18416.2	8.44	18	20	17085	35.41	92.11	-0.145
O02703	192	21259.4	5.12	23	20	36105	35.47	85.83	-0.089
Q9N0P9	313	35629.7	5.63	45	34	49305	41.85	95.88	-0.231
P63243	317	35076.7	7.60	30	31	80940	22.67	82.68	-0.251
Q05688	640	72511.4	5.30	88	70	81750	45.90	76.59	-0.413
Q02399	292	33288.4	7,57	38	39	31900	28.37	91.82	-0.302
Q5EAC7	310	33204.7	5.23	45	36	16095	44.48	90.03	-0.255
Q1RMX3	193	20774.4	5.40	21	17	33585	18.14	75.96	-0.044
Q27ID4	1125	124446.8	11.0	73	168	109135	51.62	90.13	-0.179
Q2KJE0	817	94058.7	5.33	144	111	72155	45.58	71.00	-0.854

Mol.Wt – Molecular Weight; pI- Isoelectric Point; -R – Number of negative residues; +R – Number of Positive residues; EC – Extinction Coefficient at 280 nm; II – Instability Index; AI – Aliphatic Index; GRAVY – Grand Average Hydropathicity

as a positive factor for the increase of thermal stability of globular proteins. The lower thermal stability of P62894 and P35445 is indicative of a more flexible structure when compared to other anti-apoptotic proteins (Table 4). The high aliphatic index of all anti-apoptotic proteins infers that they may be stable over a wide range of temperature.

The Grand Average Hydropathicity (GRAVY) index of antiapoptotic proteins ranged from -0.0 to -0.8 (Table 4). The very low GRAVY index of Q95108, O02703, Q1RMX3 and Q27ID4 infers that these anti-apoptotic proteins could result in a better interaction with water.

The secondary structure predicted with the help of programs SOPM and SOPMA (data not shown) infers that the anti-apoptotic proteins Q02703, Q1RMX3 and Q2KJE0 have rich adenine content and are mostly α-helices. Anti-apoptotic proteins P62894, Q01314, Q95108, Q02399, Q9N0P9 and Q27ID4 have mixed secondary structure, i.e., α – helices and coils. The very high coil structural content of P35445 (69.18%) and Q5EAC7 (52.23%) are due to rich content of more flexible glycine and hydrophobic proline amino acids. Proline has a special property of creating kinks in polypeptide chains and disrupting ordered secondary structure.

Besides all other physicochemical characterization, functional characterization of anti-apoptotic proteins was also performed including transmembrane (TM) region identification, prediction of disulphide bonding pairs etc. The SOSUI server performed the identification of transmembrane helices with their corresponding length and differentiates membrane proteins from stable proteins. The server SOSUI classifies P35445, Q05688, Q1RMX3 as membrane proteins and other anti-apoptotic proteins as soluble proteins (Table 6). SOSUI server had identified one transmembrane region in P35445 and Q05688 and two transmembrane regions in Q1RMX3. The helices of P35445, Q05668 and Q1RMX3 were visualized using EMBOSS P in wheel format (Figure 1).

Possible disulphide bond pairing and patterns with probability were predicted by CYS_REC from primary sequence and S-S bonds were identified from 3D structure by RASMOL. The tool CYS_REC recognizes the presence of 21 cys residues in Q01314, Q95108, P62243, Q5EAC7, Q27ID4 and Q2KJE0 sequences and predicted

Table 5
Disulphide (SS) bond pattern of pairs predicted, by CYS_REC (using primary structure) and identified by RASMOL (using 3D structure modelled)

