Computer-Aided Sequence Analysis and Homology Modelling Of Breast Cancer Receptor ERBB2

[1]Dr.T. V. Sai Krishna, Dr. [2]A. Yesubabu, [3]Dr. Deepak Nedunuri, [4]Ch. Madhava Rao
[1]Professor & Head, Department of CSE, QIS Institute of Technology, Ongole, Andhra Pradesh.
[2]Professor & Head, Department of CSE, Sir C R Reddy College of Engineering, Eluru, Andhra Pradesh.
[3]Associate Professor, Department of CSE, Sir C R Reddy College of Engineering, Eluru, Andhra Pradesh.
[4]Assistant Professor, Department of ECE, Sir C R Reddy College of Engineering, Eluru, Andhra Pradesh.

Abstract **-** ErbB2, a transmembrane growth factor receptor, is a member of subclass I of the superfamily of receptor tyrosine kinases. Overexpression of ErbB2 is observed in many breast cancer patients in whom normal cells turn out to be cancerous, and lead to poor prognosis. An attempt was made to study the influence of scoring matrix on alignment as well as to perform a homology derived model for the protein of interest. Hence, three different programs such as BLAST, FASTA and WU BLAST2 are employed in this work. When compared with other two programs, FASTA (with PAM 120 matrix) revealed a clear evolutionary relationship. A scan against Swiss-Prot protein sequence database for receptor protein tyrosine kinase erbb2 resulted in 5 hits, of which, Q60553 was selected for Sequence analysis and Homology modeling. The analysis identified a relevant homology with PDB protein 1N8Y (89.96% identity and 96.87% similarity). A list of similar PDB hits from FASTA program resulted in moderate to good alignments. Homology Modeling was initiated with Modeler 9v1 run in Windows operating system. Out of five generated models, fifth model was considered as the best optimized and superimposed model. It exhibited the lowest energy (2392.54 kcal/mol) and low RMSD value (0.5Aº) than the remaining models. By using Ramchandran plot, the probable number of residues that appear in different regions of the plot are identified. The number of residues in the disallowed region was seven in case of 1N8Y whereas ten in case of the modelled protein. It was observed that these residues are on the surface of the structure and they may have little or no effect on the protein.

KeyWords: Sequence analysis, homology modelling, breast cancer, BLAST, FASTA

I. INTRODUCTION

HER2 (also known as Neu, ErbB2) is a member of the epidermal growth factor receptor (EGFR; also known as ErbB) family of receptor tyrosine kinases, which in humans includes HER1 (EGFR, ERBB1), HER2, HER3 (ERBB3) and HER4 (ERBB4). ErbB receptors are essential mediators of cell proliferation and differentiation in the developing embryo and in adult tissues and their inappropriate activation is associated with the development and severity of many cancers. Over expression of HER2 is found in 20-30% of human breast cancers, and correlates with more aggressive tumors and a poorer prognosis. The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to drive breast tumor cell proliferation [1]. Over expression of the erbB-1 (EGFR, epidermal growth factor receptor) and erbB-2 (HER2/neu) proteins contributes to the aggressive behavior of malignant tumors originating from the endometrium. The level of expression appeared to be significantly higher in the malignant tumors as compared to the benign ones for erbB-1 and for erbB-2 [2]. Anticancer therapies targeting ErbB receptors have shown promise, and a monoclonal antibody against HER2, Herceptin (also known as trastuzumab), is currently in use as a treatment for breast cancer. Herceptin binds to the juxtamembrane region of HER2, identifying this site as a target for anticancer therapies [3]. Here, this paper reports sequence analysis and homology modelling of ErbB2 receptor using various computational tools and methodologies.

II. MATERIALS AND METHODS

2.1 Swiss-Prot protein sequence database

Protein sequence database was scanned for receptor protein tyrosine kinase erbb2 sequences and from the resulted sequences ERBB2 (Q60553) [4] was selected for Homology modeling.

2.2 Sequence Alignment tool- BLAST

BLAST pair wise sequence alignment tool [5] was utilized to perform a blast search against PDB protein structure database using default matrix PAM 120.

