

Protein Interactions: Computational Techniques

Assignment 1 : BE21B037

1. Get the statistics for the protein-protein complexes from RCSB PDB.

a. No. of homo and hetero dimers

Screenshot of the RCSB PDB Advanced Search Query Builder interface.

Search Bar: Search, Query History, Browse Annotations, MyPDB

Query: Number of Protein Instances (Chains) per Assembly = 2 AND Number of DNA Instances (Chains) per Assembly = 0 AND Number of RNA Instances (Chains) per Assembly = 0 AND Number of Distinct Protein Entities = 1

Advanced Search Query Builder:

- Structure Attributes:**
 - Number of Protein Instances (Chains) per Assembly = 2
 - AND Number of DNA Instances (Chains) per Assembly = 0
 - AND Number of RNA Instances (Chains) per Assembly = 0
 - AND Number of Distinct Protein Entities = 1
- Operations:** AND / OR, Add Attribute, Add Subquery, Remove Subquery, Add Subquery
- Chemical Attributes:** Chemical Attributes, Sequence Similarity, Sequence Motif, Structure Similarity, Structure Motif, Chemical Similarity

Buttons: Return, Structures, grouped by, No Grouping, Include Computed Structure Models (CSM), Count, Clear, Search

Search Summary: This query matches 48,822 Structures

Refinements: Scientific Name of Source Organism

- Homo sapiens (9,834)
- Escherichia coli (1,724)
- Severe acute respiratory syndrome coronavirus 2 (1,422)

Results: 1 to 25 of 48,822 Structures, Page 1 of 1,953, Sort by Score, Download File, View File

Complex: 10GS, HUMAN GLUTATHIONE S-TRANSFERASE P1-1, COMPLEX WITH TER117

Search Query History Browse Annotations MyPDB

QUERY: Number of Protein Instances (Chains) per Assembly = 2 AND Number of DNA Instances (Chains) per Assembly = 0 AND Number of RNA Instances (Chains) per Assembly = 0 AND Number of Distinct Protein Entities = 2

[MyPDB Login](#) [Search API](#)

Advanced Search Query Builder [Help](#)

[Full Text](#) [?](#)

[Structure Attributes](#) [?](#)

[Help](#)

Number of Protein Instances (Chains) per Assembly	x ▼ ▼ =	▼ 2	+ NOT Count x
AND	Number of DNA Instances (Chains) per Assembly	x ▼ ▼ =	▼ 0
AND	Number of RNA Instances (Chains) per Assembly	x ▼ ▼ =	▼ 0
AND	Number of Distinct Protein Entities	x ▼ ▼ =	▼ 2

AND / OR Add Attribute Add Subquery

Remove Subquery

Add Subquery

[Chemical Attributes](#) [?](#)

[Sequence Similarity](#) [?](#)

[Sequence Motif](#) [?](#)

[Structure Similarity](#) [?](#)

[Structure Motif](#) [?](#)

[Chemical Similarity](#) [?](#)

Return Structures [?](#) grouped by No Grouping [?](#)

Include Computed Structure Models (CSM) [?](#) All Selected [Count](#) [Clear](#) [Search](#)

Search Summary This query matches 17,766 Structures

Refinements [?](#) [▶](#)



-- Tabular Report --

All Selected [Download](#)

Structure Determination Methodology

experimental (17,766)

1 to 25 of 17,766 Structures

[◀◀](#) [◀](#) Page 1 of 711 [▶](#) [▶▶](#) 25 [▼](#)

Sort by [↓ Score](#)

The query gives us 48,822 structures for homodimers and 17,766 heterodimers structures

b. No. of trimers

PROTEIN DATA BANK Models (CSM) Advanced Search | Browse Annotations Help

PDB-101 PDB EMDDataResource NAKB wwPDB Foundation PDB-Dev

Search Query History Browse Annotations MyPDB

QUERY: Number of Protein Instances (Chains) per Assembly = 3 AND Number of DNA Instances (Chains) per Assembly = 0 AND Number of RNA Instances (Chains) per Assembly = 0

MyPDB Login Search API

Advanced Search Query Builder Help

Full Text

Structure Attributes Help

	Number of Protein Instances (Chains) per Assembly	x ▾ ▾ =	▼ 3	+ NOT	Count	x
AND	Number of DNA Instances (Chains) per Assembly	x ▾ ▾ =	▼ 0	+ NOT	Count	x
AND	Number of RNA Instances (Chains) per Assembly	x ▾ ▾ =	▼ 0	+ NOT	Count	x

AND / OR Add Attribute Add Subquery Remove Subquery

Add Subquery

Chemical Attributes

Sequence Similarity

Sequence Motif

Structure Similarity

Structure Motif

Chemical Similarity

Return Structures grouped by No Grouping ?

Include Computed Structure Models (CSM) ? Count Clear Search

Search Summary This query matches 13,835 Structures.

Refinements - Tabular Report - All Selected

Structure Determination Methodology 1 to 25 of 13,835 Structures Page 1 of 554 Sort by ↓ Score

We find **13,835** trimers in PDB

c. No. of tetramers

RCSB PDB Deposit Search Visualize Analyze Download Learn About Documentation Careers COVID-19 MyPDB Contact us

PDB PROTEIN DATA BANK 224,201 Structures from the PDB 1,068,577 Computed Structure Models (CSM)

3D Structures Enter search term(s), Entry ID(s), or sequence Include CSM Advanced Search | Browse Annotations Help

PDB-101 wPDB EMDataResource NAKB wwPDB Foundation PDB-Dev

Search Query History Browse Annotations MyPDB

QUERY: Number of Protein Instances (Chains) per Assembly = 4 AND Number of DNA Instances (Chains) per Assembly = 0 AND Number of RNA Instances (Chains) per Assembly = 0

MyPDB Login Search API

Advanced Search Query Builder Help

Full Text

Structure Attributes Help

Number of Protein Instances (Chains) per Assembly	x	▼	=	▼	4	+ NOT	Count	x	
AND	Number of DNA Instances (Chains) per Assembly	x	▼	=	▼	0	+ NOT	Count	x
AND	Number of RNA Instances (Chains) per Assembly	x	▼	=	▼	0	+ NOT	Count	x
AND / OR	Add Attribute	Add Subquery	Remove Subquery						
Add Subquery									

Chemical Attributes

Sequence Similarity

Sequence Motif

Structure Similarity

Structure Motif

Chemical Similarity

Return Structures grouped by No Grouping

Include Computed Structure Models (CSM) Count Clear Search

Search Summary This query matches 21,244 Structures.