Accession No	*CYC_REC*	*RasMol*
Q01314	Cys 60 – Cys 296	Cys 60 – Cys 310
		Cys 77 – Cys 296
		Cys 224 – Cys 344
Q95108	Cys 58 – Cys 90	Cys 90 – Cys 93
P63243	Cys 153 – Cys 249	Cys 138 – Cys 249
	Cys 168 – Cys 286	Cys 153 – Cys 286
	Cys 207 – Cys 240	Cys 168 – Cys 182
		Cys 207 – Cys 240
Q5EAC7	Cys 116 – Cys 283	Cys 92 – Cys 249
	Cys 235 – Cys 272	Cys 116 – Cys 235
	Cys 244 – Cys 247	Cys 244 – Cys 275
	Cys 249 – Cys 275	Cys 247 – Cys 286
		Cys 272– Cys 283
Q27ID4	Cys 7 – Cys 401	Cys 7 – Cys 89
	Cys 166 – Cys 423	Cys 54 – Cys 889
	Cys 171 – Cys 738	Cys 57 – Cys 166
	Cys 284 – Cys 313	Cys 76– Cys 401
	Cys 322 – Cys 835	Cys 156 – Cys 631
	Cys 417 – Cys 821	Cys 171 – Cys 313
		Cys 322 – Cys 734
		Cys 336 – Cys 521
		Cys 417 – Cys 423
		Cys 461 – Cys 510
		Cys 738 – Cys 1067
		Cys 789 – Cys 924
		Cys 821 – Cys 975
		Cys 835 – Cys 919
		Cys 838 – Cys 1008
Q2KJE0	Cys 34 – Cys 456	Cys 83 – Cys 640
	Cys 83 – Cys 541	Cys 103 – Cys 761
	Cys 103 – Cys 264	Cys 264 – Cys 788
	Cys 271 – Cys 747	Cys 271 – Cys 747
	Cys 640 – Cys 788	Cys 456 – Cys 541
	Cys 703 – Cys 761	Cys 552 – Cys 703
		Cys 758 – Cys 785

most probable SS bond patterns of pairs in all these proteins.

The three-dimensional structure of modelled protein Q95108 is shown in Figure 2 and ProSA analysis of antiapoptotic protein P632433 is shown in Figure 3. The structures cross verified the predicted secondary structural content and

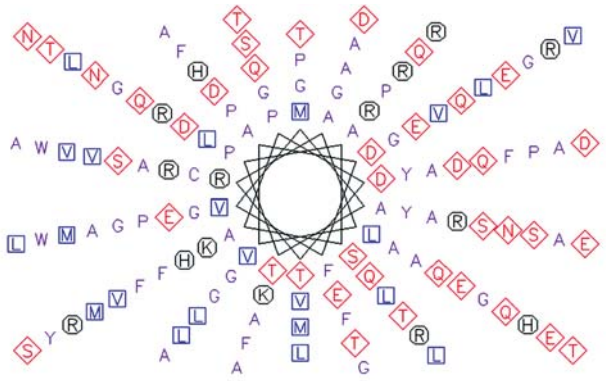

Figure 1: **Helical wheel representation of predicted helix of Q1RMX3 antiapoptotic protein. Hydrophobic residues (I,L,V,M) are represented as blue squares and violet letters (A,G,Y,P), Positively charged residue (H) are represented in octagon**

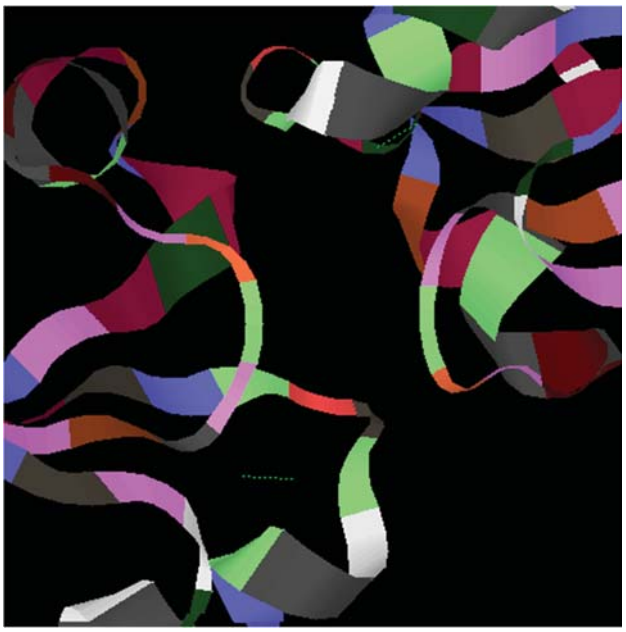

Figure 2: **RasMol (ribbon) representation of the homology modelled 3D structure of antiapoptotic protein Q95108 (using PDB template 1UVZ_A). The SS bonds (dotted lines) are shown in green colour**

presence and location of disulfide bonds. ProSa analysis indicates that the model of antiapototic protein matches with the known proteins whose structures have already been determined by NMR experiments. It signifies good quality of our model. These structures will provide a good foundation for functional analysis of experimentally derived crystal structures.

