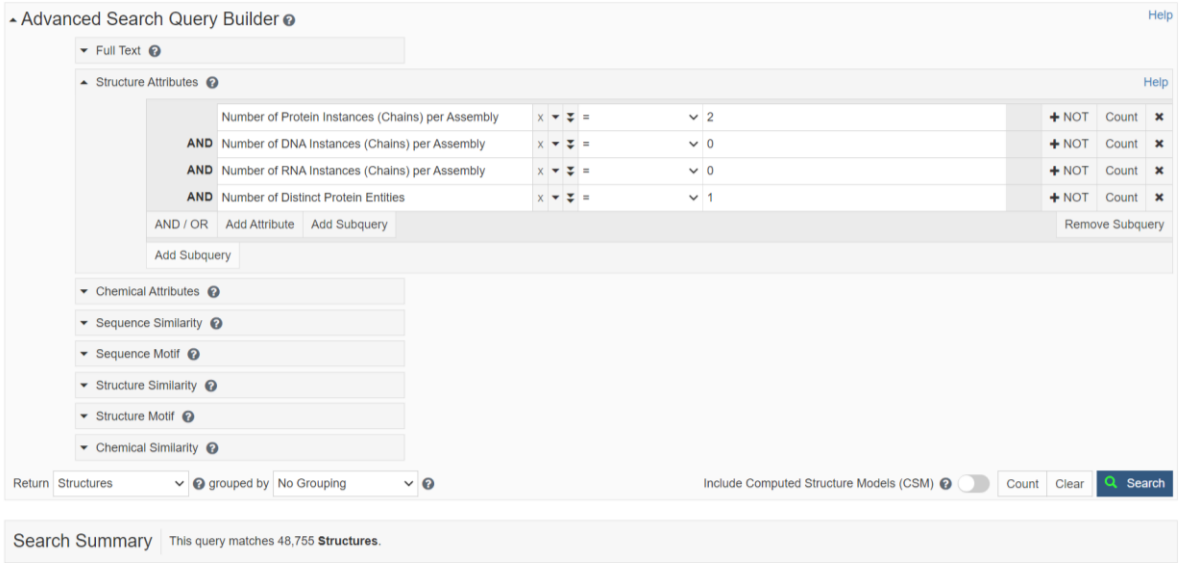
Anirudh Rao
BE21B004

BT6320 – Protein Interactions: Computational Techniques

Assignment 1

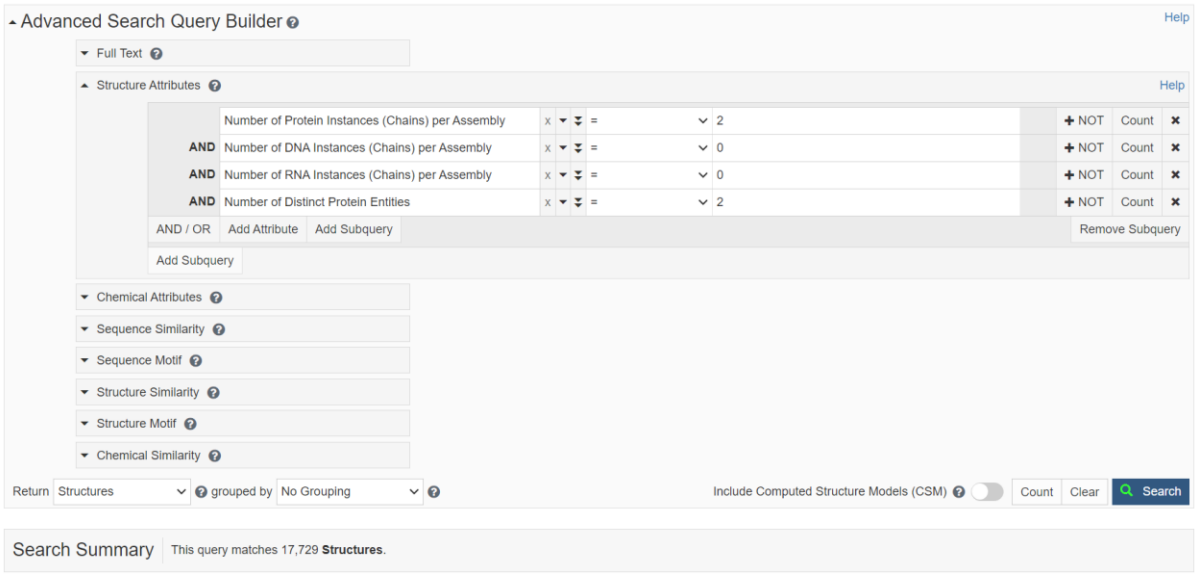
1

The number of homodimers in PDB was found using the following settings with the Advanced Search Query Builder:



