ORIGINAL ARTICLE

INTERNATIONAL JOURNAL OF RESEARCH IN PHARMACEUTICAL SCIENCES

Published by JK Welfare & Pharmascope Foundation   Journal Home Page: www.ijrps.com

Molecular effects of mutated amino acids involved in Transmembrane and Domain regions on the BCR oncogene protein using In silico techniques

Anuradha Manogharan[1], Regina Mary Rathina Samy[*1], Ramadevi Mohan[2]

[1]Department of Zoology, Auxilium College, Gandhi Nagar, Vellore-632006, Tamil Nadu, India
[2]Department of Integrative Biology, VIT University, Vellore-632014, Tamil Nadu, India

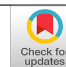
Article History:

Received on: 25 Jun 2020
Revised on: 25 Jul 2020
Accepted on: 27 Jul 2020

Keywords:

Chronic Myeloid Leukemia,
BCR gene,
Transmembrane,
insilico,
3D structure prediction,
TMRPRED,
Discovery Studio

ABSTRACT



BCR gene is expressed in patients with Philadelphia-positive Leukemias, known as chronic myeloid leukaemia (CML). Here, we focus on how the intramolecular domains and transmembrane segments are involved in the mutated sites of BCR. In this research work, we thoroughly analysed the transmembrane segments and the functional domains and predicted the 3D structure. We applied two kinds of techniques in our work. One is sequence-based, where we proved that the transmembrane segments in the functional domains contain the mutated sites. The second technique is structure-based, where we predicted the 3D structure of BCR gene-coded protein and visualised the transmembrane segments, which included the mutated sites. By using advanced molecular visualisation tools, the molecular structural properties of the respective transmembrane regions of amino acids will be determined. Both the techniques involved the use of advanced insilico tools and database. Our results elucidated that both the sequence and structure-based outcomes represented the identified transmembrane segments in the functional domains, which are potential candidates for drug docking studies. Hence, we finally concluded that this research work would play a vital role in clinical oncology and structure-based drug designing. Our research work is the first attempt to prove that potential drug binding sites are present in BCR oncogene-protein using insilico techniques. The results of this research investigation form a basic foundation for structure-based drug designing.

*Corresponding Author

Name: Regina Mary Rathina Samy
Phone: +91 9952355724
Email: fabiregi31@gmail.com

ISSN: 0975-7538

DOI: https://doi.org/10.26452/ijrps.v11i4.3234

Production and Hosted by

IJRPS | www.ijrps.com

© 2020 | All rights reserved.

INTRODUCTION

In recent years, leukaemia is a major challenge in the field of medicine. In our research, we focus on the BCR gene. Heisterkamp and his colleagues established the structural organisation of the BCR gene, which consists of 23 exons and is positioned in a region of about 135 kb on chromosome 22 (Heisterkamp *et al.*, 1985). The first exon contains a unique serine/threonine kinase activity and no less than 2 SH2 binding sites. Stam and his co-workers showed that the BCR gene is positioned with its 5-prime end toward the centromere of chromosome 22 (Stam *et al.*, 1987).

It was ascertained that the BCR gene consists of 23

BCR exons with putative alternative BCR first and second exons (Chissoe *et al.*, 1995). It was shown that BCR, when purified, contains autophosphorylation activity and transphosphorylation activity for several protein substrates (Maru and Witte, 1991). It was found out that for this new phosphotransferase activity, a region encoded by the first exon was essential.

Zhao and his researchers reported that impairment of hematopoietic stem cell renewal and reduced induction of chronic myelogenous leukaemia (CML; 608232), by the BCR-ABL1 oncoprotein were due to the loss of Smoothened (Smo; 601500) which is an essential component of the hedgehog pathway (600725) (Zhao *et al.*, 2009). Depletion of CML stem cells was caused by the loss of Smo, which propagated leukaemia, whereas constitutively CML stem cell number was increased, and active Smo accelerated disease.

