PREDICTION OF STRUCTURES OF BIOACTIVE PEPTIDES BY USING HOMOLOGY APPROACH

*Shrikant Sharma1, Raghvendar Singh2, Shashank Rana1

1. School of Life Science Singhania University, Jhunjhunu (RJ), India
2. Bioinformatics Facility Department of Immunology College of Biotechnology Sardar Vallabh Bhai Patel University of Agriculture & Technology. Meerut (U.P.) India

E-mail of Corresponding author: shribioinfo@gmail.com

Abstract

In silico study for prediction of structures of bioactive peptides by using homology approach and different bioinformatics tools as Basic Local Alignment Search Tool (BLAST) Swiss Model workspace repository and template were applied. The homology model comprises four main steps find out the homology model against query protein identification of structural template(s) alignment of target sequence and template structure(s) model building and model quality evaluation. The feasible structures of antimicrobial antithrombotic casein derived immunomodulatory and mineral binding peptides were designed. Antimicrobial peptide seq8 (NP_446175.1) was analyzed by using above stated methodology and found that 3E6U chain D showed homology with this sequence. Like this methodology was followed for all query bioactive peptide sequences.

Keywords: Swiss Model; Homology; Template; Peptides

1. Introduction

Bioactive peptides have been defined as specific protein fragments that have a positive impact on body functions and conditions and may ultimately influence health1. According to Fitz-Gerald & Murray (2006)2 bioactive peptides have been defined as 'peptides with hormone- or drug-like activity that eventually modulate physiological function through binding interactions to specific receptors on target cells leading to induction of physiological responses. Most of these bioactivities are encrypted within the primary sequence of the native protein and peptides require to be released through one of the following ways: Hydrolysis by digestive enzymes such as trypsin and pepsin3,4 Food processing 5 and through hydrolysis by proteolytic microorganisms or through the action of proteolytic enzymes derived from the microorganisms2.

2. Material and methods

The comparison of nucleotide or protein sequences from the same or different organisms is a very powerful tool in molecular biology. By finding similarities between sequences scientists can infer the function of newly sequenced genes predict new members of gene families and explore evolutionary relationships now the whole genomes are being sequenced sequence similarity searching can be used to predict the location and function of protein-coding and transcription-regulation regions in genomic DNA. Basic Local Alignment Search Tool (BLAST) is the tool most frequently used for calculating sequence similarity. BLAST comes in variations for use with different query sequences against different databases6. The way most people use BLAST is to input a nucleotide or protein sequence as a query against all (or a subset of) the public sequence databases pasting the sequence into the textbox on one of the BLAST web pages. This sends the query over the Internet the search is performed on the NCBI servers and the results are posted back to the person’s browser in the chosen display format. However many biotech companies genome scientists and bioinformatics personnel
may want to use “stand-alone” Blast to query their own local databases or want to customize BLAST in some way to make it better suit their needs. Stand-alone BLAST comes in two forms the executables that can be run from the command line or the Standalone WWW BLAST Server which allows users to set up their own in-house versions of the BLAST Web pages.

2.1. **SWISS-MODEL workspace:** The SWISS-MODEL Workspace is a web-based integrated service dedicated to protein structure homology modeling. It assists and guides the user in building protein homology models at different levels of complexity. Building a homology model comprises four main steps: identification of structural template(s) alignment of target sequence and template structure(s) model building and model quality evaluation. These steps can be repeated until a satisfying modeling result is achieved. Each of the four steps requires specialized software and access to up-to-date protein sequence and structure databases. Protein sequence and structure databases necessary for modeling are accessible from the workspace and are updated in regular intervals. Software tools for template selection model building and structure quality evaluation can be invoked from within the workspace. A personal working environment (workspace) where several modeling projects can be carried out in parallel is provided for each user. This help file provides references and illustrate the use of the individuals tools available from within the SWISS-MODEL Workspace.