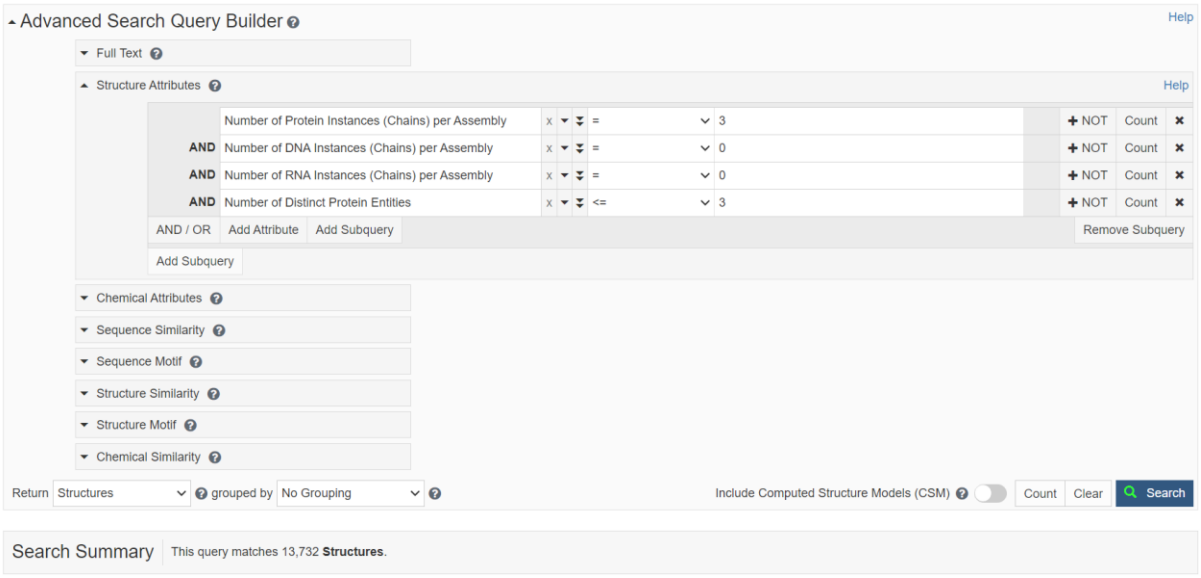
This showed the number of homodimers to be **48,755**.

The number of heterodimers in PDB was found using the following settings with the Advanced Search Query Builder:



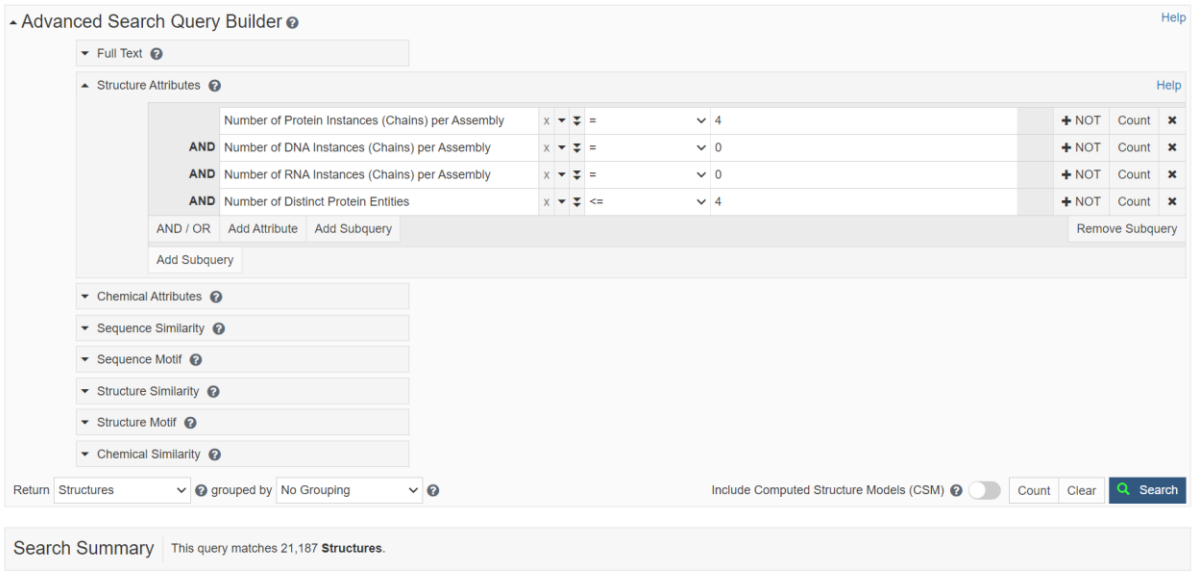
This showed the number of heterodimers to be **17,729**.

