BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 10(3)

Differences in structural elements of Bcr-Abl oncoprotein isoforms in Chronic Myelogenous Leukemia

Abdul Hai[1], Nadeem A Kizilbash[1]*, Syeda Huma H Zaidi[2], Jamal Alruwaili[1] & Khuram Shahzad[3]

[1]Department of Biochemistry, Faculty of Medicine & Applied Medical Sciences, Northern Border University; [2]Department of Chemistry, Faculty of Science, Northern Border University, P.O. Box 1321, Arar-91431, Saudi Arabia; [3]Illinois Informatics Institute, University of Illinois, Urbana-Champaign, Illinois, U.S.A; Nadeem A Kizilbash - Email: fsd707@gmail.com; Phone: +966592390610; *Corresponding author

Received January 01, 2014; Revised January 11, 2014; Accepted January 12, 2014; Published March 19, 2014

Abstract:
In silico modeling, using Psipred and ExPASy servers was employed to determine the structural elements of Bcr-Abl oncoprotein ($p210^{BCR-ABL}$) isoforms, b2a2 and b3a2, expressed in Chronic Myelogenous Leukemia (CML). Both these proteins are tyrosine kinases having masses of 210-kDa and differing only by 25 amino acids coded by the b3 exonand an amino acidsubstitution (Glu903Asp). The secondary structure elements of the two proteins show differences in five α-helices and nine β-strands which relates to differences in the SH_3, SH_2, SH_1 and DNA-binding domains. These differences can result in different roles played by the two isoforms in mediating signal transduction during the course of CML.

Background:
Chronic Myelogenous Leukemia (CML) develops when a single, hematopoietic stem cell acquires a Philadelphia (Ph) chromosome carrying the *BCR-ABL* fusion oncogene which gives its progeny an advantage for proliferation over normal RBCs and allows the Ph-positive clone gradually to displace normal RBCs during hematopoiesis **[1, 2, 3, 4].** The abnormal Ph chromosome is produced by the translocation between chromosome 9 and 22 **(Figure 1a).**The major consequence of Philadelphia translocation is the fusion of the *ABL* gene on chromosome 9 with the *BCR* gene on chromosome 22 **[5].** The *BCR-ABL* fusion oncogene encodes new fusion proteins of 190, 210 and 230 kDa molecular weight **[6, 7].** The $p210^{BCR-ABL}$ isoforms have an increased level of tyrosine kinase activity, which is important for the development of the disease **[8].** The production of fusion proteins increases the diversity of protein-protein binding domains associated with tyrosine kinase activity.

The normal product of the *BCR* gene is a 160 kDa ($p160^{BCR}$) cytosolic phosphoprotein whose physiologic role is not clearly defined. It has been shown to form cytoplasmic complexes with $p210^{BCR-ABL}$ in Ph-positive CML cells, as well as with a 53 kDa protein of unknown function in both Ph-positive and Ph-negative cell lines **[9, 10].** The sequences encoded by the first exon of *BCR* gene are responsible for the $p160^{BCR}$ serine/threonine kinase activity **[11].** The N-terminus of $p160^{BCR}$ comprises a coiled–coil domain that allows dimer formation *in vivo*. At the center, are DBL-like **[12]** and Pleckstrin-homology **[13]** domains that stimulate the exchange of guanidine triphosphate (GTP) for guanidine diphosphate (GDP) and activate transcription factors such as NF-kB. The C-terminus of $p160^{BCR}$ contains GTPase activity for Rac. In addition, $p160^{BCR}$ can be phosphorylated at several tyrosine residues, such as Tyrosine 177, which binds Grb-2, an adapter protein involved in the activation of the Ras pathway.

The normal product of the *ABL* gene is a 145-kDa protein ($p145^{ABL}$) **[14].** It is a protein tyrosine kinase that is involved in cell differentiation, cell division, cell adhesion, and stress response. The activity of $p145^{ABL}$ is regulated by its SH_3 domain, and deletion of the SH_3 domain turns it into an

oncoprotein. The Philadelphia translocation results in the head-to-tail fusion of $p145^{ABL}$ to $p160^{BCR}$ **[15].**

