

## BIOINFORMATICS PRACTICAL 2

K PAVANI  
BE19B023

1. DDBJ: 37,759,253  
Genbank: 27,979,315  
EMBL: 37,723,166
2. Given sequence: AY330867 is for Synthetic construct human lysozyme mRNA  
GC content= (Number of G nucleotides + number of C nucleotides)/ total number of nucleotides  
GC content = 46.527778 (calculated from seq2Feature tool).

3. a. DDBJ:  
Locus  
Definition  
Accession  
Version  
Keywords  
Source  
Organism  
Reference  
Authors  
Title  
Journal  
Features  
Source  
CDS

- b. Genbank:  
Locus  
Definition  
Accession  
Version  
Keywords  
Source  
Organism  
Reference  
Authors  
Title  
Journal  
Features

- Source  
CDS

- c. EMBL:  
**Organism:**  
**Accession:**  
**Mol Type:**  
**Topology:**  
**Base Count:**

**Dataclass:****Tax Division:****Md5 Checksum:**

EMBL will almost have the same contents.

EMBL, main sources for DNA and RNA sequences are direct submissions from individual researchers, genome sequencing projects and patent applications.

GenBank didn't show the base count while DDBJ and EMBL will show the base count.

4. As found in PubMed, there are 23 papers on the topic "discrimination of beta barrel membrane proteins".

They are –

1: Roumia AF, Theodoropoulou MC, Tsirigos KD, Nielsen H, Bagos PG. Landscape of Eukaryotic Transmembrane Beta Barrel Proteins. *J Proteome Res.* 2020 Mar 6;19(3):1209-1221. doi: 10.1021/acs.jproteome.9b00740. Epub 2020 Feb 11. PMID: 32008325.

2: Tsaousis GN, Hamodrakas SJ, Bagos PG. Predicting Beta Barrel Transmembrane Proteins Using HMMs. *Methods Mol Biol.* 2017;1552:43-61. doi: 10.1007/978-1-4939-6753-7\_4. PMID: 28224490.

3: Madeo G, Savojardo C, Martelli PL, Casadio R. BetAware-Deep: An Accurate Web Server for Discrimination and Topology Prediction of Prokaryotic Transmembrane  $\beta$ -barrel Proteins. *J Mol Biol.* 2021 May 28;433(11):166729. doi: 10.1016/j.jmb.2020.166729. Epub 2020 Dec 3. PMID: 33972021.

4: Ou YY, Gromiha MM, Chen SA, Suwa M. TMBETADISC-RBF: Discrimination of beta-barrel membrane proteins using RBF networks and PSSM profiles. *Comput Biol Chem.* 2008 Jun;32(3):227-31. doi: 10.1016/j.compbiolchem.2008.03.002. Epub 2008 Mar 18. PMID: 18434251.

5: Klinger A, Gosch V, Bodensohn U, Ladig R, Schleiff E. The signal distinguishing between targeting of outer membrane  $\beta$ -barrel protein to plastids and mitochondria in plants. *Biochim Biophys Acta Mol Cell Res.* 2019 Apr;1866(4):663-672. doi: 10.1016/j.bbamcr.2019.01.004. Epub 2019 Jan 8. PMID: 30633951.

6: Tsirigos KD, Elofsson A, Bagos PG. PRED-TMBB2: improved topology prediction and detection of beta-barrel outer membrane proteins. *Bioinformatics.* 2016 Sep 1;32(17):i665-i671. doi: 10.1093/bioinformatics/btw444. PMID: 27587687.

7: Gromiha MM, Suwa M. Current developments on beta-barrel membrane proteins: sequence and structure analysis, discrimination and prediction. *Curr Protein Pept Sci.* 2007 Dec;8(6):580-99. doi: 10.2174/138920307783018712. PMID: 18220845.

8: Mirus O, Schleiff E. Prediction of beta-barrel membrane proteins by searching for restricted domains. *BMC Bioinformatics.* 2005 Oct 14;6:254. doi: 10.1186/1471-2105-6-254. PMID: 16225682; PMCID: PMC1280923.