2.3 PDB protein structure database

This database [6] was used to download structural sequences in PDB format in order to perform homology modeling. The structural data, summary information, sequence length, X-ray parameters, Resolutions, Ramachandran plot and other factors were carefully studied.

2.4 Modeler 8v2 software

Homology modeling software Modeler [7] was used to build comparative homology model using Q60553 as target sequence and 1N8Y as template sequence. All the steps were performed in Window operating system.

2.5 Homology Modeling

Comparative models were constructed for Golden hamster erbb2 protein sequence (Q60553) to study the sequence in the structural context and to suggest site-directed mutagenesis experiments for elucidating specificity changes in this apparent case of convergent evolution of enzymatic specificity.

2.5.1 Protein structure modeling by satisfaction of spatial restraints

MODELLER is used for homology or comparative modeling of protein three dimensional structures. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints and can perform many additional tasks, including de novo modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexible define objective function, multiple alignment of protein sequences and/or structure, clustering, searching of sequence databases, comparison of protein structures, etc.

Step-1: Selection of target sequence:
Homology Modeling has been carried out using BLAST, FASTA, WU-BLAST2 out of which fasta is taken into consideration since it exhibits high %Identity, high % similarity, less no of gaps when compared to other entries. Thus to generate various models in homology modeling 1N8Y alignment with the query sequence Q60553 has been taken into consideration.

Query sequence
>Q60553|ERBB2_MESAU Receptor tyrosine-protein kinase erbB-2 - Mesocricetus auratus (Golden hamster).
MELAAWCGWGLLLALLSPGASGTQVCTGTDMKLRLP
ASPETHLDIVRHLYQGCQVVQGNL
ELTYLPANATLSFLQDIQEVQGYMLIAHSQVRHVPLQR
LRIVRGTQLFEDKYALAVLDNR
DPLDNVTTATGRTPEGLRELQLRSLTEILKGGVLIRGNP
QLCYQDTVLWKDVFRKNNQLA
PVDIDTNRSRACPPCAPACKDNHCWGASPEDCQTLTGT
IAPRAVPAARARLPTDCCHEQC
AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDT
FESMPNPEGRYTFGASCVTTCP
YNYLSTEVGSCTLVCPLNNQEVTAEDGTQRCEKCSKSC
ARVCYGLGMEHLRGARAITSAN
IQEFAGCKKIFGSLAFLPESFDGNPSSGIAPLTPEQLQVF
ETLEEITGYLYISAWPDSLH
DLSVFQNLRVIRGRVLHDGAYSLALQGLGIRWLGLRSL
RELGSGLVLIHRNTHLCFVHTV
PWDQLFRNPHQALLHSGNPSEEECGLKDFACYPLCAH
GHCWGPGPTQCVNCSHFLRGQEC
VKECRVWKGLPREYVNGKHCLPCHPECQPQNSTETCT
GSEADQCTACPHYKDSPFCVARC
PSGVKPDLSYMPIWKYPDEEGMCQPCPINCTHSCVDLD
ERGCPAEQRASPATSIIATVVG
ILLFLVIGVVVGILIKRRRQKIRKYTMRRLLQETELVEPL
TPSGAMPNQAQMRILKETEL
RKVKVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLREN
TSPKANKEILDEAYVMAGLGSP
YVSRLLGICLTSTVQLVTQLMPYGCLLDHVREHRGRLG
SQDLLNWCVQIAKGMSYLEDVR
LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEYH
ADGGKVPIKWIALESILRRRFT
HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEK
GERLPQPPICTIDVYMIMVKCWM
IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPSSPL
DSTFYRSLLEDDDMGDLVDA
EEYLVPQQGFFFPDPAPGAGSTAHRRHRSSSTRSGGGE
LTLGMEPSGEEPPRSPLAPSEG
AGSDVFEGELGMGATKGPQSISPRDLSPLQRYSEDPTLP
LPTETDGYVAPLACSPQPEYV
NQPEVRPQPPLTPEGPLPPVRPAGATLERPKTLSPGKNG
VVKDVFTFGGAVENPEYLVPR
GGSASQPHPPALCPAFDNLYYWDQDPSERGSPPNTFEG
TPTAENPEYLGLDVPV