The number of trimers in PDB was found using the following settings with the Advanced Search Query Builder:



This showed the number of trimers to be **13,732**.

The number of tetramers in PDB was found using the following settings with the Advanced Search Query Builder:



This showed the number of tetramers to be **21,187**.

For all subsequent questions, supplementary files can be found at this link: https://drive.google.com/drive/folders/18R_eC6mQP9UAx5gOvspCVmHshDm9yEHH?usp=sharing

2

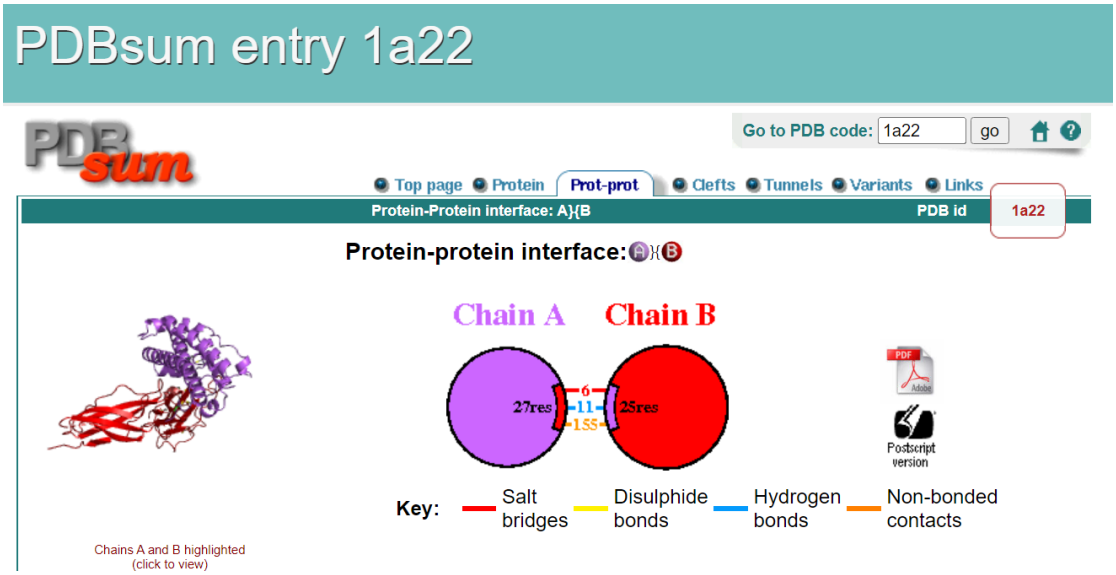
I have chosen PDB ID 1A22 for this assignment.

3

- The structure contains the human growth hormone bound to its receptor.
- It is a heterodimer with stoichiometry A1B1, i.e., it has 2 unique protein chains.
- It is an asymmetric protein complex.
- The structure was determined by x-ray diffraction at a resolution of 2.60 Å.
- The total weight of the structure is 49.66 kDa.
- The binding of the growth hormone to this receptor leads to the activation of growth-related signal transduction pathways in the cell.

4

PDBsum shows that 27 residues of the hormone interact with 25 residues of the receptor. These interactions include 6 salt bridges, 11 hydrogen bonds, and 155 non-bonded contacts.

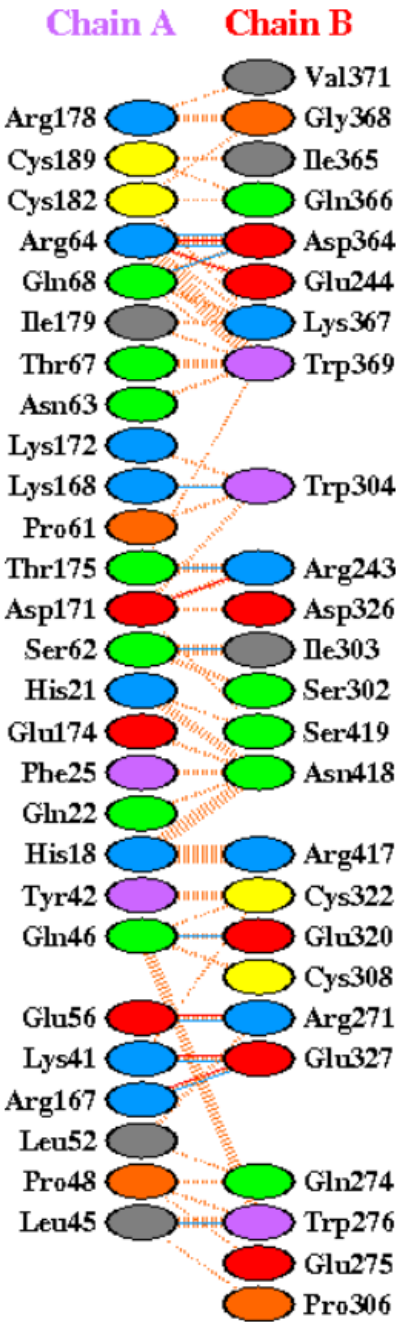


The interface surface area is approximately 1300 Å^2 for both proteins.