9: Gromiha MM, Yabuki Y. Functional discrimination of membrane proteins using machine learning techniques. *BMC Bioinformatics.* 2008 Mar 3;9:135. doi: 10.1186/1471-2105-9-135. PMID: 18312695; PMCID: PMC2375119.

10: Dewald AH, Hodges JC, Columbus L. Physical determinants of  $\beta$ -barrel membrane protein folding in lipid vesicles. *Biophys J.* 2011 May 4;100(9):2131-40. doi: 10.1016/j.bpj.2011.03.025. PMID: 21539780; PMCID: PMC3149260.

- 11: Farrand AJ, Hotze EM, Sato TK, Wade KR, Wimley WC, Johnson AE, Tweten RK. The Cholesterol-dependent Cytolysin Membrane-binding Interface Discriminates Lipid Environments of Cholesterol to Support  $\beta$ -Barrel Pore Insertion. *J Biol Chem*. 2015 Jul 17;290(29):17733-17744. doi: 10.1074/jbc.M115.656769. Epub 2015 Jun 1. PMID: 26032415; PMCID: PMC4505022.
- 12: Le NQK, Sandag GA, Ou YY. Incorporating post translational modification information for enhancing the predictive performance of membrane transport proteins. *Comput Biol Chem*. 2018 Dec;77:251-260. doi: 10.1016/j.compbiolchem.2018.10.010. Epub 2018 Oct 18. PMID: 30393099.
- 13: Liu Q, Zhu Y, Wang B, Li Y. Identification of beta-barrel membrane proteins based on amino acid composition properties and predicted secondary structure. *Comput Biol Chem*. 2003 Jul;27(3):355-61. doi: 10.1016/s1476-9271(02)00085-3. PMID: 12927109.
- 14: Bagos PG, Liakopoulos TD, Spyropoulos IC, Hamodrakas SJ. A Hidden Markov Model method, capable of predicting and discriminating beta-barrel outer membrane proteins. *BMC Bioinformatics*. 2004 Mar 15;5:29. doi: 10.1186/1471-2105-5-29. PMID: 15070403; PMCID: PMC385222.
- 15: Taylor PD, Toseland CP, Attwood TK, Flower DR. A predictor of membrane class: Discriminating alpha-helical and beta-barrel membrane proteins from non- membranous proteins. *Bioinformation*. 2006 Oct 7;1(6):208-13. PMID: 17597890; PMCID: PMC1891694.
- 16: Tran Vdu T, Chassignet P, Steyaert JM. Supersecondary structure prediction of transmembrane beta-barrel proteins. *Methods Mol Biol*. 2013;932:277-94. doi: 10.1007/978-1-62703-065-6\_17. PMID: 22987359.
- 17: Hoogerheide DP, Rostovtseva TK, Bezrukov SM. Exploring lipid-dependent conformations of membrane-bound  $\alpha$ -synuclein with the VDAC nanopore. *Biochim Biophys Acta Biomembr*. 2021 Sep 1;1863(9):183643. doi: 10.1016/j.bbamem.2021.183643. Epub 2021 May 7. PMID: 33971161; PMCID: PMC8255272.
- 18: Martelli PL, Fariselli P, Krogh A, Casadio R. A sequence-profile-based HMM for predicting and discriminating beta barrel membrane proteins. *Bioinformatics*. 2002;18 Suppl 1:S46-53. doi: 10.1093/bioinformatics/18.suppl\_1.s46. PMID: 12169530.
- 19: Pelosi P. Odorant-binding proteins: structural aspects. *Ann N Y Acad Sci*. 1998 Nov 30;855:281-93. doi: 10.1111/j.1749-6632.1998.tb10584.x. PMID: 9929622.
- 20: Deng Y, Liu Q, Li YX. Scoring hidden Markov models to discriminate beta- barrel membrane proteins. *Comput Biol Chem*. 2004 Jul;28(3):189-94. doi: 10.1016/j.compbiolchem.2004.02.004. PMID: 15261149.
- 21: Garrow AG, Westhead DR. A consensus algorithm to screen genomes for novel families of transmembrane beta barrel proteins. *Proteins*. 2007 Oct 1;69(1):8-18. doi: 10.1002/prot.21439. PMID: 17557332.
- 22: Ghequire MGK, Swings T, Michiels J, Buchanan SK, De Mot R. Hitting with a BAM: Selective Killing by Lectin-Like Bacteriocins. *mBio*. 2018 Mar 20;9(2):e02138-17. doi: 10.1128/mBio.02138-17. PMID: 29559575; PMCID: PMC5874912.