2.2. **SWISS-MODEL Repository:** The SWISS-MODEL Repository is a database of annotated three-dimensional comparative protein structure models generated by the fully automated homology-modelling pipeline SWISS-MODEL. The repository is developed at the Biozentrum Basel within the Swiss Institute of Bioinformatics. The repository currently contains three-dimensional models for sequences from the UniProt knowledge base. The content of the repository is updated on a regular basis incorporating new sequences taking advantage of new template structures becoming available and reflecting improvements in the underlying modelling algorithms. The current data status is given on the entry page. The steps of structure prediction of protein are specified as firstly we opt the protein sequence of bioactive peptide from our database i.e. IBPD. Then investigate the protein data bank (PDB) by using BLASTp. After BLASTp and selection of best match form BLASTp results we had done the homology modelling by swiss modal workplace swiss modal repository. Finally we unearth the achievable three dimensional protein models.

3. Results and discussion

The feasible structures of proteins of unknown structures of bioactive peptides were designed by via bioinformatics tools and software packages and online servers available through different websites. To know about the functions and mode of action it is necessary to study the structures of the peptides. The application of computational chemistry will result in the creation of structure and sequence databases that will enable bioactive fragments to be searched for in the protein chain. Some bioactive peptides have demonstrated multifunctional activities based on their structure and other factors including hydrophobicity charge or microelement binding properties. A peptide–peptoid hybrid (peptomer) library was designed and synthesized by Ovadia et al (2010) based on the sequence Phe-D-Phe-Arg-Trp-GlyThis sequence was previously found to specifically activate the melanocortin-4 receptor (MC4R) which participates in regulation of energy homeostasis and appetite. The library of peptomers included a peptoid bond in the Phe and/or D-Phe position and consisted of linear and backbone cyclic analogs.
differed in their ring size. Jones et al. (2010)\cite{18} proposed an innovative strategy for the development of bioactive cell-penetrating peptides. They combine computer-based design of peptides with specific targeting to elaborate a potent cell-penetrating bioactive peptide derived from cytochrome C. In light of this bioactive peptides in five categories based on function as immunomodulatory mineral binding opioid antimicrobial and antithrombotic peptides accordingly the structures of above stated peptides were predicted as given in table 1.1. Ten sequences per peptide were chosen form the above stated peptides. On the bases of study by using different software packages and online servers the results are explained as given in figure 1.1. The peptide sequences as seq1 seq2 etc. and so on. An antimicrobial seq8 (NP_446175.1) was analyzed by swiss model for structure prediction and found that this seq was found best similarity with 3E6uD (template and alignment shown in figure 1.4) as the e-value was 0.00e-1 bit score 91.479. It is again noted that when above stated seq was further analyzed by swiss repository template identification tool same structure was found similarity with seq8 of antimicrobial peptide. Evaluation of model quality is a crucial step in homology modeling. While the performance of the automated SWISS-MODEL\cite{19} pipeline in general is continuously evaluated by the EVA project\cite{20} the quality of individual models can vary significantly. Therefore graphical plots of Anolea mean force potential\cite{21} GROMOS empirical force field energy and Verify3D profile evaluation are provided to enable the user to estimate the quality of protein models and template structures. Anolea\cite{21} an atomic empirical mean force potential was used to assess packing quality of the models. The program performs energy calculations on a protein chain evaluating the "Non- Local Environment" (NLE) of each heavy atom in the molecule. The y-axis of the plot represented the energy for each amino acid of the protein chain. Negative energy values (in green) represent favorable energy environment whereas positive values (in red) unfavorable energy environment for a given amino acid. Verify3D method assesses protein structures using three-dimensional profiles. This program analyzed the compatibility of an atomic model (3D) with its own amino acid sequence (1D). Each residue assigned a structural class based on its location and environment (alpha beta loop polar a polar etc). Then a database generated from good structures used to obtain a score for each of the 20 amino acids in this structural class. The vertical axis in the plot represents the average 3D-1D profile score for each residue in a 21-residue sliding window. The ranges of scores varies from -1 (bad score) to +1 (good score) as represent in graph figure 1.3 as blue line. In Gromos the y-axis of the plot represents empirical force field energy for each amino acid of the protein chain. Negative energy values (in green) represent favorable energy environment whereas positive values (in red) unfavorable energy environment for a given amino acid. Like this all other structures were designed (Table1.1). The modelled structures of Antithrombotic peptides seq no. 1 of showed the homology with the Chain D of 3BRW. It is based on the NCBI's BLASTp against protein databank swiss modal repository and domain identification tools of swiss modal. 3BRW is a structure of Rap-Rap GAP complex from Rap1 GTPase-activating protein of Homo sapiens. On the other hand sequence no. 5 showed the similarity with the Chain B of 1KB2 a crystal structure of VDR domain a synthetic construct of Homo sapiens. Sequence no. 11 found homology with the Chain B of IYPQ a structure of LOX1 dioxane complex of immune system of Homo sapiens. Sequence no. 6 has similarity with the Chain A of 2GI7 which is the structure of human platelet...
glycoprotein. In antimicrobial peptides Sequence no. 6 found homology with the Chain A of 1178 belongs to *E. coli*'s OMPT. Sequence no. 10 showed similarity with the Chain B of 2PCJ a structure of ABC transporter of *Aquifex aeolicicus*. The sequences no. 2 and no. 8 had the resemblance with the Chain D of 3E6U which is a crystal structure of HSN1 NS1 a polymer of influenza virus. According to the studies in immunomodulatory peptides; Sequence no. 1 similitude with the Chain A of 1BG1 a crystal structure of transcription factor stat3b/DNA complex. Sequence no. 5 matched with the Chain A of 230S a structure of piscidin. This 2JOS structure is important as it belongs to antimicrobial protein found in DPC Micelles. Sequence no. 2 had equivalence with the Chain A of 2KA2 as NMR structure of transmembrane peptide of *Homo sapiens*. Sequence no. 3 found homologous with the Chain A of 2OFU a crystal structure of 2-aminopyrimidine carbamate belongs to transferase family of *Homo sapiens*. Sequence no. 4 showed identicalness with the Chain B of 3E7M a structure of murine belongs to an oxidoreductase family of *Mus Musculus*. Casein based peptides; Sequence no. 8 matched with the Chain A of 1CKJ a structure of casein kinaseI delta of phosphotransferase family of *Rattus norvegicus*. Sequence no. 1 found similarity with the Chain F of 3EZQ a crystal structure of Fas/FADD death domain complex of *Homo sapiens*. Mineral binding peptides; Sequence no. 7 showed resemblance with Chain A of 1H9D a structure of AML1/CFB-BETA/DNA complex of *Homo sapiens*. On the other hand sequence no. 5 showed similarity with the Chain C of 1565 a structure of cytockine of *Mus Musculus*. Sequence no. 8 found the homology with the Chain A of 1WW1 a crystal structure of tRNAase Z belongs to hydrolase family of *Thermotoga maritime msb8*. Sequence no. 6 showed identicalness with the Chain A of signalling protein /metal binding protein of *Homo sapiens*.