Subject sequence

>1N8Y:C|PDBID|CHAIN|SEQUENCE
TQVCTGTDMKLRLPASPETHLDMLRHLYQGCQVVQG
NLELTYVPANASLSFLQDIQEVQGYMLIAHNQVKRVPL
QRLRIVRGTQLFEDKYALAVLDNRDPQDNVAASTPGR
TPEGLRELQLRSLTEILKGGVLIRGNPQLCYQDMVLWK
DVFRKNNQLAPVDIDTNRSRACPPCAPACKDNHCWGE
SPEDCQILTGTICTSGCARCKGRLPTDCCHEQCAAGCT
GPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMH
NPEGRYTFGASCVTTCPYNYLSTEVGSCTLVCPPNNQE
VTAEDGTQRCEKCSKPCARVCYGLGMEHLRGARAITS
DNVQEFDGCKKIFGSLAFLPESFDGDPSSGIAPLRPEQL
QVFETLEEITGYLYISAWPDSLRDLSVFQNLRIIRGRILH
DGAYSLTLQGLGIHSLGLRSLRELGSGLALIHRNAHLCF
VHTVPWDQLFRNPHQALLHSGNRPEEDCGLEGLVCNS
LCAHGHCWGPGPTQCVNCSHFLRGQECVEECRVWKG
LPREYVSDKRCLPCHPECQPQNSSETCFGSEADQCAAC

```
AHYKDSSSCVARCPSGVKPDLSYMPIWKYPDEEGICQP
CPIN
```