Discussion

The three-dimensional (3D) protein structures provide valuable insights into the molecular basis

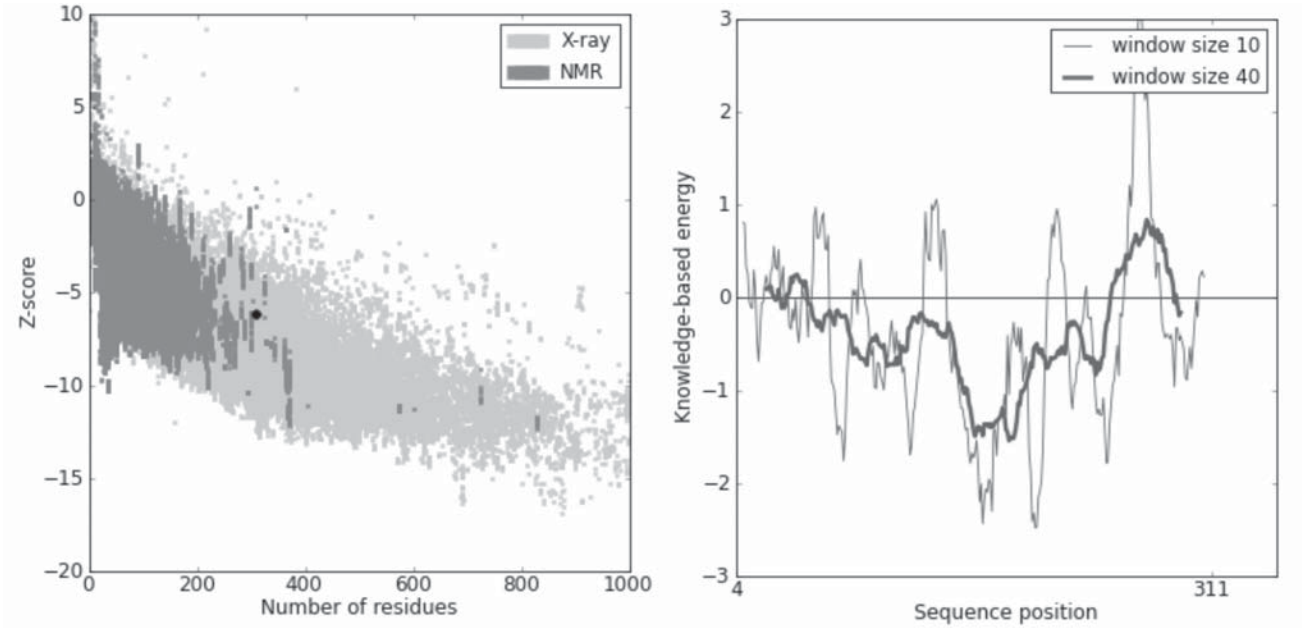

Figure 3: **ProSA-web service analysis of antiapoptotic protein P63243 (A) ProSA-web z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with respect to their length. The z-scores of P63243 highlighted as large dot. (B) Energy plot of P63243**

Table 6
Transmembrane regions identified by SOSUI server

Accession No	*Transmembrane Region*	*Type*	*Length*
P35445	MVLAAARVLLLTLAALGA	Primary	18
Q05688	IHLMIALPIAVLLIVGGLVIMLY	Primary	23
Q1RMX3	LFQGGPNWGRLVAFFVFGAALCA	Secondary	23
	ASVRTVLTGAVALGALVTVGAFF	Primary	23

of protein function, allowing an effective design of experiments. Homology models of proteins are of great interest for planning and analyzing biological experiments when no experimental three dimensional structures are available. Now-a-days, 3D structure of protein can be predicted from amino acid sequences by different web based homology modeling servers at different level of complexity. During evolution, the structure is more stable and changes much slower than the associated sequence, so that similar sequences adopt practically identical structures and distantly related sequences still fold into similar structures (Chothia and Lesk, 1986).