**Figure 1.1: Blastp of antimicrobial seq 8 against protein databank**

| Score | E value | Method | Composition Matrix | Adjusted Sensitivity | Positives | Queries | Targets |
|-------|---------|--------|--------------------|---------------------|-----------|---------|---------|
| 7746  | 0.001   | 0.001  | 0.001              | 0.001               | 0.001     | 0.001   | 0.001   |

**Object 1:**

| Score | E value | Method | Composition Matrix | Adjusted Sensitivity | Positives | Queries | Targets |
|-------|---------|--------|--------------------|---------------------|-----------|---------|---------|
| 7746  | 0.001   | 0.001  | 0.001              | 0.001               | 0.001     | 0.001   | 0.001   |

**Object 2:**

| Score | E value | Method | Composition Matrix | Adjusted Sensitivity | Positives | Queries | Targets |
|-------|---------|--------|--------------------|---------------------|-----------|---------|---------|
| 7746  | 0.001   | 0.001  | 0.001              | 0.001               | 0.001     | 0.001   | 0.001   |
Figure 1.2: Alignment of target sequence with 3e6u

Figure 1.3: Graphical representation of homology modelling of 3e6u antimicrobial peptides precursor seq no 8.
Figure 1.4: Template and alignment antimicrobial seq 8 with 3e6u