Recently, with the efficient demonstration of generic strategies which stabilise and crystallise unstable, eukaryotic membrane proteins, we have probably entered the third phase of membrane protein structure analysis. For crystallisation, such unstable membrane proteins should be stabilised either by prompt reinsertion into a lipid bilayer, such as with 2D membrane crystals (Kühlbrandt, 1992) or with lipidic phase, detergent-free crystallisation methods (Nollert *et al.*, 1999), or by the inclusion of specific active-site ligands or inhibitors (Pebay-Peyroula *et al.*, 2003; Vedadi *et al.*, 2006; Toyoshima, 2008), or by systematic mutagenesis to create a protein which has increased intrinsic stability (Magnani *et al.*, 2008; Serrano-Vega *et al.*, 2008). This third post-genomic phase should enable the determination of the structure of any membrane protein or complex of interest.

```
>sp|P11274|BCR_HUMAN Breakpoint cluster region protein
OS=Homo sapiens OX=9606 GN=BCR PE=1 SV=2
MVDPVGFAEAWKAQFPDSEPPRMELRSVGDIEQELERCKASIRRLEQEVNQERFRMIYLQ
TLLAKEKKSYDRQRWGFRRAAQAPDGASEPRASASRPQPAPADGADPPPAEEPEARPDGE
GSPGKARPGTARRPGAAASGERDDRGPPASVAALRSNFERIRKGHGQPGADAEKPFYVNV
EFHHERGLVKVNDKEVSDRISSLGSQAMQMERKKSQHGAGSSVGDASRPPYRGRSSESSC
GVDGDYEDAELNPRFLKDNLIDANGGSRPPWPPLEYQPYQSIYVGGMMEGEGKGPLLRSQ
STSEQEKRLTWPRRSYSPRSFEDCGGGYTPDCSSNENLTSSEEDFSSGQSSRVSPSPTTY
RMFRDKSRSPSQNSQQSFDSSSPPTPQCHKRHRHCPVVVSEATIVGVRKTGQIWPNDGEG
AFHGDADGSFGTPPGYGCAADRAEEQRRHQDGLPYIDDSPSSSPHLSSKGRGSRDALVSG
ALESTKASELDLEKGLEMRKWVLSGILASEETYLSHLEALLLPMKPLKAAATTSQPVLTS
QQIETIFFKVPELYEIHKEFYDGLFPRVQQWSHQQRVGDLFQKLASQLGVYRAFVDNYGV
AMEMAEKCCQANAQFAEISENLRARSNKDAKDPTTKNSLETLLYKPVDRVTRSTLVLHDL
LKHTPASHPDHPLLQDALRISQNFLSSINEEITPRRQSMTVKKGEHRQLLKDSFMVELVE
GARKLRHVFLFTDLLLCTKLKKQSGGKTQQYDCKWYIPLTDLSFQMVDELEAVPNIPLVP
DEELDALKIKISQIKNDIQREKRANKGSKATERLKKKLSEQESLLLLMSPSMAFRVHSRN
GKSYTFLISSDYERAEWRENIREQQKKCFRSFSLTSVELQMLTNSCVKLQTVHSIPLTIN
KEDDESPGLYGFLNVIVHSATGFKQSSNLYCTLEVDSFGYFVNKAKTRVYRDTAEPNWNE
EFEIELEGSQTLRILCYEKCYNKTKIPKEDGESTDRLMGKGQVQLDPQALQDRDWQRTVI
AMNGIEVKLSVKFNSREFSLKRMPSRKQTGVFGVKIAVVTKRERSKVPYIVRQCVEEIER
RGMEEVGIYRVSGVATDIQALKAAFDVNNKDVSVMMSEMDVNAIAGTLKLYFRELPEPLF
TDEFYPNFAEGIALSDPVAKESCMLNLLLSLPEANLLTFLFLLDHLKRVAEKEAVNKMSL
HNLATVFGPTLLRPSEKESKLPANPSQPITMTDSWSLEVMSQVQVLLYFLQLEAIPAPDS
KRQSILFSTEV
```