The two $p210^{BCR-ABL}$ onco-protein isoforms, b2a2 and b3a2 **(Figure 1b)**, formed by the head-to-tail fusion of $p160^{BCR}$ and $p145^{ABL}$ proteins, differ by a 25 amino acid insertion coded by the b3 exon and a Glu903Asp substitution between b2a2 and b3a2. Several studies have examined whether the type of fusion transcript has any influence on the clinical outcome **[16, 17, 18, 19, 20, 21]**. However, the data remains controversial. In fact, several groups did not succeed in demonstrating any such correlation **[6, 22, 23].** A study has shown a correlation between the b3a2 transcript and a higher platelet count at diagnosis in a group of CML patients **[21].**

Recently, the crystal structure of the oligomerization domain at the N-terminus of both b2a2 and b3a2 (residues 1–72) has been published. The investigators have reported a novel mode of oligomer formation which involves dimerization of two monomers by swapping of N-terminus helices and by formation of an antiparallel coiled coil between C-terminus helices. The two dimers then stack onto each other to form a tetramer **[24].** In the present study, we performed a comparison of the structural elements of $p210^{BCR-ABL}$ protein isoforms, b2a2 and b3a2. For this purpose, Psipred and ExPASy servers were employed **[25, 26].**

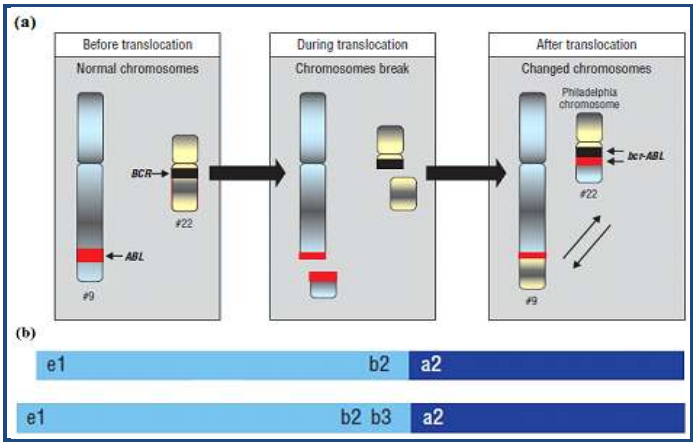


Figure 1: a) Shown is the Philadelphia chromosome (also called derivative 22) produced by reciprocal translocation **[4]**; **b)** Schematic showing the different lengths of the two p210 BCR-ABL protein isoforms (b2a2 and b3a2) expressed by the Philadelphia chromosome.

Methodology:

The amino acid sequences of the $p210^{BCR-ABL}$ protein isoforms, b2a2 and b3a2, were subjected to computer predictive analysis in order to reveal possible differences in terms of secondary structure and tertiary structure, Different tools and database were used for molecular modeling of these proteins such as GenBank-NCBI, Protein Data Bank, Psipred **[25]** and ExPASy servers **[26].** The sequence of the proteins was retrieved in FASTA format from NCBI database for homology modeling. The procedure of homology modeling procedure comprised three sequential steps: (i) template selection, (ii) target template alignment, and (iii) model building. The secondary structures of both b2a2 and b3a2 were predicted using Psipred server. The tertiary structures were predicted using the ExPASy server.