Score = 773 bits (1995) Expect = 0.0   Method: Composition-based stats.
Identities = 365/399 (91%) Positives = 381/399 (95%)
Query: 1
MAQRAFPNPYADYNKSLAE NYFDSTGRLTPEFSHRLTNKIRELLQQMERGLKSA
DPQDT 60
MAQRAFPNPYADYNKSLAE YFD+ GRLTPEFS RLTKIRELLQQMERGLKSADP
+ DGT
Sbjct: 11
MAQRAFPNPYADYNKSLAE GYFDAAGRTLPEFSQRLTNKIRELLQQMERGLKS
ADPRDGT 70
Query: 61
GYTGWAGIAVLYLHLHNVFDPAYLQMAHSYVKHSLNCLSRRSITFLCGDAGP
LAVAAL 120
GYTGWAGIAVLHLYL++VFGDPAYLQ+AH YVK SLNCL + + RSITFLCGDAGPLAVAALV
Sbjct:71

GYTGWAGIAVLHLYLDVFDPAYLQLAHGYVKQSLNCLTKRSITFLCGDAGPLLAVAALV 130
Query:121

YHKMNNSGKQAEDCITRLIHLNKiDPHVPNEMLYGRIGYI+ALLFVKNFGEEKIPQSHIQ 180
YHKMNN+ KQAEDCITRLIHLNKiDPH PNEMLYGRIGYI+ALLFVKNFG EKIPQSHIQ
Sbjct:131

YHKMNEKQAEDCITRLIHLNKiDPHAPNEMLYGRIGYIYALLFVKNFGVEKI PQSHIQ 190
Query:181

QICETILTSGKLSRKKiNTKSPLMYEWWQYETYVGAHGLAGIYYYLMQPSLHVSGKGL 240
QICETILTSGE L+RKRNF KTSPMLMYEWWQETYVGAHGLAGIYYYLMQPSL
VSQGKGL
Sbjct:191

QICETILTSGENLARKRFNTAKSPMLMYEWWQETYVGAHGLAGIYYYLMQPSL
VSQGKGL 250
Query:241

HSLVKPSVDVFCSKLFPSGNYPSCLDDTRDLHVVCHAPGVYMLIQAYKVFKEEYLC 300
HSLVKPSVD+VCQLKFPSPGNYP C+ D RDLLHVCHAPGVYMLIQAYKVF+EEYLC
Sbjct:251

HSLVKPSVDDYCQLKFPSPGNYPCCIGDNRDLLLLVVCHAPGVYMLIQAYKVF
REEKYLCL 310
Query:301

DAQQCADVIWQYGLLKKGYGLCHGAAGNAYAFLALYLNLTQDMAKLYRACKF
AEWCLDYGE 360
DA QCADVIWQYGLLKKGYGLCHGAAGNAYAFLALYLNLTQDMAKLYRACKF
KLYRACKFAEWCL + YGE
Sbjct:311

DAYQCADVIWQYGLLKKGYGLCHGSAGNAYAFLTLNYLNLTQDMKLYRACKF
EWCLEYGE 370
Query:361

HGCRTPDTPFSLFEGMAGTIYFLADLLVPTKAKFPAFEL 399
HGCRTPDTPFSLFEGMAGTIYFLADLLVPTKA+FPALF
Sbjct: 371

HGCRTPDTPFSLFEGMAGTIYFLADLLVPTKARFPAFEL 409
Table 1.1: Details of structures of different peptide precursors based on homology modeling