The modeling of 3D structure of protein was performed by Swiss model. Six anti-apoptotic proteins Q01314, Q95108, P63243, Q5EAC7, Q27ID4 and Q2KjE0 were modeled based on various PDB templates selected from the hits obtained through the BLASTP analysis. The stereo chemical quality of the predicted models and accuracy of the protein model was verified (Table 7) after the refinement process using Ramchandran Map calculation computed with

Table 7
PDB templates (hits with maximum % identity) obtained using BLASTP search against the Protein Data Bank

Accession no	*PDB Code*
Q01314	3096_A 30CB_A
Q95108	1UVZ_A
P63243	2ZKQ_AA
Q5EAC7	2LD4_A
Q27ID4	2X35_A 3L42_A
Q2KJE0	2WBT_A

Table 8
Validation parameters computed for the built 3D structures of targets Q01314, Q95108, P63243, Q5EAC7, Q27ID4 and Q2KJE0

Target	*Template (PDB) Code*	*RamPage Percentage of Residues in favoured region*			*PROQ*		*RMS Z Score*
		RFR	*RAR*	*ROR*	*LG Score*	*Max Sub*	
Q01314	3O96_A	95.80%	3.30%	0.90%	5.102	0.53	-7.98
	3OCB_A				5.986	0.475	
Q95108	1UVZ_A	97.60%	2.40%	0.00%	6.928	0.702	-7.34
P63243	2ZKQ_AA	88.20%	7.80%	3.90%	6.493	0.448	-6.16
Q5EAC7	2LD4_A	90.50%	7.70%	1.80%	-0.835	-0.113	-6.09
Q27ID4	2X35_A	92.70%	4.50%	2.70%	4.668	0.531	-6.05
	3L42_A				5.454	0.626	
Q2KJE0	2WBT_A				4.841	0.479	

Table 9
Criteria for a good (model) 3D structure

RamPage Percentage of residues in favoured region	*RMS Z*	*ProQ*		*Quality of the model*
		LG Score	*Max Sub*	
		>1.5	>0.1	Fairly good model
98	1	>2.5	>0.5	Very good model
		>4	>0.8	Extremely good model

PROCHECK program. PROCHECK suite is for assessing the stereo chemical quality of a given protein structure and to measure how normal or conversely, how unusual, the geometry of the residues in a given protein model is as compared with stereo chemical parameters derived from well refined high resolution structure (Table 8). The result revealed that, the proteins Q95108 and Q01314 modeled by Swiss model homology modeling server has average maximum residues in favoured region (RFR) which are about 97.6% and 95.8% respectively.

The modeled structure of anti-apoptotic proteins were also validated by other model verification server ProQ (Protein Quality Server), which validated the protein model based on different validation parameters. Two quality measures, LG score and MaxSub are predicted by ProQ and listed with RMS Z score. The result revealed RMS Z score, LG score, MaxSub and other criterions suggesting good model quality (Table 8). The cysteines and disulphide bonds were identified using 3D structure. Some S-S bonding pairs predicted by CYS_REC are not correlating with the S-S bond positions identified using 'RASMOL'. We speculate that, S-S bonds predicted from 3D structure might be correct and more reliable than the S-S bonds identified from the primary structure (Sivakumar *et al.*, 2007).

ProSA was used to check three dimensional models of anti-apoptotic proteins for potential errors. The program displays two quality measures of the input structure; z-score and a plot of its residue energies. The z-score indicates overall model quality and measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations. As shown in Figure 3 the Z-score of P63243 are well within the range of scores typically found for proteins of similar size indicating a highly reliable structure. The energy plot shows the local model quality by plotting energies as a function of amino acid sequence position. In general, positive values correspond to problematic or erroneous parts of a model.

Conclusion

Fifteen bovine anti-apoptotic proteins have been chosen mainly to study their physico-chemical properties, primary and secondary structures by using computational tools and servers. Primary structure analysis reveals that most of the proteins under study are hydrophobic in nature and contain disulphide linkages. Physico-chemical characterization studies give a good idea about the properties such as pI, EC, AI, GRAVY and stability that are essential and vital in providing data about the proteins and their properties. Secondary structure analysis predicts that most of them contain only α-helices and remaining of them contain mixed structure. The presence of Cys residues in few anti-apoptotic proteins indicates the presence of disulfide bridges which is also confirmed using CYS_REC and Rasmol tools. The investigation here will pave the path for experimental studies and help build testable hypotheses.

Abbreviations

IAPs, inhibitors of apoptosis proteins; BIR, baculovirus inhibitor of apoptosis protein repeat; GRAVY, grand average hydropathy; EC, extinction coefficient; pI, isolelectric point; II instability index; SSCP, secondary structural content prediction; PDB, Protein Data bank; ProSa, Protein Structure Analysis; ProCHECK, Protein Checker; Cys, cysteine; Trp, tryptophan; Tyr, tyrosine; AI, aliphatic index; TM, transmembrane; EMBOSS, European Molecular Biology Open Software Suite; BLAST, *Basic Local Alignment Search Tool; FASTA, Fast All*; ProQ, Protein Quality Server; RMS, *Root Mean Square*

References

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22, 195-201.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). *In vivo*-half-life of a protein is a function of its amino terminal residue Science 234, 179-186.