The number of tetramers being **21,244** structures

2. Find a hetero dimer complex from RCSB PDB. (Enter the PDB ID in the excel sheet)

The hetero dimer complex chosen is PDB: 2BTF

Protein details: THE STRUCTURE OF CRYSTALLINE PROFILIN-BETA-ACTIN

Classification: ACETYLATION AND ACTIN-BINDING

Organism(s): Bos taurus

Mutation(s): No

3. Briefly describe the structure and function(s) of the selected complex.

Method: X-RAY DIFFRACTION

Resolution: 2.55 Å

Total Structure Weight: 57.25 kDa

Global Symmetry: Asymmetric - C1

Global Stoichiometry: Hetero 2-mer - A1B1

Protein contains two unique chains. **Chain A, B.**

Chain Length A : 375

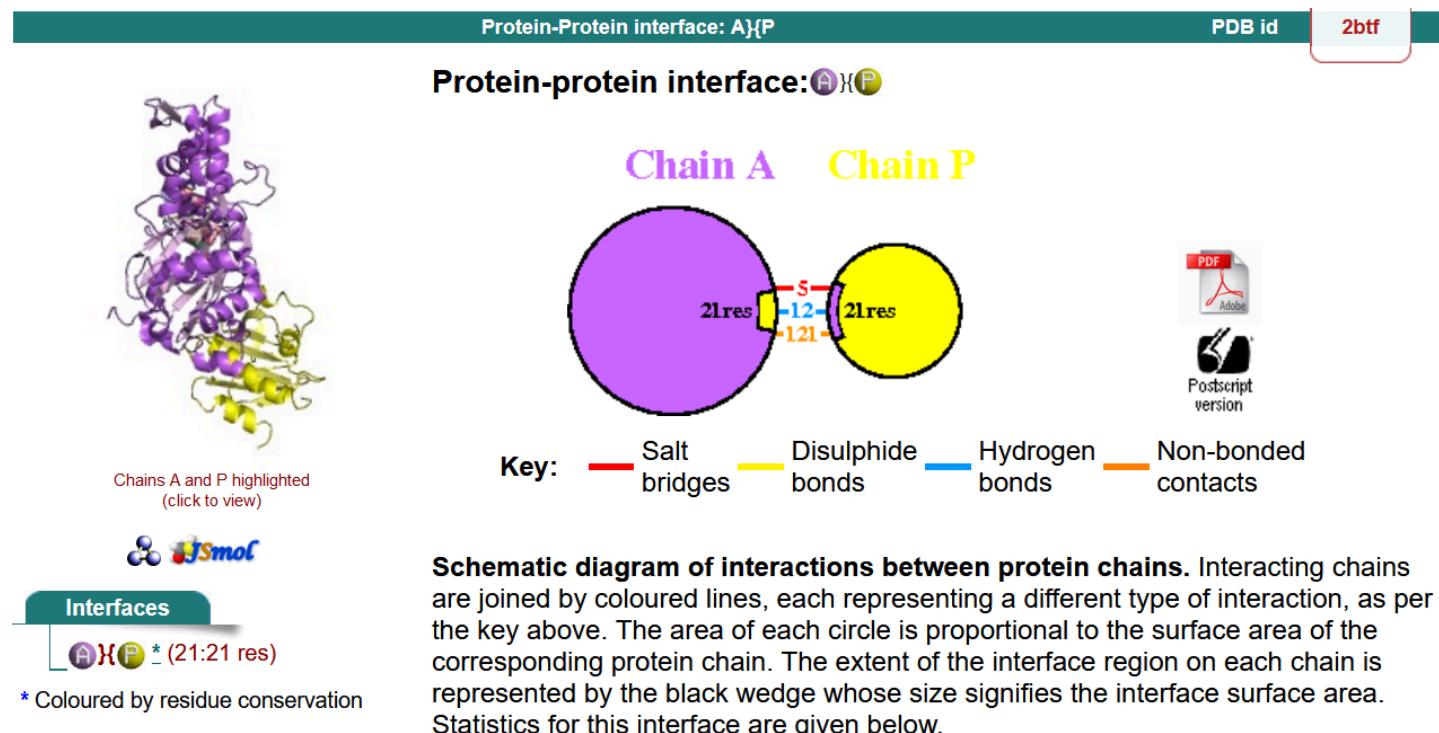
Chain Length B : 140

Ligand: ADENOSINE-5'-TRIPHOSPHATE, STRONTIUM ION

Function: The three-dimensional structure of bovine profilin-beta-actin has been solved to 2.55 Å resolution by X-ray crystallography. There are several significant local changes in the structure of beta-actin compared with alpha-actin as well as an overall 5 degrees rotation between its two major domains. Actin molecules in the crystal are organized into ribbons through intermolecular contacts like those found in oligomeric protein assemblies. Profilin forms two extensive contacts with the actin ribbon, one of which appears to correspond to the solution contact in vitro

4. Identify the binding site and interaction types using PDBSum and PDBParam

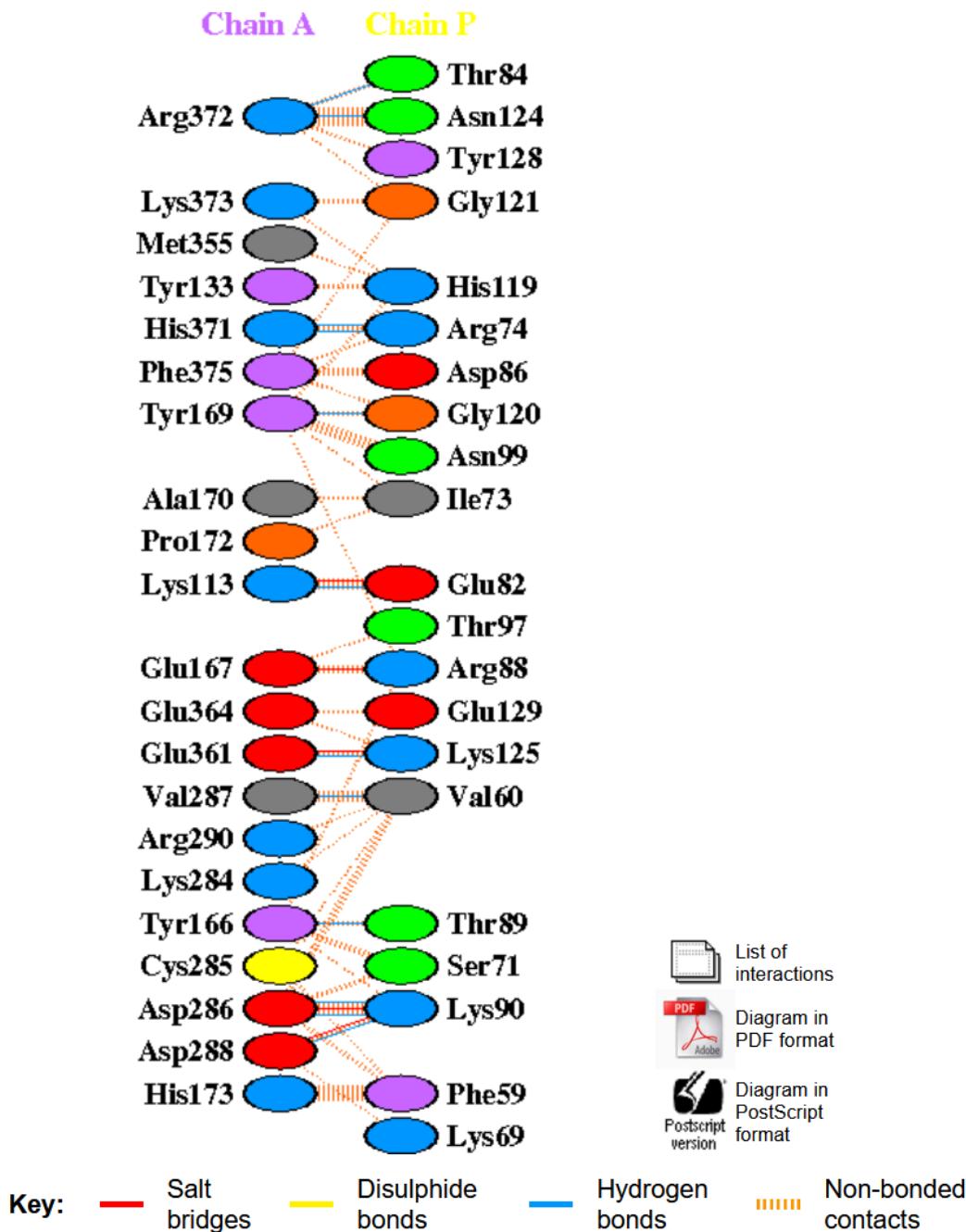
PDBSum gives the following output for the interactions between the two proteins



Interface statistics

Chain	No. of interface residues	Interface area (Å²)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	21	1052	5	-	12	121
P	21	1011				

Residue interactions across interface Coloured by residue type



The number of H-bond lines between any two residues indicates the number of potential hydrogen bonds between them. For non-bonded contacts, which can be plentiful, the width of the striped line is proportional to the number of atomic contacts.

Residue colours: Positive (H,K,R); negative (D,E); S,T,N,Q = neutral; A,V,L,I,M = aliphatic; F,Y,W = aromatic; P,G = Pro&Gly; C = cysteine.

The list of interactions are given as follows

PDB code: 2btf Chains A }{ P

Hydrogen bonds

	<---- A T O M 1 ----->					<---- A T O M 2 ----->						
	Atom	Atom	Res	Res		Atom	Atom	Res	Res			
	no.	name	name	no.	Chain	no.	name	name	no.	Chain	Distance	
1.	871	NZ	LYS	113	A	<-->	3529	OE1	GLU	82	P	3.07
2.	1275	OH	TYR	166	A	<-->	3587	O	THR	89	P	2.99
3.	1300	OH	TYR	169	A	<-->	3806	N	GLY	120	P	2.97
4.	2221	OD1	ASP	286	A	<-->	3599	NZ	LYS	90	P	3.24
5.	2222	OD2	ASP	286	A	<-->	3599	NZ	LYS	90	P	3.02
6.	2223	N	VAL	287	A	<-->	3368	O	VAL	60	P	2.82
7.	2237	OD2	ASP	288	A	<-->	3599	NZ	LYS	90	P	3.29
8.	2805	OE1	GLU	361	A	<-->	3846	NZ	LYS	125	P	3.22
9.	2877	O	HIS	371	A	<-->	3469	NH1	ARG	74	P	3.18
10.	2877	O	HIS	371	A	<-->	3470	NH2	ARG	74	P	2.82
11.	2891	NE	ARG	372	A	<-->	3836	OD1	ASN	124	P	2.59
12.	2894	NH2	ARG	372	A	<-->	3547	OG1	THR	84	P	2.82

Non-bonded contacts

	<---- A T O M 1 ----->					<---- A T O M 2 ----->						
	Atom	Atom	Res	Res		Atom	Atom	Res	Res			
	no.	name	name	no.	Chain	no.	name	name	no.	Chain	Distance	
1.	870	CE	LYS	113	A	<-->	3528	CD	GLU	82	P	3.43
2.	870	CE	LYS	113	A	<-->	3529	OE1	GLU	82	P	3.51
3.	870	CE	LYS	113	A	<-->	3530	OE2	GLU	82	P	3.09
4.	871	NZ	LYS	113	A	<-->	3528	CD	GLU	82	P	3.49
5.	871	NZ	LYS	113	A	<-->	3529	OE1	GLU	82	P	3.07
6.	871	NZ	LYS	113	A	<-->	3530	OE2	GLU	82	P	3.42
7.	1038	OH	TYR	133	A	<-->	3804	CE1	HIS	119	P	3.52
8.	1038	OH	TYR	133	A	<-->	3805	NE2	HIS	119	P	3.39
9.	1271	CD2	TYR	166	A	<-->	3443	CB	SER	71	P	3.83
10.	1271	CD2	TYR	166	A	<-->	3444	OG	SER	71	P	3.49
11.	1273	CE2	TYR	166	A	<-->	3443	CB	SER	71	P	3.84
12.	1273	CE2	TYR	166	A	<-->	3444	OG	SER	71	P	3.21
13.	1273	CE2	TYR	166	A	<-->	3577	CB	ARG	88	P	3.73
14.	1273	CE2	TYR	166	A	<-->	3578	CG	ARG	88	P	3.50
15.	1275	OH	TYR	166	A	<-->	3586	C	THR	89	P	3.67
16.	1275	OH	TYR	166	A	<-->	3587	O	THR	89	P	2.99
17.	1275	OH	TYR	166	A	<-->	3592	CA	LYS	90	P	3.55