| Sr. No. | Str. Name | E value | Score | Repository | Templates | Description |
|---------|-----------|---------|-------|------------|-----------|-------------|
| 1       | 1bg1A immuno1 | 0.00e-1 | 96.2  | 1bg1A immuno1 | >IMMUNO1 | gi|21618338||NP_003141.2| |
| 2       | 1ckjA cas8 | 7.43e-119 | 65.529 | 1ckjA cas8 | >CAS8 | gi|194680306||XP_001788962.1| |
| 3       | 1d2tA bioact1 | 4.40e-6 | 15.894 | NA | NA | >BIOACT10 |
| 4       | 1d2tA bioact9 | 7.43e-119 | 65.529 | 1d2tA bioact9 | >BIOACT9 | gi|110756225||XP_001122506.1| |
| 5       | 1d2tA cas3 | 9.80e-5 | 18.182 | NA | NA | >CAS3 |
| 6       | 1d2tA cas2 | 9.40e-7 | 22.33 | NA | NA | >CAS2 |
| 7       | 1dpea antimicro9 | 0.00e-1 | 38.856 | NA | NA | >ANTIMIOCRO9 |
| 8       | 1h9dA mb7 | 0.00e-1 | 100.00 | 1h9dA mb7 | >MB7 | gi|225690634||NP_001139392.1| |
| 9       | 1h30A antithrom9 | 0.00e-1 | 36.919 | NA | NA | >ANTITHROM9 |
| 10      | 1l78A antimicro6 | 0.00e-1 | 100.00 | 1l78A antimicro6 | >ANTIMIOCRO6 | gi|29140984||NP_804326.1| |
| 11      | 1l78A antimicro7 | 0.00e-1 | 47.458 | NA | NA | >ANTIMIOCRO7 |
| 12      | 1kb2B antithrom5 | 3.46e-49 | 100.00 | 1kb2B antithrom5 | >ANTITHROM5 | gi|31543944||NP_003350.2| |
| 13      | 1j0sA immuno5 | 3.20e-49 | 68.871 | 1j0sA immuno5 | >IMMUNO5 | gi|9506805||NP_062038.1| |
| 14      | 1ka6A immuno2 | 1.13e-56 | 100.00 | 1ka6A immuno2 | >IMMUNO2 | gi|169234945||NP_001108409.1| |
| 15      | 1katV biomilk7 | 5.63e-56 | 100.00 | 1katV biomilk7 | >BIOMILK7 | gi|76781480||NP_001020537.2| |
| 16      | 1kn6A op7 | 3.80e-12 | 31.507 | NA | NA | >OP7 |
| 17      | 1rk6A bioact5 | 0.00e-1 | 50.00 | 1rk6A bioact5 | >BIOACT5 | gi|194521198||YP_002056734.1| |
| 18      | 1rk6A bioact2 | 0.00e-1 | 50.00 | 1rk6A bioact2 | >BIOACT2 | gi|116693124||YP_838657.1| |
| 19      | 1rk6A bioact4 | 0.00e-1 | 50.00 | 1rk6A bioact4 | >BIOACT4 | gi|194526538||YP_002062063.1| |
| 20      | 1rk6A bioact6 | 0.00e-1 | 50.00 | 1rk6A bioact6 | >BIOACT6 | gi|194511112||YP_00203641.1| |
| 21      | 1rk6A bioact8 | 0.00e-1 | 50.00 | 1rk6A bioact8 | >BIOACT8 | gi|194504488||YP_002034060.1| |
| 22      | 1rk6A cas7 | 0.00e-1 | 32.707 | 1rk6A cas7 | >CAS7 | gi|21223359||NP_629138.1| |
| 23      | 1s55C mb5 | 2.98e-89 | 97.436 | 1s55C mb5 | >MB5 | gi|16924012||NP_0476490.1| |
| 24      | 1wwlA mb8 | 9.81e-167 | 100.00 | 1wwlA mb8 | >MB8 | gi|6753332||NP_039371.1| |
| 25      | 1ye01 bioact3 | 4.00e-33 | 25.000 | NA | NA | >BIOACT3 |
| 26      | 1ypqB antithrom11 | 7.84e-78 | 100.00 | 1ypqB antithrom11 | >ANTITHROM11 | gi|4505501||NP_002534.1| |
| 27      | 2bz6I antithrom2 | 0.00e-1 | 27.414 | NA | NA | >ANTITHROM2 |