Step2: Alignment between the target sequences and PDB structure template

```
>>PDB:1N8Y_C mol:protein length:608   protooncoprotein
(608 aa)
 initn: 2543 init1: 2543 opt: 3058  Z-score: 5497.8  bits:
1028.4 E():   0
Smith-Waterman score: 3058;  89.967% identity (96.875%
similar) in 608 aa overlap (23-629:1-608)

         10       20       30       40       50       60
Sequen
MELAAWCGWGLLLALLSPGASGTQVCTGTDMKLRLP
ASPETHLDIVRHLYQGCQVVQGNL
                      ::::::::::::::::::::::..::::::::::::::
PDB:1N
TQVCTGTDMKLRLPASPETHLDMLRHLYQGCQVVQG
NL
                           10       20       30

         70       80       90      100      110      120
Sequen
ELTYLPANATLSFLQDIQEVQGYMLIAHSQVRHVPLQR
LRIVRGTQLFEDKYALAVLDNR
    ::::.::::.::::::::::::::::::.::..:::::::::::::::::::::::::::
PDB:1N
ELTYVPANASLSFLQDIQEVQGYMLIAHNQVKRVPLQR
LRIVRGTQLFEDKYALAVLDNR
    40       50       60       70       80       90

        130      140      150      160      170
Sequen                                   DPLDNVTTAT-
GRTPEGLRELQLRSLTEILKGGVLIRGNPQLCYQDTVL
WKDVFRKNNQL
    :: :::...: ::::::::::::::::::::::::::::::::::: :::::::::::::
PDB:1N
DPQDNVAASTPGRTPEGLRELQLRSLTEILKGGVLIRGN
PQLCYQDMVLWKDVFRKNNQL
   100      110      120      130      140      150

   180      190      200      210      220      230
Sequen
APVDIDTNRSRACPPCAPACKDNHCWGASPEDCQTLT
GTIAPRAVPAARARLPTDCCHEQ
    :::::::::::::::::::::::::::.::::::.:::::  . .   ..::::::::::
PDB:1N
APVDIDTNRSRACPPCAPACKDNHCWGESPEDCQILTG
TICTSGCARCKGRLPTDCCHEQ
   160      170      180      190      200      210

   240      250      260      270      280      290
Sequen
CAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTD
TFESMPNPEGRYTFGASCVTTC
    ::::::::::::::::::::::::::::::::::::::::::::: ::::::::::::::
PDB:1N
CAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTD
TFESMHNPEGRYTFGASCVTTC
    220      230      240      250      260      270

   300      310      320      330      340      350
Sequen
PYNYLSTEVGSCTLVCPLNNQEVTAEDGTQRCEKCSKS
CARVCYGLGMEHLRGARAITSA
    ::::::::::::::::: :::::::::::::::::::.:::::::::::::::::::::.
PDB:1N
PYNYLSTEVGSCTLVCPPNNQEVTAEDGTQRCEKCSKP
CARVCYGLGMEHLRGARAITSD
    280      290      300      310      320      330

   360      370      380      390      400      410
Sequen
NIQEFAGCKKIFGSLAFLPESFDGNPSSGIAPLTPEQLQV
FETLEEITGYLYISAWPDSL
    :.:::.::::::::::::::::::.:::::::: ::::::::::::::::::::::::::
PDB:1N
NVQEFDGCKKIFGSLAFLPESFDGDPSSGIAPLRPEQLQ
VFETLEEITGYLYISAWPDSL
    340      350      360      370      380      390

   420      430      440      450      460      470
Sequen
HDLSVFQNLRVIRGRVLHDGAYSLALQGLGIRWLGLRS
LRELGSGLVLIHRNTHLCFVHT
    .:::::::::.::::.::::::::.::::::. :::::::::::::.:::::.:::::::
PDB:1N
RDLSVFQNLRIIRGRILHDGAYSLTLQGLGIHSLGLRSLR
ELGSGLALIHRNAHLCFVHT
    400      410      420      430      440      450

   480      490      500      510      520      530
Sequen
VPWDQLFRNPHQALLHSGNPSEEECGLKDFACYPLCA
HGHCWGPGPTQCVNCSHFLRGQE
    ::::::::::::::::::: .::.::: ...: .::::::::::::::::::::::::::
PDB:1N
VPWDQLFRNPHQALLHSGNRPEEDCGLEGLVCNSLCA
HGHCWGPGPTQCVNCSHFLRGQE
    460      470      480      490      500      510

   540      550      560      570      580      590
```

```
Sequen
CVKECRVWKGLPREYVNGKHCLPCHPECQPQNSTETC
TGSEADQCTACPHYKDSPFCVAR
      :: ::::::::::::..:.:::::::::::::.::: :::::::.::.:::::. ::::
PDB:1N
CVEECRVWKGLPREYVSDKRCLPCHPECQPQNSSETCF
GSEADQCAACAHYKDSSSCVAR
     520       530       540       550       560       570

    600       610       620       630       640       650
Sequen
CPSGVKPDLSYMPIWKYPDEEGMCQPCPINCTHSCVDL
DERGCPAEQRASPATSIIATVV
      ::::::::::::::::::::::.:::::::
PDB:1N CPSGVKPDLSYMPIWKYPDEEGICQPCPIN
     580       590       600
```

2.5.2 Generation of coordinates for SCRs and SVRs:

This model was built in three steps.

1. search model
2. Malign model
3. Get_model

2.5.2.1 SEARCH MODEL:

This step searches for structure for structures that have a match with query sequence [Q8EQB6].The length of the sequence, %identity, and scores were displayed as a result.

RELATED_SEQUENCES

```
  # CODE_1        CODE_2 LEN1 LEN2 NID    %ID    %ID
SCORE  SIGNI SIGNI2 SIGNI3
------------------------------------------------------------------------
-----
  1 Q60553_my  1N8Y     592  608 535  88.0  90.4  501185.
156.1 -999.0 -999.0
```

2.5.2.2 MALIGN:

An alignment was made by considering by complete length of sequence. A pair wise dynamic program alignment “ALIGN” was performed using a local alignment. The parameters were as given below

2.5.2.3 GET MODEL:

About five models were generated by taking into consideration the above parameters. In this step a sequence-structure alignment, number of atoms, topology and restrains were constructed.