Interface statistics

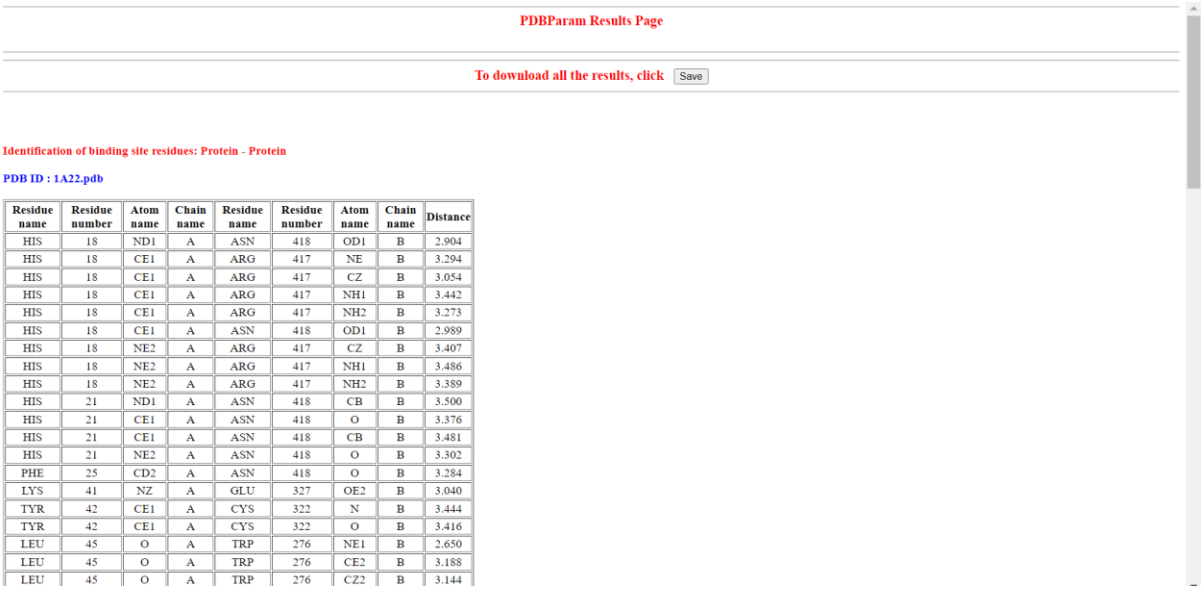
Chain	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A	27	1308	6	-	11	155
B	25	1333				

The residues involved in the interactions are:



The list of interactions is included as a supplementary file.

PDBparam is used to identify the binding site residues with the default distance threshold of 3.5 Å.



PDBParam Results Page

To download all the results, click Save

Identification of binding site residues: Protein - Protein

PDB ID : 1A22.pdb

Residue name	Residue number	Atom name	Chain name	Residue name	Residue number	Atom name	Chain name	Distance
HIS	18	ND1	A	ASN	418	OD1	B	2.904
HIS	18	CE1	A	ARG	417	NE	B	3.294
HIS	18	CE1	A	ARG	417	CZ	B	3.054
HIS	18	CE1	A	ARG	417	NH1	B	3.442
HIS	18	CE1	A	ARG	417	NH2	B	3.273
HIS	18	CE1	A	ASN	418	OD1	B	2.989
HIS	18	NE2	A	ARG	417	CZ	B	3.407
HIS	18	NE2	A	ARG	417	NH1	B	3.486
HIS	18	NE2	A	ARG	417	NH2	B	3.389
HIS	21	ND1	A	ASN	418	CB	B	3.500
HIS	21	CE1	A	ASN	418	O	B	3.376
HIS	21	CE1	A	ASN	418	CB	B	3.481
HIS	21	NE2	A	ASN	418	O	B	3.302
PHE	25	CD2	A	ASN	418	O	B	3.284
LYS	41	NZ	A	GLU	327	OE2	B	3.040
TYR	42	CE1	A	CYS	322	N	B	3.444
TYR	42	CE1	A	CYS	322	O	B	3.416
LEU	45	O	A	TRP	276	NE1	B	2.650
LEU	45	O	A	TRP	276	CE2	B	3.188
LEU	45	O	A	TRP	276	CZ2	B	3.144