Figure 1: FASTA sequence of BCR_HUMAN-UNIPROT database

PRED-TMR: Prediction of Transmembrane regions in proteins

Sequence: P11274|BCR_HUMAN

```
              1         2         3         4         5         6
     123456789012345678901234567890123456789012345678901234567890
         .    |    .    |    .    |    .    |    .    |    .    |
0000 MVDPVGFAEAWKAQFPDSEPPRMELRSVGDIEQELERCKASIRRLEQEVNQERFRMIYLQ
0060 TLLAKEKKSYDRQRWGFRRAAQAPDGASEPRASASRPQPAPADGADPPPAEEPEARPDGE
0120 GSPGKARPGTARRPGAAASGERDDRGPPASVAALRSNFERIRKGHGQPGADAEKPFYVNV
0180 EFHHERGLVKVNDKEVSDRISSLGSQAMQMERKKSQHGAGSSVGDASRPPYRGRSSESSC
0240 GVDGDYEDAELNPRFLKDNLIDANGGSRPPWPPLEYQPYQSIYVGGMMEGEGKGPLLRSQ
0300 STSEQEKRLTWPRRSYSPRSFEDCGGGYTPDCSSNENLTSSEEDFSSGQSSRVSPSPTTY
0360 RMFRDKSRSPSQNSQQSFDSSSPPTPQCHKRHRHCPVVVSEATIVGVRKTGQIWPNDGEG
0420 AFHGDADGSFGTPPGYGCAADRAEEQRRHQDGLPYIDDSPSSSPHLSSKGRGSRDALVSG
0480 ALESTKASELDLEKGLEMRKWVLSGILASEETYLSHLEALLLPMKPLKAAATTSQPVLTS
0540 QQIETIFFKVPELYEIHKEFYDGLFPRVQQWSHQQRVGDLFQKLASQLGVYRAFVDNYGV
0600 AMEMAEKCCQANAQFAEISENLRARSNKDAKDPTTKNSLETLLYKPVDRVTRSTLVLHDL
0660 LKHTPASHPDHPLLQDALRISQNFLSSINEEITPRRQSMTVKKGEHRQLLKDSFMVELVE
0720 GARKLRHVFLFTDLLLCTKLKKQSGGKTQQYDCKWYIPLTDLSFQMVDELEAVPNIPLVP
0780 DEELDALKIKISQIKNDIQREKRANKGSKATERLKKKLSEQESLLLLMSPSMAFRVHSRN
0840 GKSYTFLISSDYERAEWRENIREQQKKCFRSFSLTSVELQMLTNSCVKLQTVHSIPLTIN
0900 KEDDESPGLYGFLNVIVHSATGFKQSSNLYCTLEVDSFGYFVNKAKTRVYRDTAEPNWNE
0960 EFEIELEGSQTLRILCYEKCYNKTKIPKEDGESTDRLMGKGQVQLDPQALQDRDWQRTVI
1020 AMNGIEVKLSVKFNSREFSLKRMPSRKQTGVFGVKIAVVTKRERSKVPYIVRQCVEEIER
1080 RGMEEVGIYRVSGVATDIQALKAAFDVNNKDVSVMMSEMDVNAIAGTLKLYFRELPEPLF
1140 TDEFYPNFAEGIALSDPVAKESCMLNLLLSLPEANLLTFLFLLDHLKRVAEKEAVNKMSL
1200 HNLATVFGPTLLRPSEKESKLPANPSQPITMTDSWSLEVMSQVQVLLYFLQLEAIPAPDS
1260 KRQSILFSTEV
```

(predicted transmembrane regions are indicated in blue)

Detail of predicted transmembrane segments:

Number of predicted transmembrane segments: 1

#TM	begin	end
1	1167	1183

Figure 2: Transmembrane region prediction of BCR_HUMAN

```
>BCR_HUMAN_DOMAIN
VKIAVVTKRERSKVPYIVRQCVEEIERRGMEEVGIYRVSGVATDIQALKAAFDVNNKDVS
VMMSEMDVNAIAGTLKLYFRELPEPLFTDEFYPNFAEGIALSDPVAKESCMLNLLLSLPE
ANLLTFLFLLDHLKRVAEKEAVNKMSLHNLATVFGPTLLRPSEKESKLPANPSQPITMTD
SWSLEVMSQVQVLLY
```