- 23: Selz KA, Mandell AJ, Shlesinger MF. Hydrophobic free energy eigenfunctions of pore, channel, and transporter proteins contain beta-burst patterns. *Biophys J*. 1998 Nov;75(5):2332-42. doi: 10.1016/S0006-3495(98)77677-5. PMID: 9788928; PMCID: PMC1299907.
- 24: Hu J, Yan C. A method for discovering transmembrane beta-barrel proteins in Gram-negative bacterial proteomes. *Comput Biol Chem*. 2008 Aug;32(4):298-301. doi: 10.1016/j.compbiolchem.2008.03.010. Epub 2008 Apr 1. PMID: 18467177.
- 25: Lazaridis T. Structural Determinants of Transmembrane  $\beta$ -Barrels. *J Chem Theory Comput*. 2005 Jul;1(4):716-22. doi: 10.1021/ct050055x. PMID: 26641693.
- 26: Savojardo C, Fariselli P, Casadio R. Improving the detection of transmembrane  $\beta$ -barrel chains with N-to-1 extreme learning machines. *Bioinformatics*. 2011 Nov 15;27(22):3123-8. doi: 10.1093/bioinformatics/btr549. Epub 2011 Oct 3. PMID: 21967762.
- 27: Nugent T, Jones DT. Membrane protein orientation and refinement using a knowledge-based statistical potential. *BMC Bioinformatics*. 2013 Sep 18;14:276. doi: 10.1186/1471-2105-14-276. PMID: 24047460; PMCID: PMC3852961.
- 28: Yan RX, Chen Z, Zhang Z. Outer membrane proteins can be simply identified using secondary structure element alignment. *BMC Bioinformatics*. 2011 Mar 17;12:76. doi: 10.1186/1471-2105-12-76. PMID: 21414186; PMCID: PMC3072342.
- 29: Garrow AG, Agnew A, Westhead DR. TMB-Hunt: an amino acid composition based method to screen proteomes for beta-barrel transmembrane proteins. *BMC Bioinformatics*. 2005 Mar 15;6:56. doi: 10.1186/1471-2105-6-56. PMID: 15769290; PMCID: PMC1274253.
- 30: E-komon T, Burchmore R, Herzyk P, Davies R. Predicting the outer membrane proteome of *Pasteurella multocida* based on consensus prediction enhanced by results integration and manual confirmation. *BMC Bioinformatics*. 2012 Apr 27;13:63. doi: 10.1186/1471-2105-13-63. PMID: 22540951; PMCID: PMC3403877.
- 31: Gresock MG, Postle K. Going Outside the TonB Box: Identification of Novel FepA-TonB Interactions *In Vivo*. *J Bacteriol*. 2017 Apr 25;199(10):e00649-16. doi: 10.1128/JB.00649-16. PMID: 28264993; PMCID: PMC5405211.
- 32: Luchian T, Park Y, Asandei A, Schiopu I, Mereuta L, Apetrei A. Nanoscale Probing of Informational Polymers with Nanopores. Applications to Amyloidogenic Fragments, Peptides, and DNA-PNA Hybrids. *Acc Chem Res*. 2019 Jan 15;52(1):267-276. doi: 10.1021/acs.accounts.8b00565. Epub 2019 Jan 3. PMID: 30605305.
- 33: Ayub M, Stoddart D, Bayley H. Nucleobase Recognition by Truncated  $\alpha$ -Hemolysin Pores. *ACS Nano*. 2015 Aug 25;9(8):7895-903. doi: 10.1021/nn5060317. Epub 2015 Jul 28. PMID: 26114210; PMCID: PMC4830132.
- 34: Bigelow HR, Petrey DS, Liu J, Przybylski D, Rost B. Predicting transmembrane beta-barrels in proteomes. *Nucleic Acids Res*. 2004 May 11;32(8):2566-77. doi: 10.1093/nar/gkh580. PMID: 15141026; PMCID: PMC419468.
- 35: Mereuta L, Asandei A, Schiopu I, Park Y, Luchian T. Nanopore-Assisted, Sequence-Specific Detection, and Single-Molecule Hybridization Analysis of Short, Single-Stranded