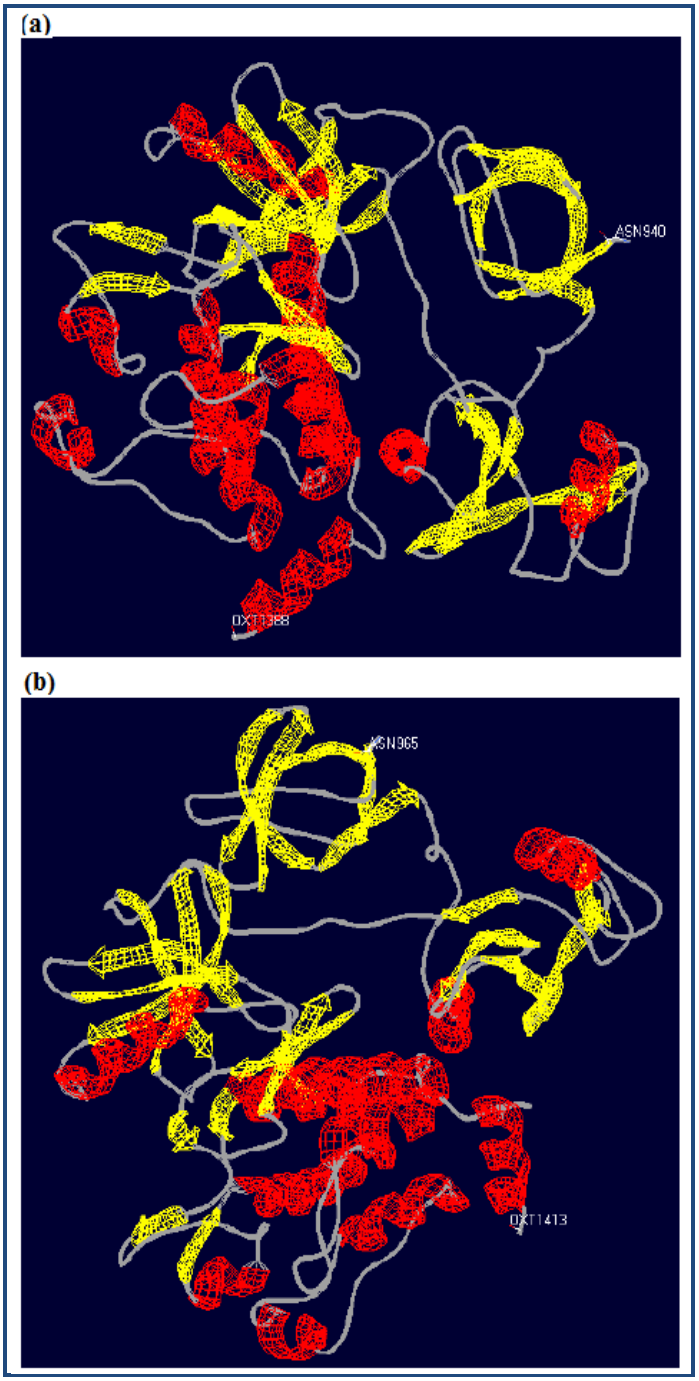


Figure 2: (a, b) Tertiary structures of b2a2 and b3a2 proteins obtained via the ExPASy server. The structures show that both the proteins possess α-helical and β-sheet domains. The b3a2 protein shows a greater amount of α-helical structure.

Results:

The use of Psipred server **[25]** provided the secondary structure elements of both b2a2 and b3a2 isoforms. The amino acid sequence of b2a2 has 2006 residues and b3a2 has 2031 residues. They differ by 26 amino acids. There is an amino acid substitution (Glu903Asp) and a 25 amino acid insertion just before Ala 904 in the sequence of b3a2. The secondary structure of b2a2 comprises 48 α-helices (α1 to α48) and 37 β-strands (β1 to β37). The b3a2 isoform also contains similar secondary structure elements along with an additional α-helix (α') and two β-strands (β' and β''). A short β-strand, β_{33} is also missing

in b3a2 **(Schematic 1, supplementary material).** In total, there are differences in five α-helices and nine β-strands of the two proteins which relates to differences in the SH_3, SH_2, SH_1 and DNA-binding domains **Table 1 (see supplementary material**). The tertiary structures of the two proteins were obtained via the ExPASy server **[26].** The tertiary structures show that both the proteins possess α-helical and β-sheet domains **(Figures 2a & b).** The tertiary structure of b3a2 protein reveals that it has greater amount of α-helical content.

Discussion:

The oncogenic potential of $p210^{BCR\text{-}ABL}$ protein isoforms is due to the fact that the normally regulated tyrosine kinase activity of the ABL protein ($p145^{ABL}$) becomes unregulated in both b2a2 and b3a2 isoforms. ABL proteins are non-receptor tyrosine kinases that have important roles in signal transduction and the regulation of cell growth **[27].** At the N-terminus, there are three SRC homology domains (SH_3, SH_2 and SH_1). SH_2 and SH_3 domains regulate the tyrosine kinase function of ABL protein and SH_1 domain contains the tyrosine kinase activity of ABL protein. SH_3 has a negative regulatory effect on the tyrosine kinase function. Deletion of SH_3 or mutation in SH_3 facilitates tyrosine kinase activity of ABL protein **[28, 29,30].** Mutations in SH_2 decrease phospho-tyrosine binding activity and reduce transforming capacities of ABL protein **[31].** The C-terminus of ABL protein contains a DNA-binding domain, nuclear localization signals, and a binding site for actin **[32] (Figure 3).**