Figure 3: BCR_HUMAN domain with transmembrane region

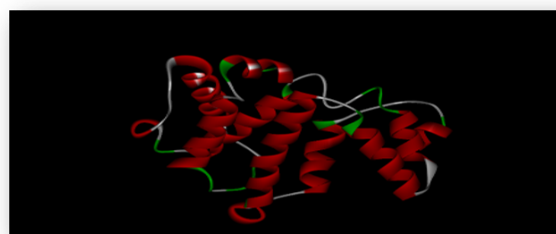


Figure 4: 3D structure prediction of BCR protein visualized by Discovery studio software

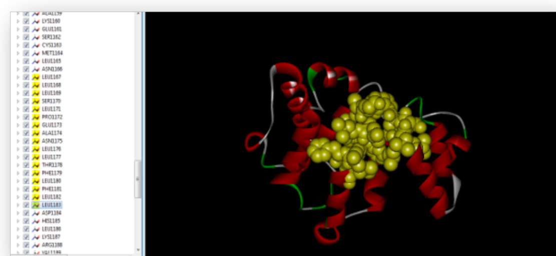


Figure 5: 3D structure of BCR protein showing Transmembrane regions

Table 1: Gene summary - BCR_HUMAN

Gene name	Chrloc	UNIPROT ID	Nucleotide length	aa length	OMIM ID	Gene ID
BCR_HUMAN Breakpoint cluster region protein	22	P11274	3816	1271	151410	613

Table 2: Domain analysis - BCR_HUMAN

Protein target	Domain length	Ref : Clinical Evidence
BCR_HUMAN Breakpoint cluster region protein	1054-1248	Chuang T.H et al., 1995

Table 3: Mutation analysis of BCR- Clinical evidence

Tm prediction Length amino acids (BCR_HUMAN Breakpoint cluster region protein)	SNP IDS (evidence 1)
1167	rs767469720
1168	COSM4917716
1169	COSM4665572
1170	-
1171	-
1172	rs1447313761
1173	rs747409015
1178	COSM3842279
1179	rs781765958
1180	rs746444301
1183	rs7704124339

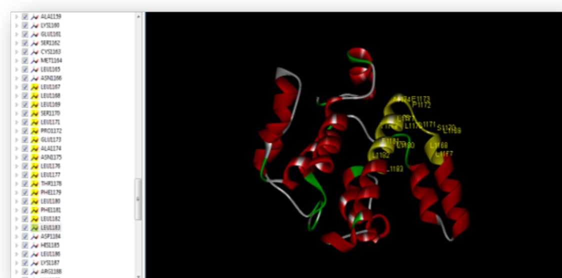


Figure 6: Residues involved in Transmembrane segments of BCR

MATERIALS AND METHODS

We, first, retrieved the gene-coded protein sequence of BCR based on various clinical literature studies (OMIM, ID, Uniprot) (). Next, the retrieved sequence was applied into TMRPRED (Pasquier *et al.*, 1999) tool to identify the potential transmembrane regions present in the BCR. Finally, the 3D structure of the retrieved sequence was studied using the CPH model 3.0 server (Nielsen *et al.*, 2010; Lund *et al.*, 2002). CPHmodels-3.0 is a webserver which predicts the 3D structure of the query protein (BCR (Breakpoint cluster region protein) in our case) by using single template homology modelling. The modelled protein structure was visualised using a molecular visualisation tool called Discovery Studio Software.

RESULTS AND DISCUSSION

The length of the amino acid sequence of human BCR is 1271 amino acids, and its various details were retrieved, as shown in Table 1. The FASTA format of BCR (Breakpoint cluster region protein) was given in Figure 1.