DNAs. *Anal Chem.* 2019 Jul 2;91(13):8630-8637. doi: 10.1021/acs.analchem.9b02080. Epub 2019 Jun 13. PMID: 31194518.

36: Garrow AG, Agnew A, Westhead DR. TMB-Hunt: a web server to screen sequence sets for transmembrane beta-barrel proteins. *Nucleic Acids Res.* 2005 Jul 1;33(Web Server issue):W188-92. doi: 10.1093/nar/gki384. PMID: 15980452; PMCID: PMC1160145.

37: Casadio R, Jacoboni I, Messina A, De Pinto V. A 3D model of the voltage- dependent anion channel (VDAC). *FEBS Lett.* 2002 Jun 5;520(1-3):1-7. doi: 10.1016/s0014-5793(02)02758-8. PMID: 12044860.

38: Magalashvili L, Pechatnikov I, Wexler HM, Nitzan Y. Isolation and characterization of the Omp-PA porin from *Porphyromonas asaccharolytica*, determination of the omp-PA gene sequence and prediction of Omp-PA protein structure. *Anaerobe.* 2007 Apr;13(2):74-82. doi: 10.1016/j.anaerobe.2006.11.003. Epub 2007 Jan 16. PMID: 17229581.

39: Bigelow H, Rost B. PROFmb: a web server for predicting bacterial transmembrane beta barrel proteins. *Nucleic Acids Res.* 2006 Jul 1;34(Web Server issue):W186-8. doi: 10.1093/nar/gkl262. PMID: 16844988; PMCID: PMC1538807.

40: Lertjuthaporn S, Cicala C, Van Ryk D, Liu M, Yolitz J, Wei D, Nawaz F, Doyle A, Horowitch B, Park C, Lu S, Lou Y, Wang S, Pan R, Jiang X, Villinger F, Byraredy SN, Santangelo PJ, Morris L, Wibmer CK, Biris K, Mason RD, Gorman J, Hiatt J, Martinelli E, Roederer M, Fujikawa D, Gorini G, Franchini G, Arakelyan A, Ansari AA, Pattanapanyasat K, Kong XP, Fauci AS, Arthos J. Select gp120 V2 domain specific antibodies derived from HIV and SIV infection and vaccination inhibit gp120 binding to  $\alpha 4\beta 7$ . *PLoS Pathog.* 2018 Aug 28;14(8):e1007278. doi: 10.1371/journal.ppat.1007278. PMID: 30153309; PMCID: PMC6130882.

41: Fahie MA, Yang B, Mullis M, Holden MA, Chen M. Selective Detection of Protein Homologues in Serum Using an OmpG Nanopore. *Anal Chem.* 2015 Nov 3;87(21):11143-9. doi: 10.1021/acs.analchem.5b03350. Epub 2015 Oct 23. PMID: 26451707; PMCID: PMC5065927.

42: Ruhe ZC, Subramanian P, Song K, Nguyen JY, Stevens TA, Low DA, Jensen GJ, Hayes CS. Programmed Secretion Arrest and Receptor-Triggered Toxin Export during Antibacterial Contact-Dependent Growth Inhibition. *Cell.* 2018 Nov 1;175(4):921-933.e14. doi: 10.1016/j.cell.2018.10.033. PMID: 30388452; PMCID: PMC6333426.