IJBAR (2011) 02(12)  www.ssjournals.com
|   |   |   |   |   |   |
|---|---|---|---|---|---|
| 28 | 2dspB biomilk 6 | 2.80e-24 | 48.864 | 2dspB biomilk 6 | >BIOMILK6 gi|4504645|NP_000587.1|
| 29 | 2ebTA biocat7 | 2.00e-22 | 57.143 | 2ebTA biocat7 | >BIOMILK7 gi|7678148|NP_001020537.2|
| 30 | 2ebTA mb6 | 2.00e-22 | 57.143 | 2ebTA mb6 | >MB6 gi|229202136|NP_690599.1|
| 31 | 2f83A antithrom3 | 0.00e-1 | 24.414 |   | >ANTITHROM3 gi|4504645|NP_000587.1|
| 32 | 2gi7A antithrom6 | 5.38e-71 | 78.022 | 2gi7A antithrom6 | >ANTITHROM6 gi|635650233|NP_145298.4|
| 33 | 2I1bA biomilk5 | 5.37e-81 | 100.00 | 2I1bA biomilk5 | >BIOMILK5 gi|10835145|NP_000567.1|
| 34 | 2IlkA biomilk4 | 4.51e-73 | 100.00 | 2IlkA biomilk4 | >BIOMILK4 gi|10835141|NP_000563.1|
| 35 | 2IpbD cas4 | 6.30e-5 | 20.00 |   | >CAS4 gi|226530019|NP_001146142.1|
| 36 | 2ofuA immuno3 | 2.13e-164 | 100.00 | 2ofuA immuno3 | >IMMUNO3 gi|112789548|NP_001036236.1|
| 37 | 3brwD antithrom1 | 1.19e-88 | 100.00 | 3brwD antithrom1 | >ANTITHROM1 gi|38857575|NP_007777.1|
| 38 | 2pcjB antimicro10 | 0.00e-1 | 44.545 | 2pcjB antimicro10 | >ANTIMIOCRO10 gi|150025006|
| 39 | 2rh1A op9 | 0.00e-1 | 25.893 | NA | >OP9 gi|148747212|NP_034207.2|
| 40 | 2rh1A op2 | 1.20e-40 | 100.00 | NA | >OP2 gi|4504648|NP_001548.1|
| 41 | 2rh1A op3 | 8.50e-41 | 15.402 | NA | >OP3 gi|4504648|NP_000625.1|
| 42 | 2rh1A op5 | 0.00e-1 | 16.29 | NA | >OP5 gi|843702363|NP_001033686.1|
| 43 | 2rh1A op10 | 0.00e-1 | 25.721 | NA | >OP10 gi|33859754|NP_034206.1|
| 44 | 2ziyA antimicro7 | 4.40e-19 | 13.782 | NA | >ANTIMIOCRO7 gi|31790978|NP_857613.1|
| 45 | 3brwD antithrom1 | 1.19e-88 | 100.00 | 3brwD antithrom1 | >ANTITHROM1 gi|38857575|NP_007777.1|
| 46 | 3e6uD antimicro2 | 0.00e-1 | 100.00 | 3e6uD antimicro2 | >ANTIMIOCRO2 gi|212274337|NP_001130046.1|
| 47 | 3e6uD antimicro8 | 0.00e-1 | 91.479 | 3e6uD antimicro8 | >ANTIMIOCRO8 gi|39930499|NP_446175.1|
| 48 | 3e07A antimicro1 | 1.40e-16 | 21.429 | NA | >ANTIMIOCRO1 gi|167736344|NP_001108066.1|
| 49 | 3e7mB immuno4 | 0.00e-1 | 100.00 | 3e7mB immuno4 | >IMMUNO4 gi|6754872|NP_035057.1|
| 50 | 3ezqF cas10 | 2.63e-48 | 100.00 | 3ezqF cas10 | >CAS10 gi|4505229|NP_003815.1|
| 51 | 3f1sB antithrom8 | 0.00e-1 | 24.281 | NA | >ANTITHROM8 gi|119887347|NP_585990.3|
| 52 | 3gflA biomilk2 | 1.57e-34 | 100.00 | 3gflA biomilk2 | >BIOMILK2 gi|11024682|NP_000651.3|
| 53 | 3kfdD biomilk3 | 1.51e-65 | 100.00 | 3kfdD biomilk3 | >BIOMILK3 gi|63025222|NP_000651.3|
Table 1.2: Accession no of peptide sequences