18.	1280	CB	GLU	167	A	<-->	3580	NE	ARG	88	P	3.55
19.	1280	CB	GLU	167	A	<-->	3581	CZ	ARG	88	P	3.78
20.	1281	CG	GLU	167	A	<-->	3580	NE	ARG	88	P	3.50
21.	1281	CG	GLU	167	A	<-->	3638	OG1	THR	97	P	3.89
22.	1283	OE1	GLU	167	A	<-->	3583	NH2	ARG	88	P	3.65
23.	1294	CG	TYR	169	A	<-->	3459	CD1	ILE	73	P	3.78
24.	1297	CE1	TYR	169	A	<-->	3658	ND2	ASN	99	P	3.74
25.	1298	CE2	TYR	169	A	<-->	3582	NH1	ARG	88	P	3.53
26.	1298	CE2	TYR	169	A	<-->	3657	OD1	ASN	99	P	3.80
27.	1298	CE2	TYR	169	A	<-->	3658	ND2	ASN	99	P	3.43
28.	1299	CZ	TYR	169	A	<-->	3657	OD1	ASN	99	P	3.84
29.	1299	CZ	TYR	169	A	<-->	3658	ND2	ASN	99	P	3.21
30.	1300	OH	TYR	169	A	<-->	3656	CG	ASN	99	P	3.59
31.	1300	OH	TYR	169	A	<-->	3657	OD1	ASN	99	P	3.32
32.	1300	OH	TYR	169	A	<-->	3658	ND2	ASN	99	P	3.24
33.	1300	OH	TYR	169	A	<-->	3797	CA	HIS	119	P	3.74
34.	1300	OH	TYR	169	A	<-->	3798	C	HIS	119	P	3.83
35.	1300	OH	TYR	169	A	<-->	3802	ND1	HIS	119	P	3.39
36.	1300	OH	TYR	169	A	<-->	3806	N	GLY	120	P	2.97
37.	1300	OH	TYR	169	A	<-->	3807	CA	GLY	120	P	3.76
38.	1304	O	ALA	170	A	<-->	3458	CG2	ILE	73	P	3.85
39.	1320	CD	PRO	172	A	<-->	3455	O	ILE	73	P	3.50
40.	1326	CG	HIS	173	A	<-->	3359	CG	PHE	59	P	3.83
41.	1326	CG	HIS	173	A	<-->	3360	CD1	PHE	59	P	3.76
42.	1327	ND1	HIS	173	A	<-->	3360	CD1	PHE	59	P	3.77
43.	1328	CD2	HIS	173	A	<-->	3358	CB	PHE	59	P	3.47
44.	1328	CD2	HIS	173	A	<-->	3359	CG	PHE	59	P	3.41
45.	1328	CD2	HIS	173	A	<-->	3360	CD1	PHE	59	P	3.63
46.	1329	CE1	HIS	173	A	<-->	3360	CD1	PHE	59	P	3.61
47.	1330	NE2	HIS	173	A	<-->	3358	CB	PHE	59	P	3.53
48.	1330	NE2	HIS	173	A	<-->	3359	CG	PHE	59	P	3.64
49.	1330	NE2	HIS	173	A	<-->	3360	CD1	PHE	59	P	3.53
50.	2203	O	LYS	284	A	<-->	3371	CG2	VAL	60	P	3.80
51.	2211	C	CYS	285	A	<-->	3357	O	PHE	59	P	3.29
52.	2212	O	CYS	285	A	<-->	3357	O	PHE	59	P	3.75
53.	2212	O	CYS	285	A	<-->	3366	CA	VAL	60	P	3.47
54.	2215	N	ASP	286	A	<-->	3357	O	PHE	59	P	2.86
55.	2216	CA	ASP	286	A	<-->	3357	O	PHE	59	P	3.18
56.	2216	CA	ASP	286	A	<-->	3366	CA	VAL	60	P	3.88
57.	2216	CA	ASP	286	A	<-->	3368	O	VAL	60	P	3.30
58.	2217	C	ASP	286	A	<-->	3368	O	VAL	60	P	3.57
59.	2219	CB	ASP	286	A	<-->	3357	O	PHE	59	P	3.78
60.	2220	CG	ASP	286	A	<-->	3368	O	VAL	60	P	3.73
61.	2220	CG	ASP	286	A	<-->	3598	CE	LYS	90	P	3.46
62.	2220	CG	ASP	286	A	<-->	3599	NZ	LYS	90	P	3.47
63.	2221	OD1	ASP	286	A	<-->	3368	O	VAL	60	P	3.44
64.	2221	OD1	ASP	286	A	<-->	3598	CE	LYS	90	P	3.40
65.	2221	OD1	ASP	286	A	<-->	3599	NZ	LYS	90	P	3.24
66.	2222	OD2	ASP	286	A	<-->	3440	CA	SER	71	P	3.75
67.	2222	OD2	ASP	286	A	<-->	3443	CB	SER	71	P	3.72
68.	2222	OD2	ASP	286	A	<-->	3444	OG	SER	71	P	3.38

69.	2222	OD2	ASP	286	A	<-->	3598	CE	LYS	90	P	3.17
70.	2222	OD2	ASP	286	A	<-->	3599	NZ	LYS	90	P	3.02
71.	2223	N	VAL	287	A	<-->	3367	C	VAL	60	P	3.79
72.	2223	N	VAL	287	A	<-->	3368	O	VAL	60	P	2.82
73.	2224	CA	VAL	287	A	<-->	3368	O	VAL	60	P	3.76
74.	2227	CB	VAL	287	A	<-->	3368	O	VAL	60	P	3.45
75.	2229	CG2	VAL	287	A	<-->	3367	C	VAL	60	P	3.83
76.	2229	CG2	VAL	287	A	<-->	3368	O	VAL	60	P	3.22
77.	2229	CG2	VAL	287	A	<-->	3370	CG1	VAL	60	P	3.61
78.	2237	OD2	ASP	288	A	<-->	3431	CE	LYS	69	P	3.86
79.	2237	OD2	ASP	288	A	<-->	3598	CE	LYS	90	P	3.58
80.	2237	OD2	ASP	288	A	<-->	3599	NZ	LYS	90	P	3.29
81.	2256	NH2	ARG	290	A	<-->	3370	CG1	VAL	60	P	3.08
82.	2749	CG	MET	355	A	<-->	3803	CD2	HIS	119	P	3.89
83.	2805	OE1	GLU	361	A	<-->	3845	CE	LYS	125	P	3.68
84.	2805	OE1	GLU	361	A	<-->	3846	NZ	LYS	125	P	3.22
85.	2835	OE2	GLU	364	A	<-->	3845	CE	LYS	125	P	3.25
86.	2835	OE2	GLU	364	A	<-->	3882	OE2	GLU	129	P	3.12
87.	2877	O	HIS	371	A	<-->	3468	CZ	ARG	74	P	3.40
88.	2877	O	HIS	371	A	<-->	3469	NH1	ARG	74	P	3.18
89.	2877	O	HIS	371	A	<-->	3470	NH2	ARG	74	P	2.82
90.	2885	CA	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.30
91.	2887	O	ARG	372	A	<-->	3811	CA	GLY	121	P	3.38
92.	2887	O	ARG	372	A	<-->	3834	CB	ASN	124	P	3.44
93.	2887	O	ARG	372	A	<-->	3835	CG	ASN	124	P	3.73
94.	2887	O	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.83
95.	2888	CB	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.33
96.	2889	CG	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.77
97.	2890	CD	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.33
98.	2890	CD	ARG	372	A	<-->	3867	CG	TYR	128	P	3.66
99.	2890	CD	ARG	372	A	<-->	3869	CD2	TYR	128	P	3.75
100.	2891	NE	ARG	372	A	<-->	3835	CG	ASN	124	P	3.71
101.	2891	NE	ARG	372	A	<-->	3836	OD1	ASN	124	P	2.59
102.	2892	CZ	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.48
103.	2894	NH2	ARG	372	A	<-->	3546	CB	THR	84	P	3.63
104.	2894	NH2	ARG	372	A	<-->	3547	OG1	THR	84	P	2.82
105.	2894	NH2	ARG	372	A	<-->	3548	CG2	THR	84	P	3.39
106.	2894	NH2	ARG	372	A	<-->	3836	OD1	ASN	124	P	3.55
107.	2894	NH2	ARG	372	A	<-->	3837	ND2	ASN	124	P	3.83
108.	2896	CA	LYS	373	A	<-->	3811	CA	GLY	121	P	3.69
109.	2898	O	LYS	373	A	<-->	3803	CD2	HIS	119	P	3.84
110.	2898	O	LYS	373	A	<-->	3805	NE2	HIS	119	P	3.37
111.	2898	O	LYS	373	A	<-->	3811	CA	GLY	121	P	3.63
112.	2912	C	PHE	375	A	<-->	3564	OD2	ASP	86	P	3.65
113.	2913	O	PHE	375	A	<-->	3562	CG	ASP	86	P	3.73
114.	2913	O	PHE	375	A	<-->	3564	OD2	ASP	86	P	3.31
115.	2914	CB	PHE	375	A	<-->	3467	NE	ARG	74	P	3.75
116.	2914	CB	PHE	375	A	<-->	3468	CZ	ARG	74	P	3.89
117.	2921	OXT	PHE	375	A	<-->	3564	OD2	ASP	86	P	3.36
118.	2921	OXT	PHE	375	A	<-->	3808	C	GLY	120	P	3.33
119.	2921	OXT	PHE	375	A	<-->	3809	O	GLY	120	P	3.46