43: Scott DC, Cao Z, Qi Z, Bauler M, Igo JD, Newton SM, Klebba PE. Exchangeability of N termini in the ligand-gated porins of *Escherichia coli*. *J Biol Chem.* 2001 Apr 20;276(16):13025-33. doi: 10.1074/jbc.M011282200. Epub 2001 Jan 19. PMID: 11278876.

44: Hayashi K, Longenecker KL, Koenig P, Prashar A, Hampl J, Stoll V, Vivona S. Structure of human DPEP3 in complex with the SC-003 antibody Fab fragment reveals basis for lack of dipeptidase activity. *J Struct Biol.* 2020 Jul 1;211(1):107512. doi: 10.1016/j.jsb.2020.107512. Epub 2020 Apr 20. PMID: 32325220.

45: Lin H. The modified Mahalanobis Discriminant for predicting outer membrane proteins by using Chou's pseudo amino acid composition. *J Theor Biol.* 2008 May 21;252(2):350-6. doi: 10.1016/j.jtbi.2008.02.004. Epub 2008 Feb 12. PMID: 18355838.

46: Nishida M, Nagata K, Hachimori Y, Horiuchi M, Ogura K, Mandiyan V, Schlessinger J, Inagaki F. Novel recognition mode between Vav and Grb2 SH3 domains. *EMBO J.* 2001 Jun 15;20(12):2995-3007. doi: 10.1093/emboj/20.12.2995. PMID: 11406576; PMCID: PMC150196.

47: Schleiff E, Shore GC, Goping IS. Human mitochondrial import receptor, Tom20p. Use of glutathione to reveal specific interactions between Tom20-glutathione S-transferase and mitochondrial precursor proteins. *FEBS Lett.* 1997 Mar 10;404(2-3):314-8. doi: 10.1016/s0014-5793(97)00145-2. PMID: 9119086.

48: Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci U S A.* 2009 May 12;106(19):7702-7. doi: 10.1073/pnas.0901054106. Epub 2009 Apr 20. PMID: 19380741; PMCID: PMC2683137.

49: Marchalonis JJ, Kaymaz H, Dedeoglu F, Schluter SF, Yocum DE, Edmundson AB. Human autoantibodies reactive with synthetic autoantigens from T-cell receptor beta chain. *Proc Natl Acad Sci U S A.* 1992 Apr 15;89(8):3325-9. doi: 10.1073/pnas.89.8.3325. PMID: 1565623; PMCID: PMC48859.

50: Kasper KJ, Xi W, Rahman AK, Nooh MM, Kotb M, Sundberg EJ, Madrenas J, McCormick JK. Molecular requirements for MHC class II alpha-chain engagement and allelic discrimination by the bacterial superantigen streptococcal pyrogenic exotoxin C. *J Immunol.* 2008 Sep 1;181(5):3384-92. doi: 10.4049/jimmunol.181.5.3384. PMID: 18714010.

5. Author: Karthik R

[https://scholar.google.com/citations?hl=en&user=qkn2yIIAAAAJ&view\\_op=list\\_works&sortby=title](https://scholar.google.com/citations?hl=en&user=qkn2yIIAAAAJ&view_op=list_works&sortby=title)

It has 74 papers.

6. For the paper - *Cell*. 2008 Dec 26;135(7):1158-9 - Mitochondrial beta-barrel proteins, 28 similar articles are reported to be found on PubMed.

7. 401 articles listed in PUBMED for the journal "Nature" and 626 article in SCOPUS.

8. The h-index = 103

Total = 47,640 Citations for Burkha Rost.

9.

### EC Tree

- └ 3 Hydrolases
  - └ 3.4 Acting on peptide bonds (peptidases)
    - └ 3.4.11 Aminopeptidases
      - └ 3.4.11.4 tripeptide aminopeptidase

**Specify your search results**

**Mark a special word or phrase in this record:**  **Mark!**

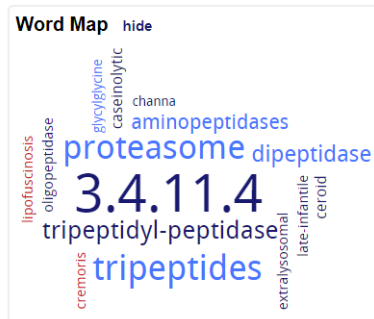
**Search Reference ID:**  **Search**

**Search UniProt Accession:**  **Search**

**Select one or more organisms in this record:** ☒

All organisms  
 Arabidopsis thaliana  
 Bacillus subtilis  
 Bos taurus  
 Caenorhabditis elegans