| Immunomodulatory | Antithrombotic | Opioid | Casein |
|------------------|---------------|--------|--------|
| Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. |
| 1 | NP_004336.2 | 1 | AAB36489.1 | 1 | P84814.1 | 1 | NP_001166365.1 |
| 2 | AAB35532.1 | 2 | P04070.1 | 2 | P01214.1 | 2 | AAB27B35385.1 |
| 3 | Q90WP7.1 | 3 | AAB31168.1 | 3 | P47969 | 3 | AAB30253.1 |
| 4 | P84815.1 | 4 | AAB31167.1 | 4 | Q93456.1 | 4 | AAB30250.1 |
| 5 | NP_001020001.1 | 5 | AAB31166.1 | 5 | O93227.1 | 5 | AAB24259.1 |
| 6 | NP_694542.1 | 6 | P01019.1 | 6 | 0610176A | 6 | AAB24258.1 |
| 7 | NP_077345.1 | 7 | NP_000307.1 | 7 | 0601259A | 7 | AAB24257.1 |
| 8 | ABO64518.1 | 8 | P01015.1 | 8 | Q13519.1 | 8 | AAB30252.1 |
| 9 | NP_001027989.1 | 9 | P11859.1 | 9 | P01210.1 | 9 | AAB23721.1 |
| 10 | Q5ZPR3.1 | 10 | Q9GLP6.1 | 10 | P22005.2 | 10 | AAB23720.1 |
| 11 | AD90622.1 | 11 | Q9TSZ0.1 | 11 | O62647.1 | 11 | AAB23719.1 |
| 12 | Q9GLN8.1 | 12 | Q62923.1 | 12 | AAB23718.1 |
| 13 | P20757.2 | 13 | Q64387.1 | 13 | AAB20364.1 |
| 14 | P01016.1 | 14 | Q28409.1 | 14 | P02662.2 |
| 15 | P01017.1 | 15 | P01212.2 | 15 | NP_620691.2 |
| Antimicrobial | | | | | | | |
| Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. |
| 1 | NP_005756.2 | 1 | NP_776863.1 | 1 | P84523.1 | 1 | NP_620229.2 |
| 2 | P_001130046.1 | 18 | O97373.1 | 18 | O35417.1 | 18 | AAA30479.1 |
| 3 | NP_752583.1 | 19 | P06300.2 | 19 | CAA25452.1 |
| 4 | NP_057396.1 | 20 | Q60478.1 | 20 | AAG60201.1 |
| 5 | NP_758367.1 | 21 | Q95104.1 | 21 | AAL34126.1 |
| Mineral binding | | | | | | | |
| Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. | Seq no. | Accession no. |
| 6 | NP_804326.1 | 22 | P01213.1 | 22 | AAL31044.1 |
| 7 | NP_857613.1 | 1 | P01009.3 | 23 | P05422.1 | 23 | AAK63938.1 |
| 8 | NP_446175.1 | 2 | P01137.2 | 24 | P05421.1 | 24 | AAB23435.1 |
| 9 | YP_002347322.1 | 3 | NP_006604.1 | 25 | P21850.1 | 25 | AAB26704.1 |
| 10 | YP_001295832.1 | 4 | NP_001007730.1 | 26 | P86285.1 | 26 | AAB34798.1 |
| 11 | NP_990323.1 | 5 | P41180.2 | 27 | AAB82584.1 | 27 | AAB34797.1 |
| 12 | NP_001158088.1 | 28 | AAB22676.1 | 28 | AAB26270.1 |
| 13 | NP_003609.2 | 29 | AAB22675.1 | 29 | NP_776953.1 |
| 14 | NP_003609.2 | 30 | AAA13386.1 | 30 | P47710.1 |
| 15 | NP_006604.1 | 31 | NP_061351.2 | 31 | P02663.2 |
| 16 | Q8N4E7.1 | 32 | CAA06546.1 | 32 | P04653.3 |
| 17 | CAA06547.1 | 33 | CAA06547.1 | 33 | AAB37142.1 |
| 18 | 0806130A | 34 | 0806130A | 34 | AAB30428.1 |
| 19 | 0806130A | 35 | CAA31448.1 |
| 20 | 0806130A | 36 | CAA27261.1 |
| 21 | 0806130A | 37 | CAA27261.1 |