120.	2921	OXT	PHE	375	A	<-->	3810	N	GLY	121	P	3.36
121.	2921	OXT	PHE	375	A	<-->	3811	CA	GLY	121	P	3.64

Salt bridges

<---- A T O M 1 ----->								<---- A T O M 2 ----->							
	Atom	Atom	Res	Res		Atom	Atom	Res	Res		Chain	Distance			
	no.	name	name	no.	Chain	no.	name	name	no.	Chain	Distance				
1.	871	NZ	LYS	113	A	<-->	3529	OE1	GLU	82	P	3.07			
2.	1283	OE1	GLU	167	A	<-->	3583	NH2	ARG	88	P	3.65			
3.	2222	OD2	ASP	286	A	<-->	3599	NZ	LYS	90	P	3.02			
4.	2237	OD2	ASP	288	A	<-->	3599	NZ	LYS	90	P	3.29			
5.	2805	OE1	GLU	361	A	<-->	3846	NZ	LYS	125	P	3.22			

Number of salt bridges: 5

Number of hydrogen bonds: 12

PDBParam gives the following for the binding site interactions

PDBparam: Online Resource for Structural Parameters of Proteins



[Home](#)
[Compute](#)
[Features](#)
[Links](#)
[Tutorial](#)
[Contact](#)

Compute

PDBparam server computes different parameters from the three dimensional structure of the protein. To calculate the properties, mark the checkboxes and enter the PDB code below.

The features are classified into four categories namely, [Inter-residue interactions](#), [Propensities](#), [Physicochemical properties](#), [Identification of binding sites](#).

The results are shown residue-wise or protein-wise whichever is applicable or both.

Input details

The given input PDB-id: 2BTF

Inter-residue interactions

<input checked="" type="checkbox"/> Short range interactions	<input type="checkbox"/> Medium range interactions	<input type="checkbox"/> Long range interactions
<input type="checkbox"/> Contact order	<input type="checkbox"/> Long range order LRO	<input type="checkbox"/> Total contact distance
<input type="checkbox"/> No. of Contacts (8A, CA atoms)		<input type="checkbox"/> No. of Contacts (14A, CA atoms)
<input type="checkbox"/> No. of Contacts (8A, CB atoms)		<input type="checkbox"/> No. of Contacts (14A, CB atoms)
<input type="checkbox"/> Multiple contact index for 2 state proteins		<input type="checkbox"/> Multiple contact index for 3 state proteins

All

[Example](#) [Submit](#) [Clear](#) [Back](#)

By clicking only the short-range interactions we are able to find all short range contacts

Short range contacts

PDB ID : 2BTF.pdb

Residue name	Residue number	Chain	Contacting residue	Residue number	Chain	Distance
ASP	2	A	ASP	3	A	3.871
ASP	2	A	ASP	4	A	5.745
ASP	3	A	ASP	4	A	3.716
ASP	3	A	ILE	5	A	5.194
ASP	4	A	ILE	5	A	3.585
ASP	4	A	ALA	6	A	6.012
ILE	5	A	ALA	6	A	3.452
ILE	5	A	ALA	7	A	6.341
ALA	6	A	ALA	7	A	3.753
ALA	6	A	LEU	8	A	6.391
ALA	7	A	LEU	8	A	3.801
ALA	7	A	VAL	9	A	6.294
LEU	8	A	VAL	9	A	3.822
LEU	8	A	VAL	10	A	6.300
VAL	9	A	VAL	10	A	3.906
VAL	9	A	ASP	11	A	6.443
VAL	10	A	ASP	11	A	3.771
VAL	10	A	ASN	12	A	6.231
ASP	11	A	ASN	12	A	3.789
ASP	11	A	GLY	13	A	6.233
ASN	12	A	GLY	13	A	3.710
ASN	12	A	SER	14	A	6.802

The pdf file is attached [here](#)

5. Check the availability of the hetero dimer complex in other databases and provide details of the same.

The PROXiMATE: PROtein-protein compleX MutAtion ThErmodynamics Database has the following information about the 2BTF complex.

PDB	Chains	Protein 1	Protein 2	Functional Class	Experimental Technique	Ionic Conditions	Temperature	pH	Pubmed/Location	Dissociation constant (M)	ΔG (kcal/mol)
2BTF	A:P	Bovine actin, cytoplasmic 1	Bovine profilin-1		High Shear Viscometry	0.01 M imidazole, 0.15mM CaCl ₂ , 1.5mM DTE, 1mM ATP	298	7	9788869 (Figure 4, p. 3267)	2.30E-06	-7.69

Pubmed provides literature information about the protein protein interaction.

The structure of crystalline profilin-beta-actin

C E Schutt ¹, J C Myslik, M D Rozycki, N C Goonesekere, U Lindberg

Affiliations + expand

PMID: 8413665 DOI: [10.1038/365810a0](https://doi.org/10.1038/365810a0)

 Full text links

 Cite

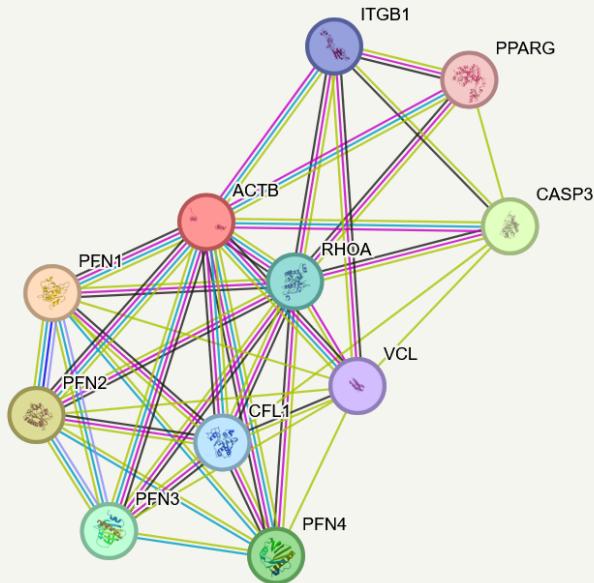
Abstract

The three-dimensional structure of bovine profilin-beta-actin has been solved to 2.55 Å resolution by X-ray crystallography. There are several significant local changes in the structure of beta-actin compared with alpha-actin as well as an overall 5 degrees rotation between its two major domains. Actin molecules in the crystal are organized into ribbons through intermolecular contacts like those found in oligomeric protein assemblies. Profilin forms two extensive contacts with the actin ribbon, one of which appears to correspond to the solution contact *in vitro*.

[PubMed](#) [Disclaimer](#)

6. Construct an interaction network for each protein in the hetero dimer complex using STRING. Write your observations.

The following is the interaction network of the hetero dimer complex observed using string
BETA-ACTIN



[Viewers](#) > [Legend](#) [Settings](#) [Analysis](#) [Exports](#) [Clusters](#) [More](#) [Less](#)

Nodes:

Network nodes represent proteins

splice isoforms or post-translational modifications are collapsed, i.e. each node represents all the proteins produced by a single, protein-coding gene locus.

Node Color



colored nodes:
query proteins and first shell of interactors



white nodes:
second shell of interactors

Node Content



empty nodes:
proteins of unknown 3D structure



filled nodes:
a 3D structure is known or predicted

Observations

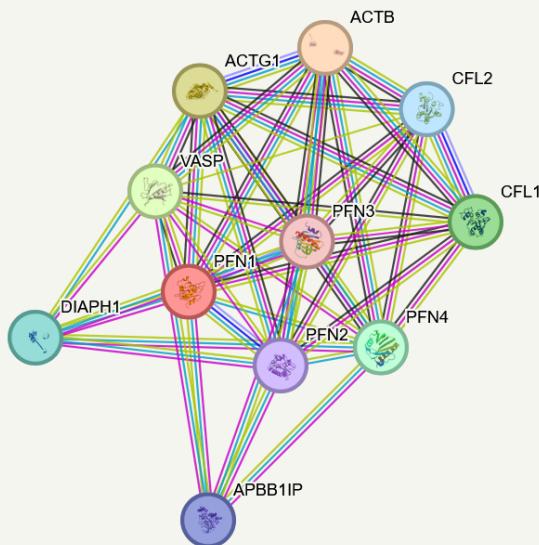
number of nodes:11, number of edges:40

average node degree:7.27, avg. local clustering coefficient:0.876

The following are the predicted functional partners

PFN1	Profilin-1; Binds to actin and affects the structure of the cytoskeleton. At high concentrations, profilin prevents the p...
PFN2	Profilin-2; Binds to actin and affects the structure of the cytoskeleton. At high concentrations, profilin prevents the p...
CASP3	Caspase-3 subunit p12; Involved in the activation cascade of caspases responsible for apoptosis execution. At the o...
PFN4	Profilin-4; Binds to phosphatidylinositol 3-phosphate (PtdIns(3)P), phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5...
PFN3	Profilin-3; Binds to actin and affects the structure of the cytoskeleton. Binds to poly-L-proline, phosphatidylinositol 3-...
RHOA	Transforming protein RhoA; Small GTPase which cycles between an active GTP-bound and an inactive GDP-bound s...
CFL1	Cofilin-1; Binds to F-actin and exhibits pH-sensitive F-actin depolymerizing activity. Regulates actin cytoskeleton dyn...
ITGB1	Integrin beta-1; Integrins alpha-1/beta-1, alpha-2/beta-1, alpha-10/beta-1 and alpha-11/beta-1 are receptors for colla...
VCL	Vinculin.
PPARG	Peroxisome proliferator-activated receptor gamma; Nuclear receptor that binds peroxisome proliferators such as hy...