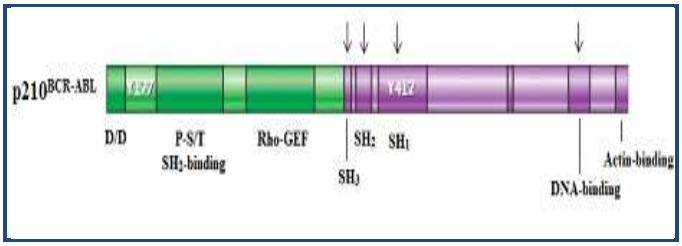


Figure 3: The domain organization of $p210^{BCR\text{-}ABL}$ proteins. The part that is derived from BCR gene is shown in *green* and the other half that is derived from *ABL* gene is shown in *purple*. The structural differences found in the SH_3, SH_2, SH_1 and DNA-binding domains of $p210^{BCR\text{-}ABL}$ are indicated by *arrows*. Also shown in the figure are oligomerization domain (D/D), actin-binding domain and the positions of Tyr 177 and Tyr 412 **[33].**

The structure of the oligomerization domain present at the N-terminus of both the $p210^{BCR\text{-}ABL}$ protein isoforms (b2a2 and b3a2) has been reported earlier **[24].** It consists of a short N-terminal helix (α_1), a flexible loop and a long C-terminal helix (α_2). The two α-helices, together form an N-shaped structure, with the loop allowing the two helices to assume a parallel orientation. The monomeric domains associate into a dimer through the formation of an antiparallel coiled coil between the α_2 helices and domain swapping of two α_1 helices, where one α_1 helix swings back and packs against the α_2 helix from the second monomer. The two dimers then associate into a tetramer **[24].** The oligomerization domain promotes clustering of a C-terminal actin-binding domain that cross-links actin filaments **[34, 35].** The two $p210^{BCR\text{-}ABL}$ protein isoforms can induce a redistribution of F-actin from the cortical cytoskeleton into aggregates **[35].** Oligomerization also plays a role in the activation of the $p210^{BCR\text{-}ABL}$ tyrosine kinase activity **[34].** A comparison of the secondary structure elements of both b2a2 and b3a2 with the structural elements determined by X-ray crystallography revealed that both similarities and differences. The length of α_1 and α_2 are longer in the crystal structure. Also, the position of α2 is shifted in the crystal structure. An additional β-strand (β_1) is present in the secondary structural elements obtained via Psipred.

The experimental data shows that $p210^{BCR\text{-}ABL}$ protein isoforms, b2a2 and b3a2, exhibit differences in their secondary structure elements **Table 1 (see supplementary material).** A twenty five amino acid insertion just before Ala 904 in the sequence of b3a2 shifts some of the secondary structure elements and also produces some additional ones **Schematic 1 (see supplementary material).** Both $p210^{BCR\text{-}ABL}$ proteins can cause pleiotropic effects on many signal transduction pathways that can affect cell survival, disease progression, and genomic stability. The signals controlled by $p210^{BCR\text{-}ABL}$ proteins are important for normal hematopoiesis. An earlier paper using the GOR, NnPredict and PHD methods to predict the structural elements of b2a2 and b3a2 did not find any difference between the two full-length protein types. However, a discrepancy was observed for the 25 amino acids coded by the b3 exon. In one case, it was seen as completely solvent accessible loop region (PHD method) whilst in the other cases (GOR and NnPredict methods) it comprised a short β-strand and a short α-helix **[36].**

Conclusion:

The $p210^{BCR\text{-}ABL}$ protein isoforms, b2a2 and b3a2 show differences in their secondary structure elements mainly due to the insertion of a 25 amino acid segment coded by the b3 exon in b3a2. In total, structural differences are found between the two proteins in five α-helices (α_{25}, α', α_{26}, α_{27} and α_{29}) and nine β-strands (β_{12}, β_{13}, β_{15}, β', β_{17}, β_{30}, β'', β_{34} and β_{35}). These differing structural elements are present in the SH_3, SH_2, SH_1 and DNA-binding domains which can result in different roles played by the two isoforms in mediating signal transduction during the course of Chronic Myelogenous Leukemia.