III. RESULTS AND DISCUSSION

Swiss-Prot protein sequence database was scanned for receptor tyrosine protein kinases-ERBB2 and the resulted hits were analyzed using BLASTP analysis. The result was given below. From the result, “Q60553”protein is taken into consideration for further analysis. (table 1).

3.1 BLAST Analysis

Blast analysis was carried out using five matrices, PAM 30, 70 and BLOSUM 80, 62 and 45 (table2)

3.2 FASTA Analysis:

FASTA analysis was carried out using five matrices PAM 120,250 BLOSUM 50, 62, 80. (table 3)

3.3 WU-BLAST2 ANALYSIS:

WU-Blast2 analysis was carried out using eight matrices PAM 30, 70, 120 and 250 BLOSUM 45, 50, 62 and 80 (table 4)

3.4 FINAL RESULT: (table 5)

3.5 Homology Modeling

Get-model resulted in five protein structures. They are given in Table 6 and Figure 1.
(table 6)

3.6 Ramachandran plot

Further analysis was supported by Ramachandran plots. The number of residues present in disallowed region is 7 in case of 1N8Y. They are His512, His568, Leu529, Ser201, Lys10, Glu327, Ser572. The number of residues present in disallowed region is 10.They are Leu280, His496, Asn519, Ser565, Asp88, Ser349, Glu311, Lys10, Ser556, Leu101. It was observed that these residues are on the surface of the structure and they may have little or no effect on the protein.

Table-1: Sequence analysis of ERBB2 showing score, E-value, %identities and positives, gaps, and overlaps

Swiss-prot	%identity	%positives	Number of gaps	Score (bits)	E-value	PDB ID	Residue overlap
ERBB2_CANFA (O18735)	91	95	1	1192	0.0	1S78-A	23-645
ERBB2_HUMAN (P04626)	100	100	0	1299	0.0	1S78-A	23-646
ERBB2_MESAU (Q60553)	89	92	1	1145	0.0	1N8Y-C	23-629
ERBB2_MOUSE (P70424)	94	96	0	1212	0.0	1N8Y-C	23-647
ERBB2_RAT (P06494)	99	99	1	1262	0.0	1N8Y-C	23-631

Table-2: Blast Analysis using five different matrices showing score, E-value, % identities, % similarities, gaps and overlaps

Matrix	%Identity	%positives	No: of gaps	Score (bits)	E-value	PDB ID	Residue overlap
PAM30	87	89	1	1693	0.0	1N8Y-C	23-629
PAM70	87	89	1	1486	0.0	1N8Y-C	23-629
BLOSUM80	87	89	1	1355	0.0	1N8Y-C	23-629
BLOSUM62	87	89	1	1135	0.0	1N8Y-C	23-629
BLOSUM45	87	89	1	1012	0.0	1N8Y-C	23-629

From the above analysis, PAM30 was selected as scoring matrix for its high percentage of Identity & positivity, Low E-value, few gaps, high score when compared to other matrices.

Table-3: FASTA analysis using five matrices showing % identities, % similarities, E-value, score, Gaps & Overlaps

Matrix	%Identity	%similarity	No.of gaps	Score bit	E-value	PDB ID	Residue overlap
PAM120	89.967	96.875	1	1046.5	0	1N8Y-C	23-629
PAM250	89.967	97.862	1	593.9	1.2E-168	1N8Y-C	23-629
BLOSUM50	89.967	95.230	1	810.9	0	1N8Y-C	23-629
BLOSUM62	89.967	95.559	1	1008.1	0	1N8Y-C	23-629
BLOSUM80	89.967	95.230	1	1130.7	0	1N8Y-C	23-629

From the above analysis, PAM120 was selected as scoring matrix for its high percentage of Identity & similarity Low E-value, less no. of gaps, high score when compared to other matrices.

