

IQB Assignment 3

Name : Aayush Gakhar
Roll no : 2020006

>protein

MALPAGPAEAAACALCQRAPREPVRADCGHRFCRACVVRFWAEEDGPFPCPECADDCWQRA
VEPGRPPLSRRLLEAAAAAPARDGPASEAALQLLCRADAGPLCAACRMAAGPEPPEWE

1.

Submitting the given sequence for blastp, we get the following result:

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

DownloadSelect columnsShow100

select all

17 sequences selected

GenPeptGraphicsDistance tree of resultsMultiple alignmentMSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Solution structure of the RING domain of the human tripartite motif-containing protein 39 [Homo sapiens]	Homo sapiens	45.1	45.1	27%	3e-07	51.52%	58	2ECJ_A
<input checked="" type="checkbox"/> Chain A, BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN [Homo sapiens]	Homo sapiens	45.1	45.1	35%	1e-06	35.71%	112	1JM7_A
<input checked="" type="checkbox"/> Chain M, Isoform 7 of Breast cancer type 1 susceptibility protein [Homo sapiens]	Homo sapiens	45.1	45.1	35%	2e-06	35.71%	124	7LYB_M
<input checked="" type="checkbox"/> Chain A, BRCA1, Ubiquitin-conjugating enzyme E2 D3 [Homo sapiens]	Homo sapiens	43.5	43.5	35%	1e-05	35.71%	258	7JZV_A
<input checked="" type="checkbox"/> Complex of TRIM25 RING with UbcH5-Ub [Homo sapiens]	Homo sapiens	41.2	41.2	27%	2e-05	45.45%	85	5FER_A
<input checked="" type="checkbox"/> TRIM25 RING domain in complex with Ubc13-Ub conjugate [Homo sapiens]	Homo sapiens	40.8	40.8	27%	3e-05	45.45%	86	5EYA_F
<input checked="" type="checkbox"/> Structure of the TRIM25 coiled-coil [Homo sapiens]	Homo sapiens	41.6	41.6	27%	9e-05	45.45%	630	4CFG_A
<input checked="" type="checkbox"/> Solution structure of the RING domain (1-66) from tripartite motif-containing protein 31 [Homo sapiens]	Homo sapiens	38.5	38.5	35%	2e-04	35.71%	73	2YSL_A
<input checked="" type="checkbox"/> Solution structure of the Zinc finger, C3HC4 type (RING finger) domain Tripartite motif protein 30 [Mus musculus]	Mus musculus	37.4	37.4	27%	7e-04	47.22%	85	2ECW_A
<input checked="" type="checkbox"/> Structure of the Trim69 RING domain [Homo sapiens]	Homo sapiens	37.7	37.7	30%	8e-04	43.24%	128	6YXE_A
<input checked="" type="checkbox"/> Solution structure of the RING domain (1-56) from tripartite motif-containing protein 31 [Homo sapiens]	Homo sapiens	36.6	36.6	27%	0.001	42.42%	63	2YSJ_A
<input checked="" type="checkbox"/> Solution Structure of the Ring Domain of Human TRAF6 [Homo sapiens]	Homo sapiens	33.1	33.1	27%	0.019	45.45%	63	2JMD_A
<input checked="" type="checkbox"/> Chain A, TNF receptor-associated factor 6 [Homo sapiens]	Homo sapiens	33.1	33.1	17%	0.023	57.14%	86	2ECI_A
<input checked="" type="checkbox"/> Crystal structure of TRAF6 in complex with Ubc13 in the P1 space group [Homo sapiens]	Homo sapiens	33.1	33.1	27%	0.036	45.45%	118	3HCT_A
<input checked="" type="checkbox"/> Chain B, TNF receptor-associated factor 6 [Homo sapiens]	Homo sapiens	33.1	33.1	27%	0.037	45.45%	107	7L3L_B
<input checked="" type="checkbox"/> Cryo-EM structure of RAG in complex with 12-RSS and 23-RSS nicked DNA intermediates [Escherichia coli K-12]	Escherichia coli...	33.9	33.9	26%	0.043	43.75%	1159	6DBI_A
<input checked="" type="checkbox"/> Cryo-EM structure of RAG in complex with 12-RSS and 23-RSS substrate DNAs [Escherichia coli K-12]	Escherichia coli...	33.9	33.9	26%	0.043	43.75%	1159	6DBV_A

We can see that Chain A, BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN [Homo sapiens] can serve as the best template for modelling the E3 ubiquitin-protein ligase structure.

1. Query cover: It has the highest query cover of all the results of 35%. this indicates it is more aligned to our query protein.
2. E value: E value is also second highest. e value shows what percentage of local alignment is there. The high E value makes it a good choice.
3. Score: The score is also very good of 45.1%.

2.

Below is the alignment of the template we chose.

Descriptions Graphic Summary **Alignments** Taxonomy

Alignment view Pairwise [Restore defaults](#) Download

1 sequences selected

[Download](#) [GenPept](#) [Graphics](#) [Next](#) [Previous](#) [Descriptions](#)

Chain A, BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN [Homo sapiens]
Sequence ID: [1JM7_A](#) Length: 112 Number of Matches: 1

Range 1: 32 to 73 [GenPept](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
45.1 bits(105)	1e-06	Compositional matrix adjust.	15/42(36%)	26/42(61%)	0/42(0%)

Query 20 REPVRADCGHRCRACVVRFWAEEGGPFPCPECADDCWQRAV 61
+EPV C H FC+ C+++ ++ GP CP C +D +R++
Sbjct 32 KEPVSTKCDHIFCKFCMLKLLNQKKGPSQCPLCKNDITKRS 73

Related Information
[Structure](#) - 3D structure displays

We don't get coverage after protein number 60 in any of the .
The coverage is highest among all other options.