The analysis of Transmembrane segments by PRED-TMR in the BCR human protein sequence revealed the start position as 1167 and end position as 1183. The domain length of potential domain regions of

BCR was reported to be 1054-1248 (Chuang *et al.*, 1995). Chuang and his co-workers also explained that the region involved in binding to ABL1-SH2-domain was rich in serine residues, and prior to SH2 binding, it needs to be Ser/Thr phosphorylated. This region plays a vital role in the activation of the ABL1 tyrosine kinase and transforming potential of the chimeric BCR-ABL oncogene.

The domain length of human BCR was shown in Table 2 and Figure 2. The DH domain was involved in an interaction with CCPG1. The amino terminus consists of intrinsic kinase activity. The central Dbl homology (DH) domain acts as guanine nucleotide exchange factor (GEF) which modulates the GTPases CDC42, RHOA and RAC1, and promotes the conversion of CDC42, RHOA and RAC1 from the GDP-bound to the GTP-bound form. The C-terminus is a Rho-GAP domain which promotes GTP hydrolysis by RAC1, RAC2 and CDC42. The protein has a special structure having two opposing regulatory activities towards small GTP-binding proteins.

Table 3 showed that the identified Transmembrane segments fall within the domain regions of BCR (Chuang *et al.*, 1995). The research proved that the identified segments were potentially involved in the various mutations of BCR. The identified SNPs were found to be potential targets for drugs. These SNPs are the over-expression of the BCR gene (Diekmann *et al.*, 1991).

The human BCR protein-containing transmembrane region (LLSLPEANLLTF

LFLL) (1167-1183) that was found within the domain region (1054-1248) was shown in Figure 3.

3D Structure Prediction

The three dimensional structure of the protein that was predicted using CPH model server was shown in the Figure 4.

The Figure 4 showed the secondary structure view of BCR in which red colour indicates alpha-helix, blue indicates beta sheets, green indicates turns and white indicates coils.

In the above depicted Figure 5, we can visualize the transmembrane segments of BCR protein representedin space-fill model.

As shown in Figure 5 and Figure 6, amino acids involved in the transmembrane segments of BCR were labelled and found to be L1167, L1168, L1169, S1170, L1171, P1172, E1173, A1174, N1175, L1176, L1177, T1178, F1179, L1180, F1181, L1182 and L1183.

The breakpoints in CML to sub-bands 22q11.21 and 9q34.1 were located by Prakash and Yunis (1984). Even if the breakpoint's position in chromosome 9 is quite variable, the breakpoint in chromosome 22 is grouped in an area called BCR for breakpoint cluster region. Shtivelman and his group referred to BCR as a gene and declared that the ABL oncogene is transferred into the BCR gene on chromosome 22. They found that an 8-kb RNA which was specific to CML is a fused transcript of the two genes. It is probably the fused protein which is involved in the malignant process (Shtivelman *et al.*, 1985).

Our results clearly showed that transmembrane segments are present within the functional domain of Oncogene BCR- Breakpoint cluster region protein. We also proved it with the help of 3D structures and sequences. The identified regions could act as potential drug binding sites for the upcoming drugs.

CONCLUSION

In this research, we found out that Transmembrane amino acid segments act as potential candidates for drug docking studies. These segments are directly involved in various mutations. The predicted 3D structure clearly showed the target drug-binding sites (BCR-Breakpoint cluster region protein). Hence, we finally conclude that our results play a vital role in structure-based drug designing and clinical oncology studies.

ACKNOWLEDGEMENT

I want to thank Dr Balaji Munivelan (bioinfobalaji@gmail.com) for his assistance towards *Insilico* works in this paper.

Conflict of Interest

The authors declare that they have no conflict of interest for this study.

Funding Support

The authors declare that they have no funding support for this study.