Table-4: Wu-blast analysis was carried out using eight different matrices, score, E-value, % identities, % similarities, gaps & overlaps

Martix	**%identity**	**%positives**	**No: of gaps**	**Score (bits)**	**E-value**	**PDB ID**	**Residue overlap**
PAM30	94	96	0	1560	0	IN8Y-C	23-450
PAM70	90	93	1	2537	0	IN8Y-C	23-450
PAM120	87	91	1	2929	0	IN8Y-C	23-629
PAM250	87	92	1	2916	1.9C-287	IN8Y-C	23-629
BLOSUM45	87	90	1	3590	0	IN8Y-C	23-629
BLOSUM50	87	90	1	3828	0	IN8Y-C	23-629
BLOSUM62	87	89	1	2941	0	IN8Y-C	23-629
BLOSUM80	87	89	1	4692	0	IN8Y-C	23-629

From the above analysis, BLOSUM80 was selected as scoring matrix for its high percentage of Identity & positivity, Low E-value, less no. of gaps, high score when compared to other matrices.

Table-5: Blast, Fasta, WU Blast2, Matrices were used displaying Score, E-value, % identity, % positive, and gaps.

Methods	**Matrix**	**%Identity**	**%Positives**	**No: of gaps**	**Score (bits)**	**E-value**	**PDB-ID**	**Residue overlap**
BLAST	PAM30	87	89	1	1693	0	1N8Y-C	23-629
FASTA	PAM120	89.967	96.875	1	1046.5	0	1N8Y-C	23-629
WU-BLAST	BLOSUM80	87	89	1	4692	0	1N8Y-C	23-629

Here score and E_value are not considered as it represents the values for particular method so the rest parameters are considered. As residue overlap and number of gaps are similar for 3 sequences they are not considered. Only maximum %identity and % positives are taken into consideration, hence FASTA method was selected to perform homology modeling.

Table 6: Summary of successfully produced models.

Filename	Energy (kcal/mol)
Q60553.B99990001.pdb	4238.07764
Q60553.B99990002.pdb	4155.61768
Q60553.B99990003.pdb	4142.48633
Q60553.B99990004.pdb	4274.60498
Q60553.B99990005.pdb	4317.16064

Table 7: Summary of optimized models

THEORETICAL MODELS	**ENERGY(kcal/mol)**	**RMSD(Aº)**
Q60553.B99990001.pdb	2397.3853	1.5993
Q60553.B99990002.pdb	2390.4692	2.6017
Q60553.B99990003.pdb	2445.6414	0.6495
Q60553.B99990004.pdb	2398.9771	1.0013
Q60553.B99990005.pdb	**2392.5422**	**0.5022**

Fifth model with lowest energy and rmsd value less than 2.0A° was considered as best optimized and superimposed model among the five generated models. Though the second generated model has lowest energy it is not considered as it has highest RMSD value.