References:

[1] Eaves CJ & Eaves AC, *Baillieres Clin Haematol*. 1997 **10**: 233 [PMID: 9376662]
[2] Fialkow PJ *et al. Proc Natl Acad Sci USA*. 1967 **58**: 1468 [PMID: 5237880]
[3] Nowell P & Hungerford D, *Science* 1960 **142:** 1497
[4] Hazlehurst LA *et al. Cancer Control* 2009 **16**: 100 [PMID: 19337196]
[5] Shtivelman E *et al. Nature* 1985 **315**: 550 [PMID: 2989692]
[6] Melo JV, *Leukemia.* 1996 **10**: 751 [PMID: 8656667]
[7] Faderl S *et al. N Engl J Med*. 1999 **341**: 164 [PMID: 10403855]
[8] Daley GQ *et al. Science* 1990 **247**: 824 [PMID: 2406902]
[9] Campbell ML *et al. Oncogene* 1990 **5:** 773 [PMID: 2140598]
[10] Li W *et al. Oncogene* 1989 **4**: 127 [PMID: 2648252]
[11] Maru Y & Witte ON, *Cell.* 1991 **67**: 459 [PMID: 1657398]
[12] Aghazadeh B *et al. Nat Struct Biol.* 1998 **5**: 1098 [PMID: 9846881]
[13] Mayer BJ *et al. Cell* 1993 **73**: 629 [PMID: 8500161]
[14] Tani K, *J Biol Chem*. 2003 **278**: 21685 [PMID: 12672821]
[15] Shtivelman E *et al. Cell* 1986 **47**: 277 [PMID: 3021337]
[16] Mills KI *et al. Blood* 1991 **78**: 1155 [PMID: 1878582]
[17] Futaki M *et al. Leukemia Res*. 1992 **16**: 1071 [PMID: 1434743]
[18] Zaccaria A *et al. Br J Haematol*. 1993 **84**: 265 [PMID: 8398828]
[19] Rozman C *et al. Leukemia* 1995 **9**: 1104 [PMID: 7596178]

[20] Elliott SL *et al. Leukemia* 1995 **9**: 946 [PMID: 7596182]
[21] Shepherd P *et al. Br J Haematol*. 1995 **89**: 546 [PMID: 7734353]
[22] Shepherd PC *et al. Blood* 1992 **80**: 556 [PMID: 1378324]
[23] Opalka B *et al. Blood* 1992 **80**: 1854 [PMID: 1391949]
[24] Zhao X *et al. Nat Struct Biol*. 2002 **9**: 117 [PMID: 11780146]
[25] McGuffin LJ *et al. Bioinformatics* 2000 **16**: 404 [PMID: 10869041]
[26] Gasteiger E *et al. Nucleic Acids Res.* 2003 **31**: 3784 [PMID: 12824418]
[27] Wang JY, *Curr Opin Genet Dev*. 1993 **3**: 35 [PMID: 8453272]
[28] Jackson P & Baltimore D, *Embo J.* 1989 **8:** 449 [PMID: 2542016]
[29] Sawyers CL, *Cancer Surv.* 1992 **15**: 37 [PMID: 1451113]
[30] Franz WM *et al. Embo J.* 1989 **8**: 137 [PMID: 2496972]
[31] Gale RP *et al. Leukemia* 1993 **7**: 653 [PMID: 8464245]
[32] Chung SW *et al. Crit Rev Oncog*. 1996 **7**: 33 [PMID: 9109496]
[33] Goldman, JM & Melo JV, *New England J Med*. 2003 **349**: 1451
[34] McWhirter JR *et al. Mol Cell Biol.* 1993 **13**: 7587 [PMID: 8246975]
[35] McWhirter JR & Wang JY, *EMBO J*. 1993 **12:** 1533 [PMID: 8467803]
[36] Perego RA *et al. Eur J Cancer*. 2000 **36:** 1395 [PMID: 10899653]

Edited by P Kangueane
Citation: **Hai *et al.*** Bioinformation 10(3): 108-114 (2014)