Residue name	Residue number	Atom name	Chain name	Residue name	Residue number	Atom name	Chain name	Distance
HIS	18	ND1	A	ASN	418	OD1	B	2.904
HIS	18	CE1	A	ARG	417	NE	B	3.294
HIS	18	CE1	A	ARG	417	CZ	B	3.054
HIS	18	CE1	A	ARG	417	NH1	B	3.442
HIS	18	CE1	A	ARG	417	NH2	B	3.273
HIS	18	CE1	A	ASN	418	OD1	B	2.989
HIS	18	NE2	A	ARG	417	CZ	B	3.407
HIS	18	NE2	A	ARG	417	NH1	B	3.486
HIS	18	NE2	A	ARG	417	NH2	B	3.389
HIS	21	ND1	A	ASN	418	CB	B	3.500
HIS	21	CE1	A	ASN	418	O	B	3.376
HIS	21	CE1	A	ASN	418	CB	B	3.481
HIS	21	NE2	A	ASN	418	O	B	3.302
PHE	25	CD2	A	ASN	418	O	B	3.284
LYS	41	NZ	A	GLU	327	OE2	B	3.040
TYR	42	CE1	A	CYS	322	N	B	3.444
TYR	42	CE1	A	CYS	322	O	B	3.416
LEU	45	O	A	TRP	276	NE1	B	2.650
LEU	45	O	A	TRP	276	CE2	B	3.188
LEU	45	O	A	TRP	276	CZ2	B	3.144

GLN	46	CA	A	TRP	276	NE1	B	3.217
GLN	46	NE2	A	CYS	308	SG	B	3.424
GLN	46	NE2	A	GLU	320	OE2	B	3.047
GLN	46	NE2	A	CYS	322	SG	B	3.491
LEU	52	CD1	A	ARG	271	CZ	B	3.430
LEU	52	CD1	A	ARG	271	NH2	B	3.210
GLU	56	CD	A	ARG	271	NH2	B	3.437
GLU	56	OE2	A	ARG	271	NH2	B	2.883
SER	62	O	A	SER	302	CA	B	3.500
SER	62	O	A	ILE	303	N	B	2.713
ARG	64	N	A	TRP	369	CZ2	B	3.433
ARG	64	NE	A	ASP	364	OD2	B	2.899
ARG	64	CZ	A	ASP	364	OD2	B	3.425
ARG	64	NH2	A	ASP	364	OD2	B	3.269
ARG	64	NH2	A	LYS	367	NZ	B	3.044
GLN	68	NE2	A	ASP	364	OD2	B	2.982
GLN	68	NE2	A	LYS	367	O	B	3.398
ARG	167	NH2	A	GLU	327	OE2	B	3.041
LYS	168	NZ	A	TRP	304	O	B	3.001
ASP	171	OD1	A	ARG	243	NH1	B	3.324
ASP	171	OD2	A	ARG	243	NH1	B	3.187
ASP	171	OD2	A	TRP	304	NE1	B	3.390
GLU	174	OE1	A	ASN	418	OD1	B	3.405
THR	175	OG1	A	ARG	243	NH2	B	2.646
ARG	178	CG	A	GLY	368	O	B	3.218
CYS	182	SG	A	GLY	368	N	B	3.468
ARG	243	NH1	B	ASP	171	OD1	A	3.324
ARG	243	NH1	B	ASP	171	OD2	A	3.187
ARG	243	NH2	B	THR	175	OG1	A	2.646
ARG	271	CZ	B	LEU	52	CD1	A	3.430
ARG	271	NH2	B	LEU	52	CD1	A	3.210
ARG	271	NH2	B	GLU	56	CD	A	3.437
ARG	271	NH2	B	GLU	56	OE2	A	2.883
TRP	276	NE1	B	LEU	45	O	A	2.650
TRP	276	NE1	B	GLN	46	CA	A	3.217
TRP	276	CE2	B	LEU	45	O	A	3.188
TRP	276	CZ2	B	LEU	45	O	A	3.144
SER	302	CA	B	SER	62	O	A	3.500