PROFILIN (P02584)



[Viewers](#) > [Legend](#) [Settings](#) > [Analysis](#) > [Exports](#) > [Clusters](#) > [More](#) [Less](#)

Nodes:

Network nodes represent proteins
splice isoforms or post-translational modifications are collapsed, i.e. each node represents all the proteins produced by a single, protein-coding gene locus.

Node Color
colored nodes:
query proteins and first shell of interactors
white nodes:
second shell of interactors

Node Content
empty nodes:
proteins of unknown 3D structure
filled nodes:
a 3D structure is known or predicted

Edges:**Observations:**

number of nodes: 11, number of edges: 47

average node degree: 8.55, avg. local clustering coefficient: 0.912

The following are the predicted functional partners

● ACTB	Actin, cytoplasmic 1, N-terminally processed; Actin is a highly conserved protein that polymerizes to produce filamentous actin.
● ACTG1	Actin, cytoplasmic 2, N-terminally processed; Actins are highly conserved proteins that are involved in various types of cellular processes.
● VASP	Vasodilator-stimulated phosphoprotein; Ena/VASP proteins are actin-associated proteins involved in a range of processes.
● CFL1	Cofilin-1; Binds to F-actin and exhibits pH-sensitive F-actin depolymerizing activity. Regulates actin cytoskeleton dynamics.
● PFN4	Profilin-4; Binds to phosphatidylinositol 3-phosphate (PtdIns(3)P), phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P2), and poly-L-proline.
● DIAPH1	Uncharacterized protein.
● CFL2	Cofilin-2; Controls reversibly actin polymerization and depolymerization in a pH-sensitive manner. It has the ability to bind to actin monomers.
● APBB1IP	Amyloid beta precursor protein binding family B member 1 interacting protein.
● PFN2	Profilin-2; Binds to actin and affects the structure of the cytoskeleton. At high concentrations, profilin prevents the formation of filaments.
● PFN3	Profilin-3; Binds to actin and affects the structure of the cytoskeleton. Binds to poly-L-proline, phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P2), and phosphatidylinositol 3-phosphate (PtdIns(3)P).

7. Obtain the thermodynamic information for the selected hetero dimer complex.

The thermodynamics information for the selected hetero dimer complex can be found in proximate:

Experimental Technique: High Shear Viscometry

Ionic Conditions: 0.01 M imidazole, 0.15mM CaCl₂, 1.5mM DTE, 1mM ATP

Temperature: 298K

pH: 7

Dissociation constant (M): 2.30E-06

ΔG (kcal/mol) : -7.69

High Shear Viscometry	0.01 M imidazole, 0.15mM CaCl ₂ , 1.5mM DTE, 1mM ATP	298	7	9788869 (Figure 4, p. 3267)	2.30E-06	-7.69
-----------------------	---	-----	---	-----------------------------------	----------	-------

8. Identify the binding sites in the hetero dimer complex using

a. ASA based method

We first input the pdb ID for the entire protein in the GETAREA server to get all the ASA for all the residues using a water molecule of radius 1.4 Angstroms as a probe

Query "3703" for be21b037@smail.iitm.ac.in submitted to webserver.

Residue	Total	Apolar	Backbone	Sidechain	Ratio(%)	In/Out
ASP 2	171.52	59.50	47.26	124.26	100.0	o
ASP 3	80.87	36.26	11.09	69.78	61.8	o
ASP 4	111.93	36.31	15.00	96.92	85.8	o
ILE 5	107.29	101.22	6.07	101.22	68.7	o
ALA 6	20.29	20.29	14.53	5.76	8.9	i
ALA 7	2.40	1.63	0.76	1.63	2.5	i
LEU 8	0.00	0.00	0.00	0.00	0.0	i
VAL 9	0.00	0.00	0.00	0.00	0.0	i
VAL 10	0.42	0.42	0.00	0.42	0.3	i
ASP 11	5.57	0.00	0.00	5.57	4.9	i
ASN 12	9.21	6.70	1.84	7.37	6.4	i
GLY 13	19.00	19.00	19.00	0.00	21.8	
SER 14	12.86	8.57	2.45	10.40	13.4	i
GLY 15	14.56	7.75	14.56	0.00	16.7	i
MET 16	67.45	66.97	0.48	66.97	42.3	

We store this information (excel)

POLAR area/energy	=	8379.36
APOLAR area/energy	=	13090.80
UNKNOWN area/energy	=	0.00

Total area/energy	=	21470.16

Number of surface atoms	=	2007
Number of buried atoms	=	1944
Number of atoms with ASP=0 =		0

We now get the ASA for each of the specific protein molecules that form a complex

Isolating the two proteins in the complex is done using the unique uniprot code provided for each protein and selecting the alpha-fold predicted protein structure and then using them as inputs in GETAREA to find the ASA for the two proteins.

PDB1: BETA-ACTIN :

```
-----  
POLAR area/energy      =      6740.12  
APOLAR area/energy     =      9797.01  
UNKNOW area/energy     =      0.00  
-----  
Total area/energy       =    16537.13  
-----  
Number of surface atoms =      1505  
Number of buried atoms =      1420  
Number of atoms with ASP=0 =      0
```

PDB2: PROFILIN :

```
-----  
POLAR area/energy      =      2565.80  
APOLAR area/energy     =      4695.37  
UNKNOW area/energy     =      0.00  
-----  
Total area/energy       =    7261.17  
-----  
Number of surface atoms =      592  
Number of buried atoms =      460  
Number of atoms with ASP=0 =      0
```

The following are the residues that are responsible for the protein protein interaction ([Sheet](#))

In **BETA-ACTIN**, we have

Residue	no	ASA_Bind	Residue	ASA_free	free-bind
ASP	2	80.87	ASP	146.77	65.9
ASP	3	111.93	ASP	139.5	27.57
ALA	5	20.29	ALA	43.56	23.27
ALA	6	2.4	ALA	6.1	3.7
ASN	11	9.21	ASN	11.18	1.97
MET	15	67.45	MET	73.49	6.04
LYS	17	16.95	LYS	23.86	6.91
PHE	20	11.6	PHE	24.17	12.57
ALA	21	9.27	ALA	12.49	3.22
ASP	23	45.1	ASP	47.09	1.99
ASP	24	128.79	ASP	133.16	4.37
PRO	26	13.45	PRO	15.88	2.43
ARG	27	137.32	ARG	142.31	4.99
VAL	29	45.51	VAL	53.67	8.16
PHE	30	0.51	PHE	1.65	1.14
SER	32	0.02	SER	0.22	0.2
PRO	37	9	PRO	42.32	33.32
GLN	40	138.19	GLN	185.14	46.95
GLY	41	30.76	GLY	39.14	8.38
VAL	42	46.55	VAL	70.69	24.14
GLY	45	52.04	GLY	82.25	30.21
MET	46	69.56	MET	134.48	64.92
GLY	47	53.17	GLY	55.46	2.29
LYS	49	79.22	LYS	111.31	32.09
ASP	50	28.21	ASP	129.67	101.46
VAL	53	7.47	VAL	13.14	5.67
ASP	55	66.32	ASP	83.94	17.62
SER	59	94.11	SER	95.31	1.2
GLY	62	47.13	GLY	64.03	16.9
THR	65	60.59	THR	65.05	4.46
LYS	67	87.32	LYS	105.01	17.69
PRO	69	0.04	PRO	0.35	0.31
GLY	72	14.63	GLY	15.68	1.05