**Submit**



The enzyme appears in viruses and cellular organisms

### Reaction Schemes [hide](#)

release of the N-terminal residue from a tripeptide [×](#)

### Synonyms

alanine-phenylalanine-proline arylamidase, alpha-aminoacyl-dipeptide hydrolase, aminoexotripeptidase, aminopeptidase, human liver, aminotripeptidase, arginy imidoendopeptidase, lymphopeptidase, [more](#)

[△ top](#) [print](#) [hide](#) [Go to Reaction Search](#)

REACTION ▲▼	REACTION DIAGRAM	COMMENTARY ▲▼ <a href="#">×</a>	ORGANISM ▲▼	UNIPROT ▲▼	LITERATURE ▲▼
release of the N-terminal residue from a tripeptide <a href="#">🔗</a>		-	-	-	-

## 3- Hydrolases (class)

## 3.4- Acting on peptide bonds (peptidases)

## 3.4.11- Aminopeptidases

## 3.4.11.4- Tripeptide aminopeptidase

Function- Release of the N-terminal residue from a tripeptide.

10.

### Catalytic Residues Roles

UniProt	PDB* (1ct9)		
Cys2 (N-term)	Ala1A (N-term)	Acts as a general acid/base to activate the cysteine nucleophile.	proton acceptor, proton donor
Leu51 (main-C)	Leu50A (main-C)	Helps stabilise the reactive intermediates formed.	hydrogen bond acceptor, electrostatic stabiliser
Thr322, Arg325	Thr321A, Arg324A	Bind and stabilise the phosphate groups of the ATP and reactive intermediates formed.	hydrogen bond donor, electrostatic stabiliser
Cys2	Ala1A	Acts as a catalytic nucleophile in the glutaminase domain reaction.	covalently attached, hydrogen bond acceptor, nucleofuge, nucleophile, proton acceptor, proton donor
Gly76 (main-N), Asn75	Gly75A (main-N), Asn74A	Forms the oxyanion hole that stabilises the reactive intermediates and transition states formed.	hydrogen bond donor, electrostatic stabiliser

CATH superfamily 3.60.20.10 is of Aminohydrolase, N-terminal nucleophile domain. And CATH superfamily 3.40.50.620 is of HUPs.

11.

Organism	Scientific name	Taxonomy ID	Number of Chromosomes
Human	Homo sapiens	9606	23
Cat	<i>Felis catus</i>	9685	20
Dog	Canis lupus familiaris	9615	40
House mouse	Mus musculus	10090	40
Onion	Allium cepa. L	4679	9
Thale cress	Arabidopsis thaliana	3702	5

12. The Entrez Programming Utilities (E-utilities) are a group of nine server-side programmes that have a secure interface into the National Center for Biotechnology Information's

Entrez query and database system. The E-utilities employ a standard URL syntax that converts a string of characters into a URL. A common set of inputs that are converted into the values needed by different NCBI software components to look for it and collect the data that has been requested.

#### Basic Downloading

```
efetch.fcgi?db=<database>&id=<uid_list>&rettype=<retrieval_type>
&retmode=<retrieval_mode>
```

Input: List of UIDs (&id); Entrez database (&db); Retrieval type (&rettype); Retrieval mode (&retmode)

Output: Formatted data records as specified

*Example: Download nuccore GIs 34577062 and 24475906 in FASTA format*

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?db=nuccore&id=34577062,24475906&rettype=fasta&retmode=text>

13.

A. Protein Properties: (i) AAindex (ii) Cybase

B. Small molecules (Structure related): (i) ChEBI-Chemical Entities of Biological Interest (ii) Hemolytik

C. Cancer gene databases: (i) Cancer3D (ii) BCCTBbp