ILE	303	N	B	SER	62	O	A	2.713
TRP	304	O	B	LYS	168	NZ	A	3.001
TRP	304	NE1	B	ASP	171	OD2	A	3.390
CYS	308	SG	B	GLN	46	NE2	A	3.424
GLU	320	OE2	B	GLN	46	NE2	A	3.047
CYS	322	N	B	TYR	42	CE1	A	3.444
CYS	322	O	B	TYR	42	CE1	A	3.416
CYS	322	SG	B	GLN	46	NE2	A	3.491
GLU	327	OE2	B	LYS	41	NZ	A	3.040
GLU	327	OE2	B	ARG	167	NH2	A	3.041
ASP	364	OD2	B	ARG	64	NE	A	2.899
ASP	364	OD2	B	ARG	64	CZ	A	3.425
ASP	364	OD2	B	ARG	64	NH2	A	3.269
ASP	364	OD2	B	GLN	68	NE2	A	2.982
LYS	367	O	B	GLN	68	NE2	A	3.398
LYS	367	NZ	B	ARG	64	NH2	A	3.044
GLY	368	N	B	CYS	182	SG	A	3.468
GLY	368	O	B	ARG	178	CG	A	3.218
TRP	369	CZ2	B	ARG	64	N	A	3.433
ARG	417	NE	B	HIS	18	CE1	A	3.294
ARG	417	CZ	B	HIS	18	CE1	A	3.054
ARG	417	CZ	B	HIS	18	NE2	A	3.407
ARG	417	NH1	B	HIS	18	CE1	A	3.442
ARG	417	NH1	B	HIS	18	NE2	A	3.486
ARG	417	NH2	B	HIS	18	CE1	A	3.273
ARG	417	NH2	B	HIS	18	NE2	A	3.389
ASN	418	O	B	HIS	21	CE1	A	3.376
ASN	418	O	B	HIS	21	NE2	A	3.302
ASN	418	O	B	PHE	25	CD2	A	3.284
ASN	418	CB	B	HIS	21	ND1	A	3.500
ASN	418	CB	B	HIS	21	CE1	A	3.481
ASN	418	OD1	B	HIS	18	ND1	A	2.904
ASN	418	OD1	B	HIS	18	CE1	A	2.989
ASN	418	OD1	B	GLU	174	OE1	A	3.405

In addition to this, all inter-residue interactions were found using PDBparam.

PDBParam Results Page

To download all the results, click Save

Short range contacts

PDB ID : 1A22.pdb

Residue name	Residue number	Chain	Contacting residue	Residue number	Chain	Distance
PHE	1	A	PRO	2	A	3.842
PHE	1	A	THR	3	A	6.776
PRO	2	A	THR	3	A	3.789
PRO	2	A	ILE	4	A	6.999
THR	3	A	ILE	4	A	3.782
THR	3	A	PRO	5	A	6.357
ILE	4	A	PRO	5	A	3.816
ILE	4	A	LEU	6	A	6.261
PRO	5	A	LEU	6	A	3.829
PRO	5	A	SER	7	A	5.602
LEU	6	A	SER	7	A	3.800
LEU	6	A	ARG	8	A	5.825
SER	7	A	ARG	8	A	3.809
SER	7	A	LEU	9	A	5.349
ARG	8	A	LEU	9	A	3.776
ARG	8	A	PHE	10	A	5.565
LEU	9	A	PHE	10	A	3.795
LEU	9	A	ASP	11	A	5.286
PHE	10	A	ASP	11	A	3.769
PHE	10	A	ASN	12	A	5.433
ASP	11	A	ASN	12	A	3.794
ASP	11	A	ALA	13	A	5.612
ASN	12	A	ALA	13	A	3.801
ASN	12	A	MET	14	A	5.339