ILE	73	28.04	ILE	32.04	4
ASN	76	51.45	ASN	52.32	0.87
TRP	77	43.78	TRP	44.12	0.34
ASP	78	97.03	ASP	98.98	1.95
MET	80	0	MET	0.22	0.22
HIS	86	12.5	HIS	21.74	9.24
TYR	89	83.94	TYR	89.95	6.01
GLU	91	57.02	GLU	63.66	6.64
ARG	93	187.03	ARG	192.82	5.79
VAL	94	18.72	VAL	29.5	10.78
GLU	98	102.02	GLU	102.38	0.36
HIS	99	7.18	HIS	23.02	15.84
GLU	105	10.58	GLU	11.82	1.24
LEU	108	100.73	LEU	106.03	5.3
ASN	109	16.02	ASN	18.88	2.86
LYS	111	103.91	LYS	142.03	38.12
ARG	114	22.57	ARG	25.81	3.24
GLU	115	45.33	GLU	53.78	8.45
GLN	119	53.03	GLN	53.4	0.37
THR	124	72.82	THR	73.5	0.68
ASN	126	76.34	ASN	79.82	3.48
TYR	131	8.31	TYR	17.03	8.72
ILE	134	18.76	ILE	19.2	0.44
GLY	144	66.75	GLY	73.22	6.47
ARG	145	58.54	ARG	61.99	3.45
THR	146	62.22	THR	69.58	7.36
HIS	159	1.74	HIS	2.57	0.83
ILE	163	0	ILE	0.15	0.15
TYR	164	12.6	TYR	74.76	62.16
GLU	165	54.09	GLU	105.34	51.25
TYR	167	11.12	TYR	150.43	139.31
ALA	168	0	ALA	13.27	13.27
LEU	169	1.91	LEU	31.64	29.73
PRO	170	35.73	PRO	92.67	56.94
HIS	171	47.03	HIS	121.97	74.94
ILE	173	8.44	ILE	12.7	4.26

MET	188	36.84	MET	46.86	10.02
LYS	189	106.32	LYS	121.99	15.67
THR	192	65.41	THR	65.79	0.38
TYR	196	45.2	TYR	59.77	14.57
THR	201	74.51	THR	81.3	6.79
GLU	203	23.06	GLU	50.78	27.72
ILE	206	28.06	ILE	30.98	2.92
ASP	209	25.43	ASP	35.7	10.27
GLU	212	87.22	GLU	105.19	17.97
LYS	213	122.45	LYS	136.81	14.36
LEU	214	21.4	LEU	28.48	7.08
LEU	219	96.99	LEU	97.63	0.64
ALA	226	65.01	ALA	65.35	0.34
ALA	228	11.96	ALA	25.61	13.65
ALA	229	90.42	ALA	95.03	4.61
SER	230	74.83	SER	98.45	23.62
SER	231	44.02	SER	47.87	3.85
LYS	236	91.06	LYS	104.38	13.32
SER	237	68.97	SER	74.37	5.4
GLU	239	113.48	GLU	114.81	1.33
LEU	240	12.36	LEU	20.74	8.38
PRO	241	101.29	PRO	114.55	13.26
ASP	242	105.53	ASP	113.81	8.28
VAL	245	86.24	VAL	92.61	6.37
ILE	246	1.59	ILE	2.07	0.48
THR	247	40.62	THR	47.94	7.32
GLY	249	0.97	GLY	3.38	2.41
GLU	251	3.06	GLU	6.46	3.4
ARG	252	14.94	ARG	22.79	7.85
GLU	257	0	GLU	0.62	0.62
PHE	260	3.88	PHE	5.13	1.25
GLN	261	58.93	GLN	68.2	9.27
PHE	264	59.27	PHE	70.57	11.3
LEU	265	40.27	LEU	53.06	12.79
GLY	266	67.8	GLY	69.33	1.53
GLU	268	176.83	GLU	186.54	9.71

GLU	274	48.65	GLU	60.06	11.41
PHE	277	38.85	PHE	39.25	0.4
LYS	282	90.79	LYS	107.51	16.72
CYS	283	0.06	CYS	2.74	2.68
ASP	284	2.14	ASP	106.76	104.62
VAL	285	72.5	VAL	106.91	34.41
ASP	286	65.9	ASP	127.14	61.24
ILE	287	6.36	ILE	23.24	16.88
LYS	289	110.35	LYS	131.08	20.73
ASP	290	55.36	ASP	57.37	2.01
ALA	293	20.45	ALA	27.38	6.93
ASN	294	23.59	ASN	29.4	5.81
GLY	300	27.15	GLY	41.49	14.34
THR	301	8.87	THR	10.18	1.31
PRO	305	82.26	PRO	93.5	11.24
GLY	306	27.07	GLY	27.31	0.24
ASP	309	80.94	ASP	84.51	3.57
ARG	310	9.36	ARG	11.63	2.27
LYS	313	80.24	LYS	82.8	2.56
GLU	314	27.22	GLU	32.96	5.74
THR	316	72.2	THR	75.79	3.59
ALA	317	82.49	ALA	83.98	1.49
LEU	318	63.81	LEU	70.37	6.56
ALA	319	7.34	ALA	8.95	1.61
THR	322	125.29	THR	128.23	2.94
LYS	324	137.26	LYS	156.03	18.77
ALA	329	29.37	ALA	31.27	1.9
PRO	331	111.14	PRO	115.34	4.2
LYS	334	59.01	LYS	79.39	20.38
SER	342	1.65	SER	2.02	0.37
ILE	343	62.44	ILE	72.46	10.02
SER	346	57.89	SER	61.78	3.89
LEU	347	70.32	LEU	81.74	11.42
THR	349	81.48	THR	114.23	32.75
PHE	350	1.37	PHE	8.82	7.45
GLN	352	131.29	GLN	152.33	21.04

MET	353	27.09	MET	75.53	48.44
SER	356	22.15	SER	36.11	13.96
LYS	357	64.25	LYS	67.58	3.33
GLU	359	27.89	GLU	47.9	20.01
GLU	362	107.04	GLU	129.18	22.14
GLY	364	18.65	GLY	23.21	4.56
PRO	365	37.7	PRO	43.61	5.91
ILE	367	1.95	ILE	10.63	8.68
HIS	369	11.65	HIS	41.38	29.73
ARG	370	13.64	ARG	146.25	132.61
LYS	371	14.3	LYS	51.01	36.71
PHE	373	15.29	PHE	165.41	150.12

In PROFILIN we have

Residue	no	ASA_Bind	Residue	ASA_free	free-bind
ASN	4	60.5	ASN	95.22	34.72
TYR	6	49.46	TYR	57.85	8.39
ILE	7	8.07	ILE	8.62	0.55
ASP	8	91.95	ASP	95.25	3.3
ASN	9	80.65	ASN	84.67	4.02
ALA	12	86.29	ALA	88.06	1.77
THR	15	20.59	THR	29.15	8.56
GLY	23	1.85	GLY	2.53	0.68
TYR	24	35.84	TYR	36.6	0.76
LYS	25	96.01	LYS	107.26	11.25
ASP	26	144.47	ASP	149.31	4.84
SER	27	86.57	SER	88.64	2.07
PRO	28	28.59	PRO	34.27	5.68
VAL	34	12.92	VAL	21.73	8.81
PRO	35	111.28	PRO	119.19	7.91
GLY	36	73.88	GLY	79.27	5.39
LYS	37	88.46	LYS	90.73	2.27
VAL	40	61.07	VAL	63.85	2.78
ASN	41	82.94	ASN	85.85	2.91
THR	43	57.29	THR	62.68	5.39

PRO	44	79.55	PRO	83.74	4.19
ALA	45	74.25	ALA	75.47	1.22
GLY	48	36.11	GLY	36.32	0.21
LEU	50	0.02	LEU	0.13	0.11
GLY	52	18.94	GLY	23.13	4.19
LYS	53	196.48	LYS	203.47	6.99
ARG	55	43.47	ARG	55.31	11.84
SER	56	69.71	SER	76.19	6.48
SER	57	54.11	SER	54.29	0.18
PHE	59	28.22	PHE	107.73	79.51
VAL	60	51.62	VAL	136.39	84.77
ASN	61	105.96	ASN	108.69	2.73
GLY	66	7.34	GLY	18.25	10.91
GLN	68	47.79	GLN	54.87	7.08
LYS	69	82.57	LYS	132.03	49.46
CYS	70	0	CYS	0.15	0.15
SER	71	2.38	SER	41.38	39
VAL	72	2.6	VAL	13.72	11.12
ILE	73	0.83	ILE	93.08	92.25
ARG	74	44.08	ARG	138.66	94.58
ASP	75	24.91	ASP	27.25	2.34
LEU	78	71.68	LEU	77.36	5.68
GLU	82	21.69	GLU	85.07	63.38
PHE	83	33.33	PHE	42.53	9.2
THR	84	1.52	THR	16.37	14.85
MET	85	0	MET	0.28	0.28
ASP	86	1.64	ASP	8.46	6.82
LEU	87	0.03	LEU	0.14	0.11
ARG	88	31.87	ARG	109.55	77.68
LYS	90	22.29	LYS	83.99	61.7
THR	92	116.2	THR	131.21	15.01
GLY	94	74.71	GLY	80.54	5.83
THR	97	12.46	THR	41.82	29.36
PHE	98	36.3	PHE	47.76	11.46
ASN	99	2.62	ASN	22.69	20.07
ILE	100	0.35	ILE	0.85	0.5

THR	105	1.57	THR	4.03	2.46
LYS	107	76.02	LYS	84.37	8.35
MET	113	6.98	MET	7.14	0.16
GLU	116	107.17	GLU	114.9	7.73
GLY	117	73.85	GLY	81.12	7.27
HIS	119	28.09	HIS	140.65	112.56
GLY	120	1.12	GLY	16.35	15.23
GLY	121	0	GLY	36.98	36.98
MET	122	69.13	MET	110.93	41.8
ASN	124	0	ASN	22.68	22.68
LYS	125	25.22	LYS	132.24	107.02
LYS	126	30.64	LYS	37.34	6.7
TYR	128	47.17	TYR	82.88	35.71
GLU	129	72.16	GLU	90.42	18.26
ARG	136	165.34	ARG	171.15	5.81
SER	137	51.42	SER	52	0.58
GLN	138	134.08	GLN	135.59	1.51

b. Distance based method

We use PDBparam to find all binding sites using the identify binding site option, which sets the interaction distance cutoff at 3.5 Angstroms.