REFERENCES

Chissoe, S. L., Bodenteich, A., Wang, Y.-F., Wang, Y.-P., Burian, D., Clifton, S. W., Crabtree, J., Freeman, A., Iyer, K., Jian, L., Ma, Y., McLaury, H.-J., Pan, H.-Q., Sarhan, O. H., Toth, S., Wang, Z., Zhang, G., Heisterkamp, N., Groffen, J., Roe, B. A. 1995. Sequence and Analysis of the Human ABL Gene, the BCR Gene, and Regions Involved in the Philadelphia Chromosomal Translocation. *Genomics*, 27(1):67–82.

Chuang, T. H., Xu, X., Kaartinen, V., Heisterkamp, N., Groffen, J., Bokoch, G. M. 1995. Abr and Bcr

are multifunctional regulators of the Rho GTP-binding protein family. *Proceedings of the National Academy of Sciences*, 92(22):10282–10286.

Diekmann, D., Brill, S., Garrett, M. D., Totty, N., Hsuan, J., Monfries, C., Hall, C., Lim, L., Hall, A. 1991. Bcr encodes a GTPase-activating protein for p21rac. *Nature*, 351(6325):400–402.

Heisterkamp, N., Stam, K., Groffen, J., de Klein, A., Grosveld, G. 1985. Structural organization of the bcr gene and its role in the Ph′ translocation. *Nature*, 315(6022):758–761.

Kühlbrandt, W. 1992. Two-dimensional crystallization of membrane proteins. *Quarterly Reviews of Biophysics*, 25(1):1–49.

Lund, O., Nielsen, M., Lundegaard, C., Worning, P. 2002. CPHmodels 2.0: X3M a Computer Program to Extract 3D Models. Abstract at the CASP5 conference A 102.

Magnani, F., Shibata, Y., Serrano-Vega, M. J., Tate, C. G. 2008. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proceedings of the National Academy of Sciences*, 105(31):10744–10749.

Maru, Y., Witte, O. N. 1991. The BCR gene encodes a novel serine/threonine kinase activity within a single exon. *Cell*, 67(3):459–468.

Nielsen, M., Lundegaard, C., Lund, O., Petersen, T. N. 2010. CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Research*, 38(suppl_2):W576–W581.

Nollert, P., Royant, A., Pebay-Peyroula, E., Landau, E. M. 1999. Detergent-free membrane protein crystallization. *FEBS Letters*, 457(2):205–208.

Pasquier, C., Promponas, V. J., Palaios, G. A., Hamodrakas, J. S., Hamodrakas, S. J. 1999. A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. *Protein Engineering, Design and Selection*, 12(5):381–385.

Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G. J.-M., Brandolin, G. 2003. Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature*, 426(6962):39–44.

Prakash, O., Yunis, J. J. 1984. High resolution chromosomes of the t(9;22) positive leukemias. *Cancer Genetics and Cytogenetics*, 11(4):361–367.

Serrano-Vega, M. J., Magnani, F., Shibata, Y., Tate, C. G. 2008. Conformational thermostabilization of the 1-adrenergic receptor in a detergent-resistant form. *Proceedings of the National Academy of Sciences*, 105(3):877–882.

Shtivelman, E., Lifshitz, B., Gale, R. P., Canaani, E. 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature*, 315(6020):550–554.

Stam, K., Heisterkamp, N., Reynolds, F. H., Groffen, J. 1987. Evidence that the phl gene encodes a 160,000-dalton phosphoprotein with associated kinase activity. *Molecular and Cellular Biology*, 7(5):1955–1960.

Toyoshima, C. 2008. Structural aspects of ion pumping by Ca2+-ATPase of sarcoplasmic reticulum. *Archives of Biochemistry and Biophysics*, 476(1):3–11.

Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., Edwards, A. M. 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences*, 103(43):15835–15840.

Zhao, C., Chen, A., Jamieson, C. H., Fereshteh, M., Abrahamsson, A., Blum, J., Kwon, H. Y., Kim, J., Chute, J. P., Rizzieri, D., Munchhof, M., VanArsdale, T., Beachy, P. A., Reya, T. 2009. Hedgehog signalling is essential for maintenance of cancer stem cells in myeloid leukaemia. *Nature*, 458(7239):776–779.