Chothia, C., and Lesk, A.M. (1986). The relation between the divergence of sequence and structure in proteins. EMBO J. 5, 823-826.

Ciechanover, A., and Schwartz, A.L. (1989). How are substrates recognized by the ubiquitin-mediated proteolytic system? Trends Biochem. Sci. 14, 483-488.

CLC bio. (2006). CLC free Workbench. *http://www.clcbio.com/index.php?id=28*, (27/10/2006).

Combet, C., Blanchet, C., Geourjon, C., and Deléage, G.(2000). NPS@: Network Protein Sequence Analysis. Trends Biochem. Sci. 25, 147-150.

Cristobal, S., Zemla, A., Fischer, D., Rychlewski, L., and Elofsson, A. (2001). A study of quality measures for protein threading models. BMC Bioinform. 2, 5.

Crook, N.E., Clem, R. J., and Miller, L.K. (1993). An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. J. Virol. 67, 2168–2174.

CYS_REC. (2006) *http://sun1.softberry.com/berry.phtml?topic=cys_rec& group=help&subgroup=propt.* (27/10/2006).

Eisenhaber, F., Imperiale, F., Argos, P., and Froemmel, C. (1996). Prediction of secondary structural sontent of sroteins from their amino acid composition alone. I. New analytic vector decomposition methods. Proteins: Struct. Funct. Design 25,157-168.

Fan, T.J., Han, L.H., Cong, R.S., and Liang, J. (2005). Caspase family proteases and apoptosis. Acta Biochem Biophys Sin. 37, 719–727.

Gill, S.C., Von Hippel, and P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182, 319.

Gonda, D.K., Bachmair, A., Wunning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. J. (1989) Universality and structure of the N-end rule. J. Biol. Chem. 264, 16700-16712.

Guruprasad, K., Reddy, B. V. B., and Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Prot. Eng. 4, 155-161.

Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classiûcation and secondary structure prediction system for membrane proteins. Bioinformatics 14, 378–379.

Huang, Q., Deveraux, Q.L., Maeda, S., Salvese, G.S., Stennicke, H.R., Hammock, B.D., and Reeds, J.C. (2000). Evolutionary conservation of apoptosis mechanisms: Lepidopteran and baculoviral inhibitor of apoptosis proteins are inhibitors of mammalian caspase-9. Proc. Natl. Acad. Sci. USA 97, 1427–1432.

Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. J. Biochem. 88, 1895-1896.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR 8, 477-486.

Liu, Y., Qin, X., Song, H, Jiang, H., Shen, Y., Durbin, J.K, Lien, S., Kent, MP, Sodeland, M, Ren, Y., Zhang, L., Sodergren, E., Havlak, P., Worley, K,C, Weinstock G.M., and Gibbs, R.A. (2009). *Bos taurus* genome assembly. *BMC Genom.* 10, 180.

Lovell, S.C., Davis, I.W., Arendall, B., De Bakker, P.I.W., Word, M., Prisant, M.G., Richardson, J, S., and Richardson, D.C. (2002). Structure validation by Cα geometry: φ,ψ and Cβ deviation. Proteins: Struct. Funct.Gen. 50, 437–450.

Nachmias, B., Ashhab, Y., and Ben-Yehuda, D. (2004). The inhibitor of apoptosis protein family (IAPs) : an emerging therapeutic target in cancer. Semin. Canc. Biol. 14, 231–243.

Ramachandran, G.N., and Sasiskharan, V. (1968). Conformation of polypeptides and proteins. Adv. Prot. Chem. 23, 283–437.

Schimmer, A.D. (2004). Inhibitor of apoptosis proteins: Translating basic knowledge into clinical practice. Cancer Res. 64, 7183–7190.

Sivakumar, K., Balaji, S., and Ganga, R. (2007). *In silico* characterization of antifreeze proteins using computational tools and servers. J. Chem. Sci. 119, 571–579.

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. Science 254, 1374-1377.

Vriend, G. (1990). WHAT IF: A molecular modeling and drug design program. J. Mol. Graph 8, 52-56.