Medium range contacts

PDB ID : 1A22.pdb

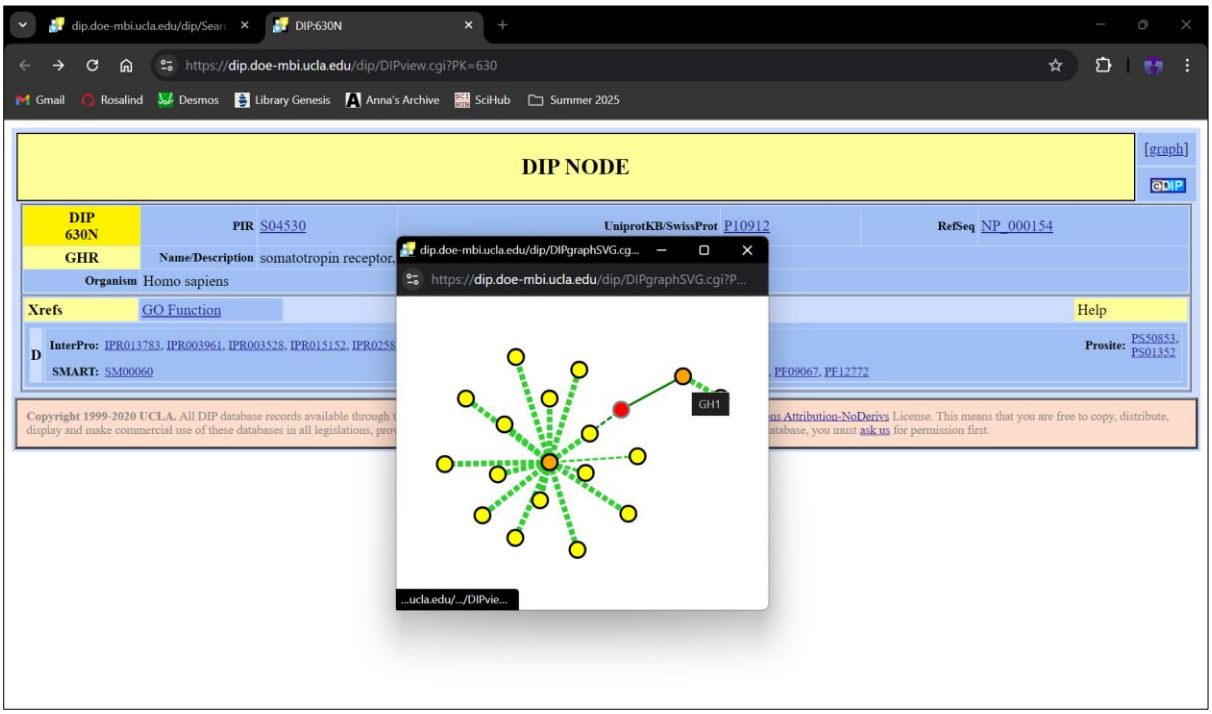
Residue name	Residue number	Chain	Contacting residue	Residue number	Chain	Distance
PRO	5	A	ARG	8	A	7.438
LEU	6	A	LEU	9	A	5.501
LEU	6	A	PHE	10	A	6.071
SER	7	A	PHE	10	A	4.970
SER	7	A	ASP	11	A	5.722
ARG	8	A	ASP	11	A	4.998
ARG	8	A	ASN	12	A	6.329
LEU	9	A	ASN	12	A	5.230
LEU	9	A	ALA	13	A	6.226
PHE	10	A	ALA	13	A	5.089
PHE	10	A	MET	14	A	5.942
ASP	11	A	MET	14	A	5.233
ASP	11	A	LEU	15	A	6.356
ASN	12	A	LEU	15	A	5.224
ASN	12	A	ARG	16	A	6.551
ALA	13	A	ARG	16	A	5.447
ALA	13	A	ALA	17	A	6.293
MET	14	A	ALA	17	A	5.105
MET	14	A	HIS	18	A	6.187
LEU	15	A	HIS	18	A	5.293
LEU	15	A	ARG	19	A	6.132
ARG	16	A	ARG	19	A	4.872
ARG	16	A	LEU	20	A	6.063
ALA	17	A	LEU	20	A	5.145