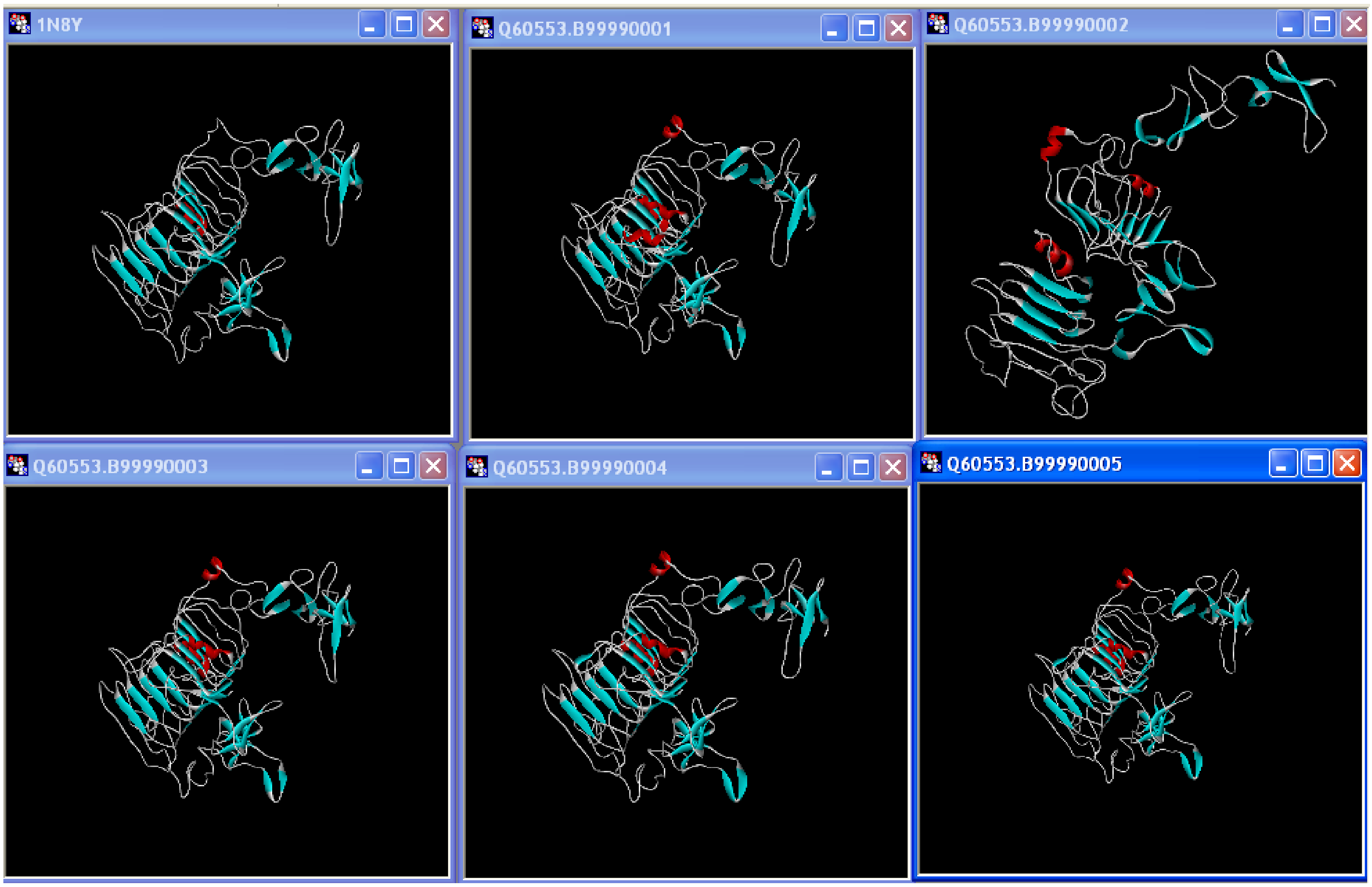

Figure 1: Five models generated from Modeller 8v2 software, 1N8Y is given for comparison.

Energy optimization for all models resulted in model-2 with lowest energy state. However, RMS superposition with template structure 1N8Y resulted in model-5 as best homologous model (Table 7)