BIOINFORMATION

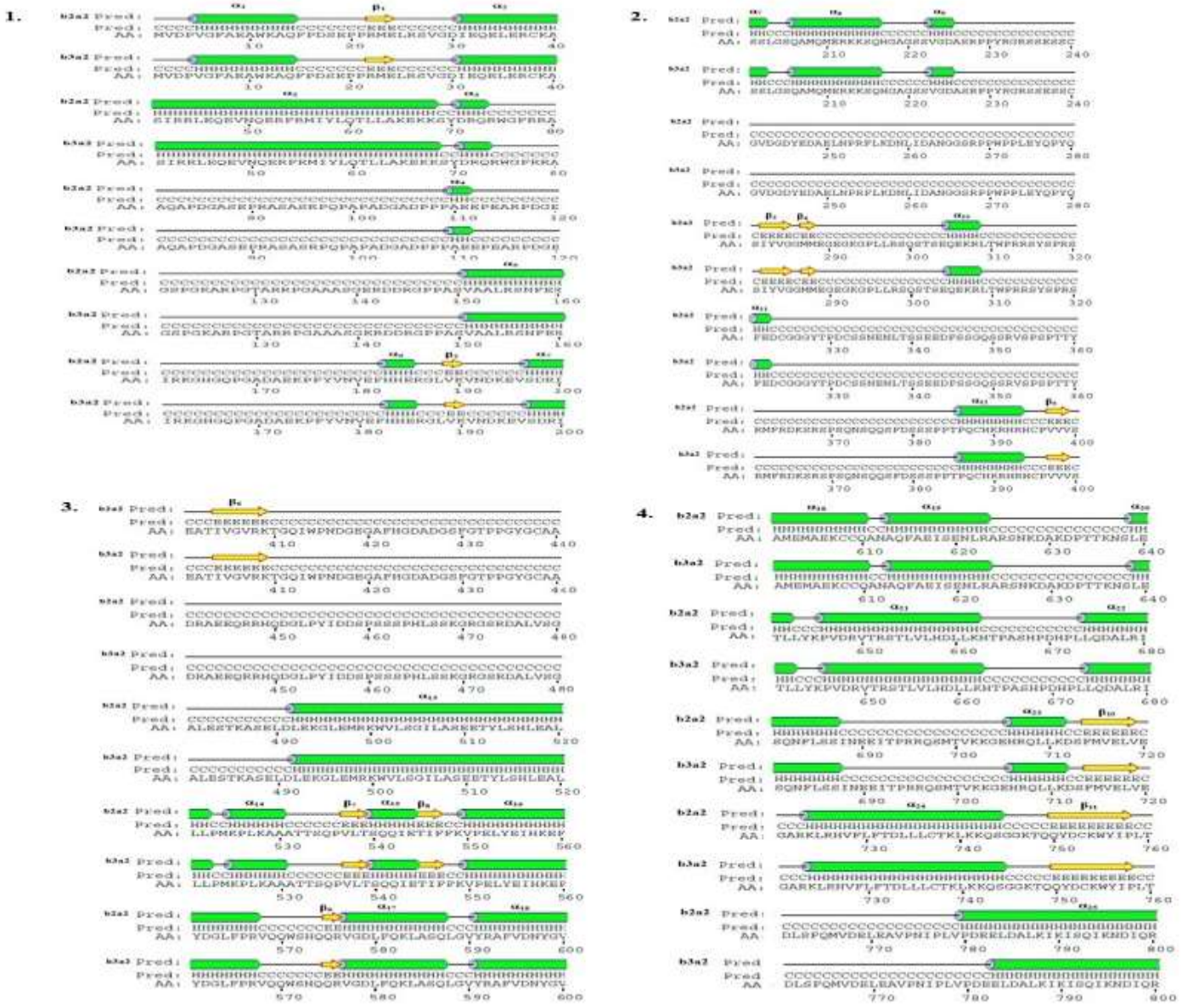
open access

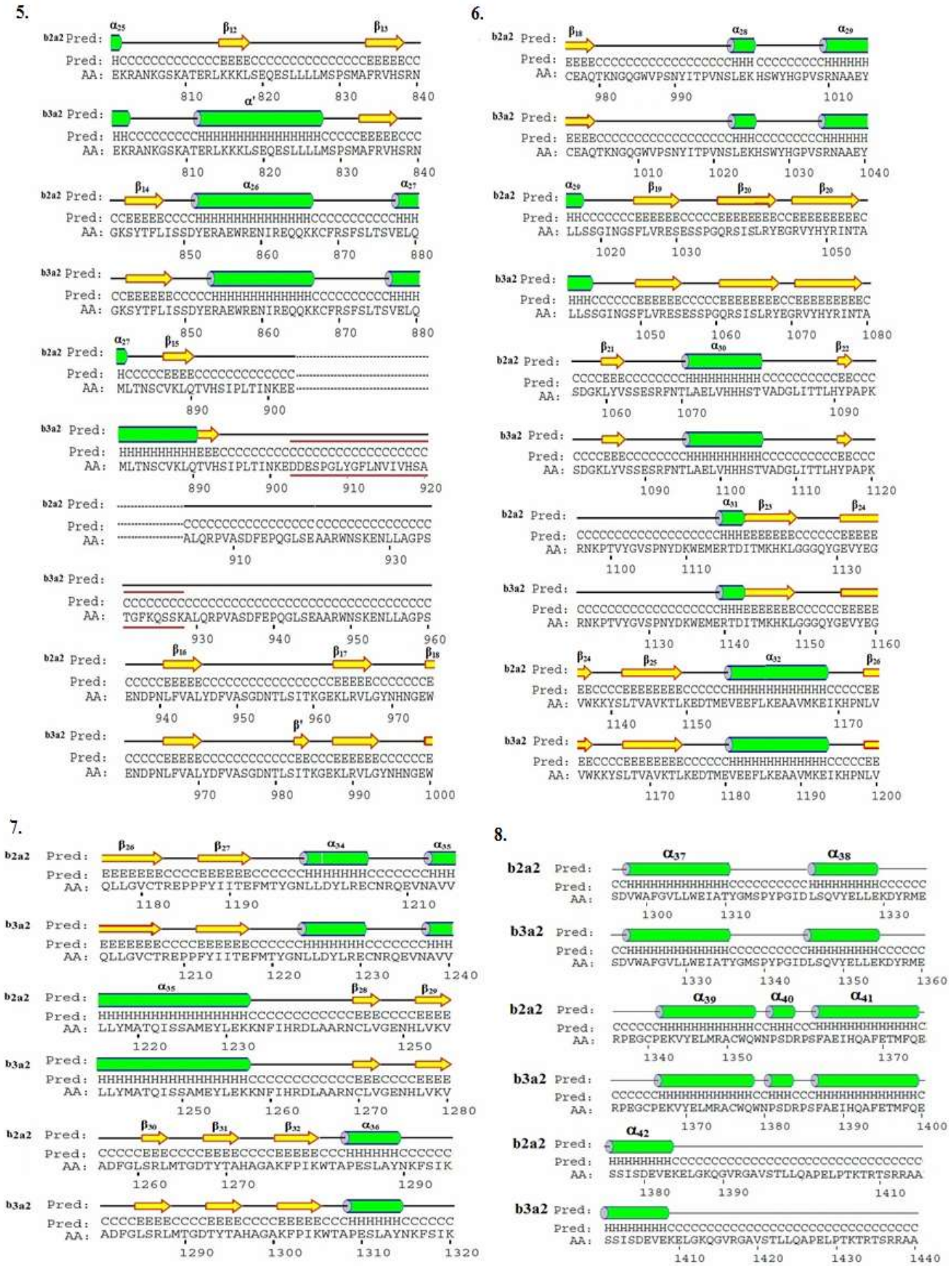
Supplementary material:

Table 1: Summary of structural differences between b2a2 and b3a2 isoforms.

Secondary Structure Element	Position of structural element in b2a2	Position of structural element in b3a2
α_{25}	Pro780 - Glu801	Glu 782 – Lys 802
α'	absent	Pro 812 – Glu 827
α_{26}	Tyr 852 – Lys 866	Arg 854 – Lys 866
α_{27}	Glu 878 – Met 881	Val 877 - Gln 890
α_{29}	Arg 1010 – Leu 1017	Arg 1035 – Ser 1043
β_{12}	Lys 815 – Leu 818	absent
β_{13}	Phe 834 - Ser 838	Ala 833 - His 837
β_{15}	Val 887 - Gln 890	Thr 891 - His 893
β'	absent	Ile 983 - Thr 984
β_{17}	Lys 963 - Leu 967	Lys 988 - Gly 993
β_{30}	Ser 1261 - Leu 1263	Leu 1285 - Leu 1288
β''	absent	between Asn1671 and Arg1672
β_{34}	Ser 1735 - Pro 1738	Gly 1761 - Pro 1763
β_{35}	Lys 1755 - His 1756	absent