[Home](#)[Compute](#)[Features](#)[Links](#)[Tutorial](#)[Contact](#)

Compute

PDBparam server computes different parameters from the three dimensional structure of the protein. To calculate the properties, mark the checkboxes and enter the PDB code below.

The features are classified into four categories namely, **Inter-residue interactions**, **Propensities**, **Physicochemical properties**, **Identification of binding sites**.

To calculate all the properties in a single page, click [here](#)

The results are shown residue-wise or protein-wise whichever is applicable or both.

Please enter the following

Enter a PDB-id

Ex: PDB-id=2TRX, Chain=A

Chain name (optional)

Select one of the following

- Identification of binding site
- Inter-residue interactions
- Secondary structure propensities
- Physicochemical properties

[Home](#)[Compute](#)[Features](#)[Links](#)[Tutorial](#)[Contact](#)

Compute

PDBparam server computes different parameters from the three dimensional structure of the protein. To calculate the properties, mark the checkboxes and enter the PDB code below.

The features are classified into four categories namely, **Inter-residue interactions**, **Propensities**, **Physicochemical properties**, **Identification of binding sites**.

The results are shown residue-wise or protein-wise whichever is applicable or both.

Input details

The given input PDB-id: 2BTF

Identification of binding site

Protein-Protein

Protein-Ligand

Protein-DNA/RNA

Chain Name (optional) Distance threshold

The output is all the residues that are identified as binding sites

PDB ID : 2BTF.pdb

Residue name	Residue number	Atom name	Chain name	Residue name	Residue number	Atom name	Chain name	Distance
LYS	113	CE	A	GLU	82	CD	P	3.426
LYS	113	CE	A	GLU	82	OE2	P	3.091
LYS	113	NZ	A	GLU	82	CD	P	3.489
LYS	113	NZ	A	GLU	82	OE1	P	3.071
LYS	113	NZ	A	GLU	82	OE2	P	3.418
TYR	133	OH	A	HIS	119	NE2	P	3.394
TYR	166	CD2	A	SER	71	OG	P	3.488
TYR	166	CE2	A	SER	71	OG	P	3.206
TYR	166	CE2	A	ARG	88	CG	P	3.498
TYR	166	OH	A	THR	89	O	P	2.990
TYR	169	CE2	A	ASN	99	ND2	P	3.429
TYR	169	CZ	A	ASN	99	ND2	P	3.206
TYR	169	OH	A	ASN	99	OD1	P	3.315
TYR	169	OH	A	ASN	99	ND2	P	3.237
TYR	169	OH	A	HIS	119	ND1	P	3.386
TYR	169	OH	A	GLY	120	N	P	2.973
HIS	173	CD2	A	PHE	59	CB	P	3.470
HIS	173	CD2	A	PHE	59	CG	P	3.410
CYS	285	C	A	PHE	59	O	P	3.287
CYS	285	O	A	VAL	60	CA	P	3.468
ASP	286	N	A	PHE	59	O	P	2.861
ASP	286	CA	A	PHE	59	O	P	3.177
ASP	286	CA	A	VAL	60	O	P	3.298
ASP	286	CG	A	LYS	90	CE	P	3.463
ASP	286	CG	A	LYS	90	NZ	P	3.468
ASP	286	OD1	A	VAL	60	O	P	3.438
ASP	286	OD1	A	LYS	90	CE	P	3.404
ASP	286	OD1	A	LYS	90	NZ	P	3.239
ASP	286	OD2	A	SER	71	OG	P	3.376
ASP	286	OD2	A	LYS	90	CE	P	3.173
ASP	286	OD2	A	LYS	90	NZ	P	3.016
VAL	287	N	A	VAL	60	O	P	2.816
VAL	287	CB	A	VAL	60	O	P	3.445
VAL	287	CG2	A	VAL	60	O	P	3.223
ASP	288	OD2	A	LYS	90	NZ	P	3.286
ARG	290	NH2	A	VAL	60	CG1	P	3.081
GLU	361	OE1	A	LYS	125	NZ	P	3.222
GLU	364	OE2	A	LYS	125	CE	P	3.249
GLU	364	OE2	A	GLU	129	OE2	P	3.117

HIS	371	O	A	ARG	74	CZ	P	3.404
HIS	371	O	A	ARG	74	NH1	P	3.184
HIS	371	O	A	ARG	74	NH2	P	2.816
ARG	372	CA	A	ASN	124	OD1	P	3.300
ARG	372	O	A	GLY	121	CA	P	3.377
ARG	372	O	A	ASN	124	CB	P	3.437
ARG	372	CB	A	ASN	124	OD1	P	3.334
ARG	372	CD	A	ASN	124	OD1	P	3.326
ARG	372	NE	A	ASN	124	OD1	P	2.587
ARG	372	CZ	A	ASN	124	OD1	P	3.483
ARG	372	NH2	A	THR	84	OG1	P	2.821
ARG	372	NH2	A	THR	84	CG2	P	3.393
LYS	373	O	A	HIS	119	NE2	P	3.366
PHE	375	O	A	ASP	86	OD2	P	3.305
PHE	375	OXT	A	ASP	86	OD2	P	3.357
PHE	375	OXT	A	GLY	120	C	P	3.333
PHE	375	OXT	A	GLY	120	O	P	3.461
PHE	375	OXT	A	GLY	121	N	P	3.360
PHE	59	O	P	CYS	285	C	A	3.287
PHE	59	O	P	ASP	286	N	A	2.861
PHE	59	O	P	ASP	286	CA	A	3.177
PHE	59	CB	P	HIS	173	CD2	A	3.470
PHE	59	CG	P	HIS	173	CD2	A	3.410
VAL	60	CA	P	CYS	285	O	A	3.468
VAL	60	O	P	ASP	286	CA	A	3.298
VAL	60	O	P	ASP	286	OD1	A	3.438
VAL	60	O	P	VAL	287	N	A	2.816
VAL	60	O	P	VAL	287	CB	A	3.445
VAL	60	O	P	VAL	287	CG2	A	3.223
VAL	60	CG1	P	ARG	290	NH2	A	3.081
SER	71	OG	P	TYR	166	CD2	A	3.488
SER	71	OG	P	TYR	166	CE2	A	3.206
SER	71	OG	P	ASP	286	OD2	A	3.376
ARG	74	CZ	P	HIS	371	O	A	3.404
ARG	74	NH1	P	HIS	371	O	A	3.184
ARG	74	NH2	P	HIS	371	O	A	2.816
GLU	82	CD	P	LYS	113	CE	A	3.426
GLU	82	CD	P	LYS	113	NZ	A	3.489
GLU	82	OE1	P	LYS	113	NZ	A	3.071
GLU	82	OE2	P	LYS	113	CE	A	3.091
GLU	82	OE2	P	LYS	113	NZ	A	3.418
THR	84	OG1	P	ARG	372	NH2	A	2.821
THR	84	CG2	P	ARG	372	NH2	A	3.393

ASP	86	OD2	P	PHE	375	O	A	3.305
ASP	86	OD2	P	PHE	375	OXT	A	3.357
ARG	88	CG	P	TYR	166	CE2	A	3.498
THR	89	O	P	TYR	166	OH	A	2.990
LYS	90	CE	P	ASP	286	CG	A	3.463
LYS	90	CE	P	ASP	286	OD1	A	3.404
LYS	90	CE	P	ASP	286	OD2	A	3.173
LYS	90	NZ	P	ASP	286	CG	A	3.468
LYS	90	NZ	P	ASP	286	OD1	A	3.239
LYS	90	NZ	P	ASP	286	OD2	A	3.016
LYS	90	NZ	P	ASP	288	OD2	A	3.286
ASN	99	OD1	P	TYR	169	OH	A	3.315
ASN	99	ND2	P	TYR	169	CE2	A	3.429
ASN	99	ND2	P	TYR	169	CZ	A	3.206
ASN	99	ND2	P	TYR	169	OH	A	3.237
HIS	119	ND1	P	TYR	169	OH	A	3.386
HIS	119	NE2	P	TYR	133	OH	A	3.394
HIS	119	NE2	P	LYS	373	O	A	3.366
GLY	120	N	P	TYR	169	OH	A	2.973
GLY	120	C	P	PHE	375	OXT	A	3.333
GLY	120	O	P	PHE	375	OXT	A	3.461
GLY	121	N	P	PHE	375	OXT	A	3.360
GLY	121	CA	P	ARG	372	O	A	3.377
ASN	124	CB	P	ARG	372	O	A	3.437
ASN	124	OD1	P	ARG	372	CA	A	3.300
ASN	124	OD1	P	ARG	372	CB	A	3.334
ASN	124	OD1	P	ARG	372	CD	A	3.326
ASN	124	OD1	P	ARG	372	NE	A	2.587
ASN	124	OD1	P	ARG	372	CZ	A	3.483
LYS	125	CE	P	GLU	364	OE2	A	3.249
LYS	125	NZ	P	GLU	361	OE1	A	3.222
GLU	129	OE2	P	GLU	364	OE2	A	3.117

9. Calculate the propensity of the binding site residues (use results from question 8)

We calculate the propensity using the binding sites identified using the ASA method.