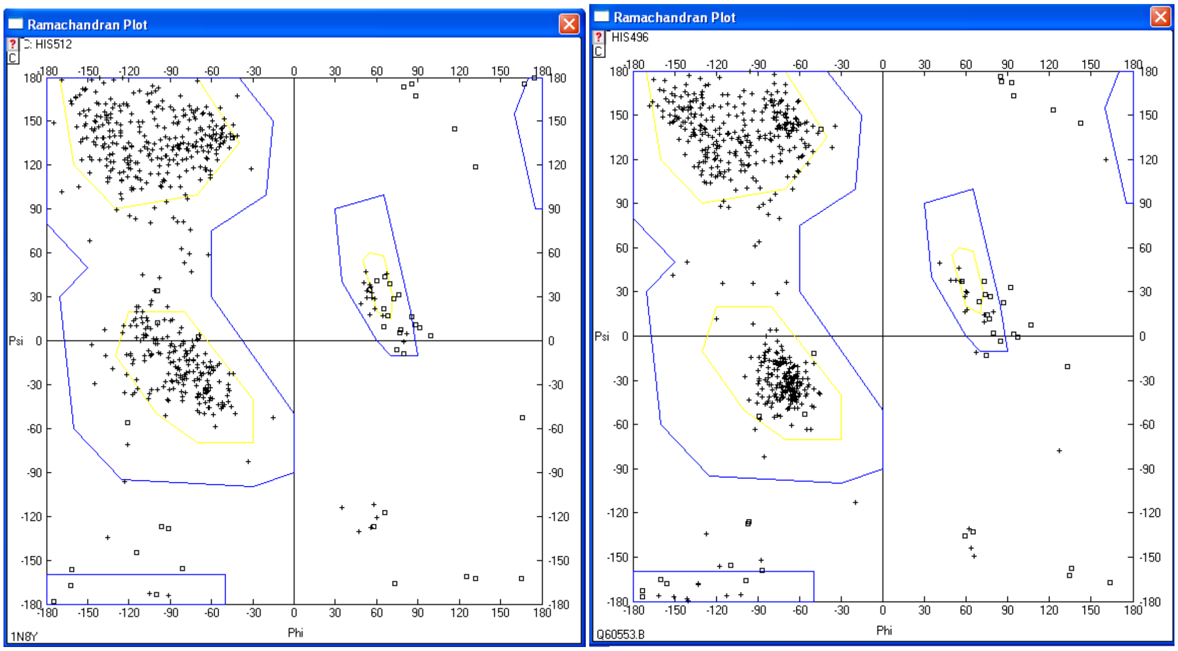

Figure-2 : Ramachandran Plots of 1N8Y and model-5

IV. CONCLUSION

The physico-chemical properties of amino acids are necessary to maintain the structure and function of proteins. The residues that are likely to be conserved can be detected by using scoring matrices. In order to use any sequence alignment tool with different scoring matrices it is necessary to quantify the scoring matrices. Therefore an attempt was made to study the influence of scoring matrix on alignment as well as to perform a homology derived model for the protein of interest. Hence, three different programs such as BLAST, FASTA and WU BLAST2 are employed in this work. When compared with other two programs, FASTA (with PAM 120 matrix) revealed a clear evolutionary relationship. A scan against Swiss-Prot protein sequence database for receptor protein tyrosine kinase erbb2 resulted in 5 hits, of which, Q60553 was selected for Sequence analysis and Homology modeling. The analysis identified a relevant homology with PDB protein 1N8Y (89.96% identity and 96.87% similarity). A list of similar PDB hits from FASTA program resulted in moderate to good alignments. Homology Modeling was initiated with Modeler 8v2 run in Windows operating system. Out of five generated models, fifth model was considered as the best optimized and superimposed model. It exhibited the lowest energy (2392.54 kcal/mol) and low RMSD value (0.5A^{o}) than the remaining models. By using Ramchandran plot, the probable number of residues that appear in different regions of the plot are identified. Therefore this study suggests the fact that a fast and reliable homology model was possible by considering the sequences with profound similarity at sequence level as the method employed is customizable and result-oriented.

V. REFERENCES

[1]. Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, Leahy DJ. Structure of the extra cellular region of HER2 alone and in complex with the Herceptin Fab Nature. 2003 Feb 13;421(6924):756-60.

[2]. Brys M, Semczuk A, Rechberger T, Krajewska WM.Expression of erbB-1 and erbB-2 genes in normal and pathological human endometrium. Oncol Rep. 2007 Jul;18(1):261-5.

[3]. Allen SD, Garrett JT, Rawale SV, Jones AL, Phillips G, Forni G, Morris JC, Oshima RG, Kaumaya PT. Program.Peptide vaccines of the her-2/neu dimerization loop are effective in inhibiting mammary tumor growth in vivo. J Immunol. 2007 Jul 1;179(1):472-82.

[4]. (www.expasy.org)

[5]. (www.ncbi.nlm.nih.gov/blast)

[6]. (www.rcsb.org/pdb)

[7]. (https://salilab.org/modeller)