References:
1. Kitts D.D. and. Weiler K.A, “Bioactive proteins and peptides from food sources Applications of bioprocesses used in isolation and recovery ”. Current Pharmaceutical Design 2003; 9 :1309–1323.
2. FitzGerald R, and Murray B.A., “Bioactive peptides and lactic fermentations”. International Journal of Dairy Technology 2006; 59:118–125.

3. FitzGerald R. J, Murray B. A, and Walsh D. J, (). “Hypotensive peptides from milk proteins”. Journal of Nutrition, 2004;134:980S–988S.

4. Korhonen H and Pihlanto-Leppala A, “Food-derived bioactive peptides: Opportunities for designing future foods”. Current Pharmaceutical Design 2003; 9:1297–1308.

5. Van-Gunsteren W. F. et al. “Biomolecular Simulations: the GROMOS96 Manual and User Guide.” VdF Hochschulverlag ETHZ Zurich 1996.

6. Altschul S F, Gish W, Miller W, Myers E W and Lipman D J. "Basic local alignment search tool". J Mol Biol 1990; 215 (3): 403–410.

7. http://blast.ncbi.nlm.nih.gov/

8. Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. “SWISS-MODEL: an automated protein homology-modeling server”. Nucleic Acids Res 2003; 31:3381-3385.

9. Arnold K, Bordoli L, Kopp J, and Schwede T. “The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling.” Bioinformatics2006; 22:195-201.

10. http://swissmodel.expasy.org/workspac e/

11. Schwede, T., J. Kopp, et al. "SWISS-MODEL: An automated protein homology-modeling server". Nucleic Acids Res. 2003; 31(13): 3381-3385.

12. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T. “The SWISS-MODEL Repository and associated resources”. Nucleic Acids Res2009; 37: D387-D392.

13. http://swissmodel.expasy.org/repositor y/

14. Dziuba J and Iwaniak A, “Database of protein and bioactive peptide sequences.” In Nutraceutical Proteins and Peptides in Health and Disease. Edited by Mine V Shahidi F. Taylor & Francis 2006:543-563.

15. Korhonen H and Pihlanto-Leppala A. “Food-derived bioactive peptides: Opportunities for designing future foods”. Current Pharmaceutical Design 2003; 9:1297–1308.

16. Cho H. S, Lee H H, Choi S J, Kim K. J, Jeun S H, Li Q Z, and Sung K W. “Forskolin Enhances Synaptic Transmission in Rat Dorsal Striatum through NMDA Receptors and PKA in Different Phases”. Korean J Physiol Pharmacology 2008; 12: 293-297.

17. Ovadia O, Linde Y, Haskell L. C, Dirain M. L. Sheynis T, and Jelinek R. “The effect of backbone cyclization on PK/PD properties of bioactive peptide-peptoid hybrids: The melanocortin agonist paradigm”. Bioorganic & Medicinal Chemistry 2010; 18: 580–589.

18. Jones S, Holm T, Mäger I, Langel Ü, and Howl J. “Characterization of Bioactive Cell Penetrating Peptides from Human Cytochrome c: Protein Mimicry and the Development of a Novel Apoptogenic Agent” Chemistry & Biology 2010; 17: 735-744.

19. Schwede T., Kopp J., Guex N. and Peitsch M C. “SWISS-MODEL: an automated protein homology-modeling server”. Nucleic Acids Res.2003; 31:3381–3385.

20. Koh I. Y, et al. “EVA: evaluation of protein structure prediction servers”. Nucleic Acids. Res. 2003; 31:3311–3315.

21. Melo F. and Feytmans E. “Assessing protein structures with non-local atomic interaction energy”. Journal of Molecular Biology 1998; 277:1141–1152.