A = BETA-ACTIN

B = PROFILIN

Count = number of residues in the identified binding sites

Count_mol = number of residues in the protein molecule

No	Residue	Count_A	Count_B	Count_mol	PropensityA	PropensityB
1	ALA	11	2	40	0.28	0.05
2	ARG	7	4	23	0.30	0.17
3	ASN	5	6	15	0.33	0.40
4	ASP	13	4	31	0.42	0.13
5	CYS	1	1	9	0.11	0.11
6	GLU	15	3	30	0.50	0.10
7	GLN	4	2	16	0.25	0.13
8	GLY	11	9	44	0.25	0.20
9	HIS	5	1	10	0.50	0.10
10	ILE	9	3	35	0.26	0.09
11	LEU	8	3	38	0.21	0.08
12	LYS	14	8	28	0.50	0.29
13	MET	5	3	22	0.23	0.14
14	PHE	7	3	18	0.39	0.17
15	PRO	8	3	23	0.35	0.13
16	SER	8	5	34	0.24	0.15
17	THR	10	6	38	0.26	0.16
18	TRP	1	0	6	0.17	0.00
19	TYR	5	3	19	0.26	0.16
20	VAL	6	4	33	0.18	0.12

10. Identify the top 5 interacting pairs in the hetero dimer complex.

To do this we use PDB sum to find all interactions in the complex and use python to find the top 5 interacting pairs.

The csv file containing all interactions can be found [here](#).

The code to find the top 5 interactions can be found [here](#).

The top 5 interactions are

Residue 1	Residue 2	Number of interactions
TYR1169	ASN99	8
VAL2287	VAL60	7
ARG3372	ASN124	6
HIS1173	PHE59	6
LYS113	GLU82	6

11. Calculate the following for the hetero dimer complex

a. Binding free energy

Dissociation constant (M): 2.30E-06

ΔG (kcal/mol) : -7.69 [Binding Free Energy]

T = 298 K

as mentioned in the proximate database, we can calculate the binding energy from the dissociation constant as $\Delta G = RT \ln(k_d)$

Which gives us

$$\Delta G = 8.314 \text{ J/mol.K} \cdot 298\text{K} \cdot \ln(2.30\text{E-06}) \cdot 0.000239006(\text{cal/J}) = -7.6875 \text{ kcal/mol}$$

b. Change in binding free energy upon mutation

The proximate database provides 4 known mutations for the PDB : 2BTF, and their corresponding dissociation constants and the binding free energy

SEARCH RESULTS

You have searched for: **PDB=2BTF**

Found **4 entries** matching your search criteria. Click [here](#) to download all results.

Entry	PDB	Mutation(s)	Protein 1	Protein 2	Secondary Structure	Relative Accessibility	Wild-type K_D (M)	Mutant K_D (M)	Wild-type ΔG (kcal/mol)	$\Delta \Delta G$ (kcal/mol)	Pubmed/Reference	Database Source
5880	2BTF	P:F59A	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:F59A=Alpha-helix	P:F59A=0.15	2.30E-06	3.40E-05	-7.69	1.6	9788869 (Figure 4, p. 3267)	SKEMPI;ASEdb;
5881	2BTF	P:G120F	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:G120F=Alpha-helix	P:G120F=0.00	2.30E-06	4.20E-05	-7.69	1.72	9788869 (Figure 4, p. 3267)	SKEMPI
5882	2BTF	P:K125A	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:K125A=Alpha-helix	P:K125A=0.11	2.30E-06	5.00E-06	-7.69	0.46	9788869 (Figure 4, p. 3267)	SKEMPI;ASEdb;
5883	2BTF	P:V60E	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:V60E=Alpha-helix	P:V60E=0.33	2.30E-06	1.20E-05	-7.69	0.98	9788869 (Figure 4, p. 3267)	SKEMPI

Entry	PDB	Mutation(s)	Protein 1	Protein 2	Secondary Structure
5880	2BTF	P:F59A	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:F59A=Alpha-helix
5881	2BTF	P:G120F	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:G120F=Alpha-helix
5882	2BTF	P:K125A	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:K125A=Alpha-helix
5883	2BTF	P:V60E	Bovine actin, cytoplasmic 1 P60712	Bovine profilin-1 P02584	P:V60E=Alpha-helix

Wild-type K _D (M)	Mutant K _D (M)	Wild-type ΔG (kcal/mol)	ΔΔG (kcal/mol)	Pubmed/Reference	Database Source
2.30E-06	3.40E-05	-7.69	1.6	9788869 (Figure 4, p. 3267)	SKEMPI; ASEdb;
2.30E-06	4.20E-05	-7.69	1.72	9788869 (Figure 4, p. 3267)	SKEMPI
2.30E-06	5.00E-06	-7.69	0.46	9788869 (Figure 4, p. 3267)	SKEMPI; ASEdb;
2.30E-06	1.20E-05	-7.69	0.98	9788869 (Figure 4, p. 3267)	SKEMPI

The change in free energy is in the ΔΔG column.

12. Predict the structure for the given protein-protein complex. Write your observations.

>Subunit1

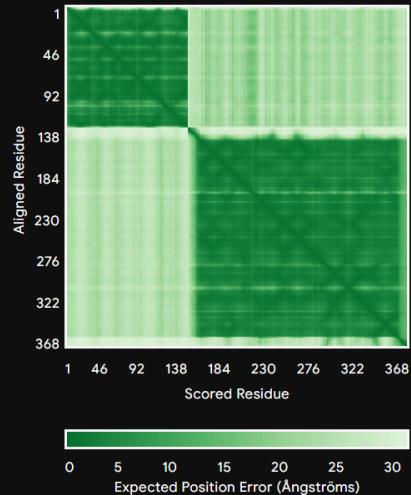
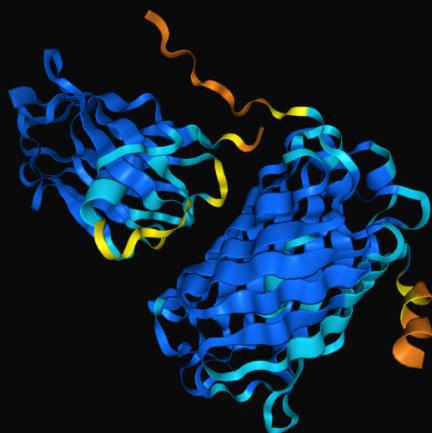
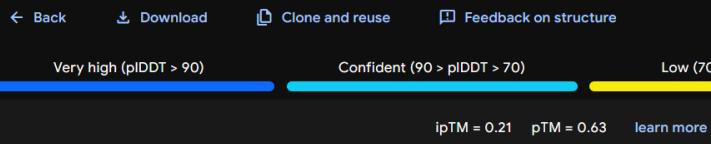
SGEVQLLESGGGLVQAGGSLRLSCAASDRTFSSYAVGWFRQAPGREREFVAATSWRGDSTYYADSVKGR
FTISRDNAKNTVHLQMNSLKPEDTAVYYCAAKWGPVPTNSVMNYYTREYDYWGQQGTQVTVSS

>Subunit2

SGMVKGEEDNMASLPATHELHIFGSINGVDFDMVGQGTGNPNDGYEELNLKSTKGDLQFSPWILVPHI
GYGFHQYLPPDGMSPFQAAMVDGSGYQVHRTMQFEDGASLTNVYRYTYEGSHIKGEAQVKGTGFPA
DGPVMTNSLTAADWCRSKKTYPNDKTIISTFKWSYTTGNGKRYRSTARTTYTFAKPMAANYLKNQPMYVF
RKTELKHSKTELFKEWQKAFTDVMGMDELYK

Using alpha fold3 protein structure prediction online tool, we get the following protein protein complex

2024-09-04_12:18



Information

Type	Copies	Sequence
Protein	1	<pre> 10 20 30 40 50 60 SGEVQLLES GGLVQAGGSL RLSCAADRT FSSYAVGWFR QAPGREREFV AATSWRGDST 70 80 90 100 110 120 YYADSVKG TISRDNAKNT VHLQMNSLKP EDTAVYYCAA KWGPVPTNSV MNYYTREYDY 130 131 WGQGTQVTVS S </pre>
Protein	1	<pre> 10 20 30 40 50 60 SGMVSKGEE NMASLPATHE LHIFGSINGV DFDMVGGTG NPNDGYEELN LKSTKGDLQF 70 80 90 100 110 120 SPWILVPHIG YGFHQYLPPY DGMSPFQAAAM VDGSGYQVHR TMQFEDGASL TVNYRYTYEG 130 140 150 160 170 180 SHIKGEAQVK GTGFPAADGPV MTNSLTAAADW CRSKKTPND KTIISTFKWS YTTGNGKRYR 190 200 210 220 230 238 STARTTYTFA KPMAANYLKN QPMYVFRKTE LKHSKTELNF KEWQKAFTDV MGMDELYK </pre>

Seed: 69

Observations:

The protein complex has two subunits, both are beta barrel in structure and one is smaller than the other, the structure prediction is high in confidence. The bigger beta subunit has a alpha helix enclosed inside as shown below