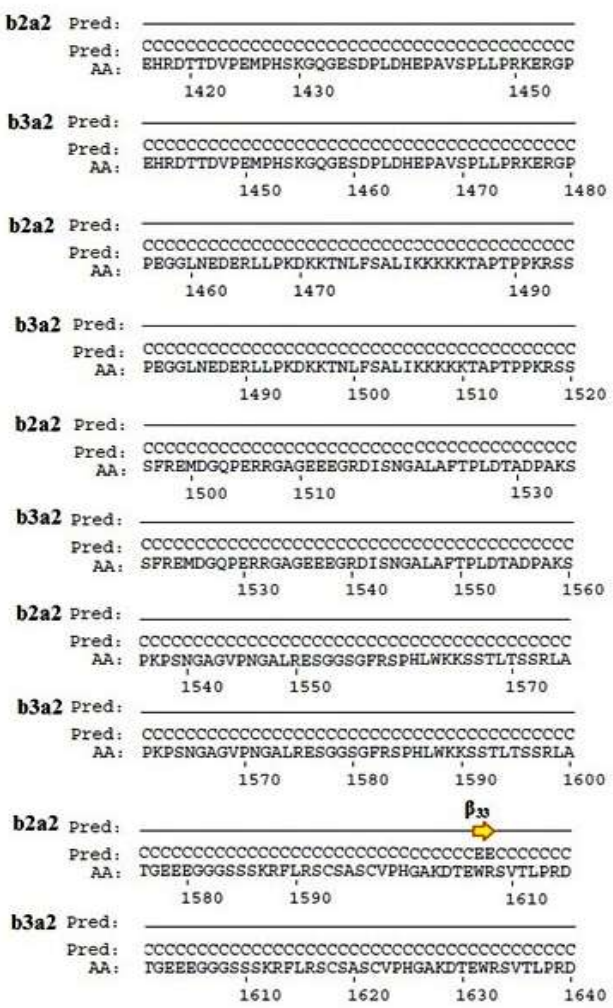
Schematic 1: Comparison of secondary structure elements of b2a2 and b3a2 protein isoforms obtained by the use of Psipred. The 25 amino acid insertion just before Ala 904 in the sequence of b3a2 is underlined in red (the panels are numbered 1-11 according to the order in which they appear).



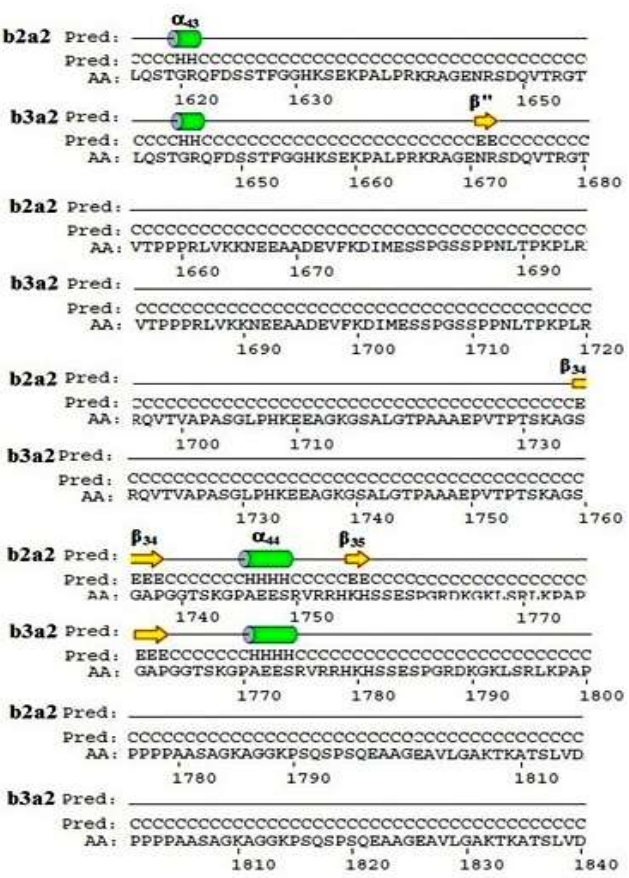


ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 10(3): 108-114 (2014)

9.



10.



11.