Descriptions **Graphic Summary** Alignments Taxonomy

[hover to see the title](#) [click to show alignments](#) ☒ Show Conserved Domains Alignment Scores ☐ < 40 ☐ 40 - 50 ☐ 50 - 80 ☐ 80 - 200 ☐ >= 200

1 sequences selected **Putative conserved domains have been detected, click on the image below for detailed results.**

Query seq. [View alignment](#)
Superfamilies [View alignment](#)

Distribution of the top 1 Blast Hits on 1 subject sequences

Query

1 20 40 60 80 100 120

So we can only use this as our template.

3.

SOLUTION NMR method was used to solve this structure. There are 2 chains in this structure a and b. We need the chain a.

Structure Summary3D ViewAnnotationsExperimentSequenceGenomeVersions

NMR Ensemble

3D View: Structure | 1D-3D View | Validation Report | Ligand Interaction

Macromolecule Content

- Total Structure Weight: 26.25 kDa
- Atom Count: 1573
- Modelled Residue Count: 200
- Deposited Residue Count: 229
- Unique protein chains: 2

1JM7

Solution structure of the BRCA1/BARD1 RING-domain heterodimer

PDB DOI: 10.2210/pdb1JM7/pdb

Classification: ANTITUMOR

Organism(s): Homo sapiens

Expression System: Escherichia coli BL21(DE3)

Mutation(s): No

Deposited: 2001-07-17 Released: 2001-10-03

Deposition Author(s): Brzovic, P.S., Rajagopal, P., Hoyt, D.W., King, M.-C., Klevit, R.E.

Experimental Data Snapshot

Method: SOLUTION NMR

Conformers Calculated: 25

Conformers Submitted: 14

Selection Criteria: structures with the least restraint violations, structures with the lowest energy

wwPDB Validation

3D ReportFull Report

Metric	Percentile Ranks	Value
Clashscore		24
Ramachandran outliers		5.1%
Sidechain outliers		15.2%

WorseBetter

Percentile relative to all structures

Percentile relative to all NMR structures

This is version 1.3 of the entry. See complete history.

LiteratureDownload Primary Citation

Macromolecules

Find similar proteins by: Sequence (by identity cutoff) | 3D Structure

Entity ID: 1

Molecule	Chains	Sequence Length	Organism	Details	Image
BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN	A	112	Homo sapiens	Mutation(s): 0 Gene Names: BRCA1, RN F53 EC: 2.3.2.27	

UniProt & NIH Common Fund Data Resources

Find proteins for P38398 (Homo sapiens)

Go to UniProtKB: P38398

PHAROS: P38398GTEX: ENSG0000012046

Entity Groups

Sequence Clusters30% Identity50% Identity70% Identity90% Identity95% Identity100% Identity

UniProt GroupP38398

Protein Feature ViewExpand

Reference Sequence1JM7_1

PDB ENTITY 1JM7_1

UNIPROT P38398

UNMODELED A

ARTIFACT

HYDROPATHY

DISORDER

DISORDERED BINDING

PFAM

MDLSALRVEEVQNVINAMOKILECPICLELIKPEPVSTKCDHIFCKFCMLKLLNQKKGPSQCPLCKNDITKRSLOESTRFSQLVEELKIIICAFQDITGLYANSYNFAKKGK

MDLSALRVEEVQNVINAMOKILECPICLELIKPEPVSTKCDHIFCKFCMLKLLNQKKGPSQCPLCKNDITKRSLOESTRFSQLVEELKIIICAFQDITGLYANSYNFAKK

Entity ID: 2

Molecule	Chains	Sequence Length	Organism	Details	Image
BRCA1-ASSOCIATED RING DOMAIN PROTEIN 1	B	117	Homo sapiens	Mutation(s): 0 Gene Names: BARD1 EC: 2.3.2.27	

UniProt & NIH Common Fund Data Resources

Find proteins for Q99728 (Homo sapiens)

Go to UniProtKB: Q99728

PHAROS: Q99728GTEX: ENSG00000138376

Entity Groups

Sequence Clusters30% Identity50% Identity70% Identity90% Identity95% Identity100% Identity

UniProt GroupQ99728

Protein Feature ViewExpand

Reference Sequence1JM7_2

PDB ENTITY 1JM7_2

UNIPROT Q99728

UNMODELED B

ARTIFACT

HYDROPATHY

DISORDER

DISORDERED BINDING

PFAM

MEPDGRGAWAHSRAALDRLEKLLRCSRCTNIIIREPVCLGGCEHIFCSNCVSDCIGTGCPVCYTPAWIQDLKINRQLDSMIQLCSKLRNLLHONELSDLKEDKPKRSLFNDAGNKKGK

MEPDGRGAWAHSRAALDRLEKLLRCSRCTNIIIREPVCLGGCEHIFCSNCVSDCIGTGCPVCYTPAWIQDLKINRQLDSMIQLCSKLRNLLHONELSDLKEDKPKRSLFNDAGNKK

The chain B is different from our chain of interest.