Long range contacts

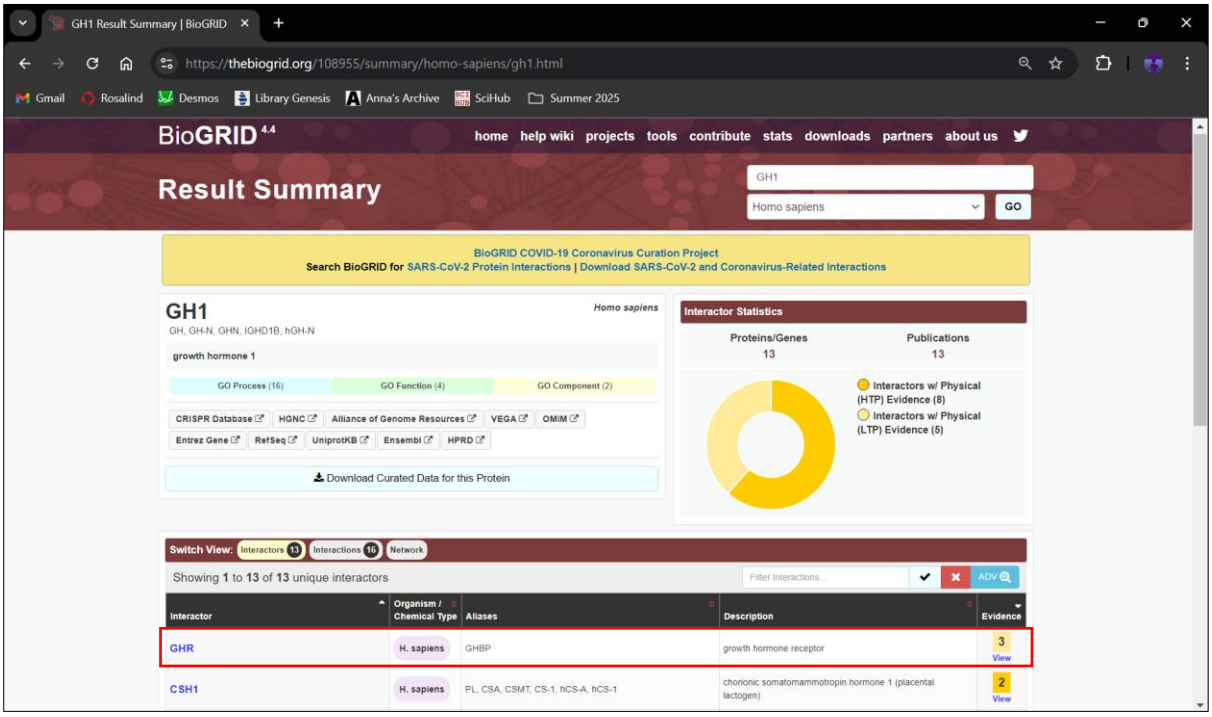
PDB ID : 1A22.pdb

Residue name	Residue number	Chain	Contacting residue	Residue number	Chain	Distance
PRO	5	A	VAL	185	A	7.793
LEU	6	A	SER	184	A	7.307
LEU	6	A	VAL	185	A	6.728
SER	7	A	VAL	185	A	6.083
LEU	9	A	ARG	120	A	7.693
LEU	9	A	LEU	124	A	7.660
PHE	10	A	LEU	124	A	7.918
PHE	10	A	GLN	181	A	6.439
ALA	13	A	LEU	117	A	6.070
ALA	13	A	ARG	120	A	6.223
ALA	13	A	ILE	121	A	7.325
MET	14	A	GLU	174	A	7.012
MET	14	A	LEU	177	A	7.089
MET	14	A	ARG	178	A	7.135
ARG	16	A	LEU	113	A	6.638
ARG	16	A	ASP	116	A	7.959
ARG	16	A	LEU	117	A	6.482
ALA	17	A	LEU	117	A	7.174
ALA	17	A	MET	170	A	7.151
ALA	17	A	VAL	173	A	6.810
ALA	17	A	GLU	174	A	5.214
HIS	18	A	GLU	174	A	6.250
HIS	18	A	ASN	418	B	7.827

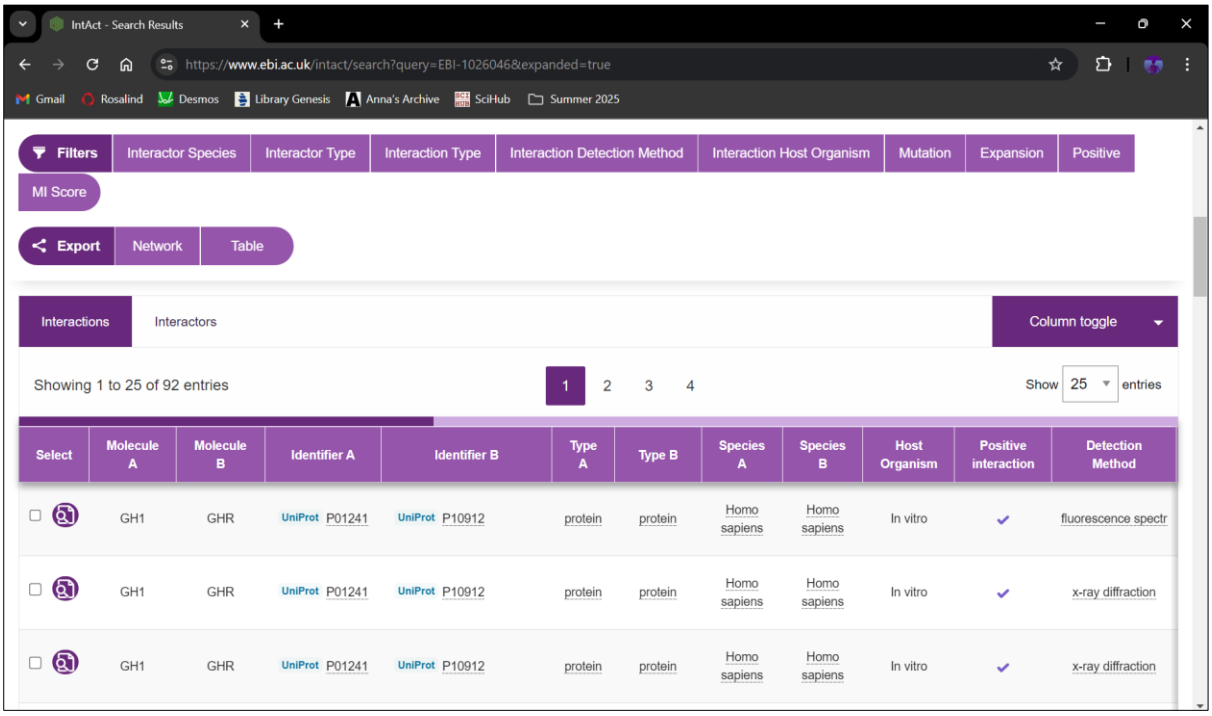
The interaction between the human growth hormone (GH1 / somatotropin) and its receptor was found in the Database of Interacting Proteins (DIP).



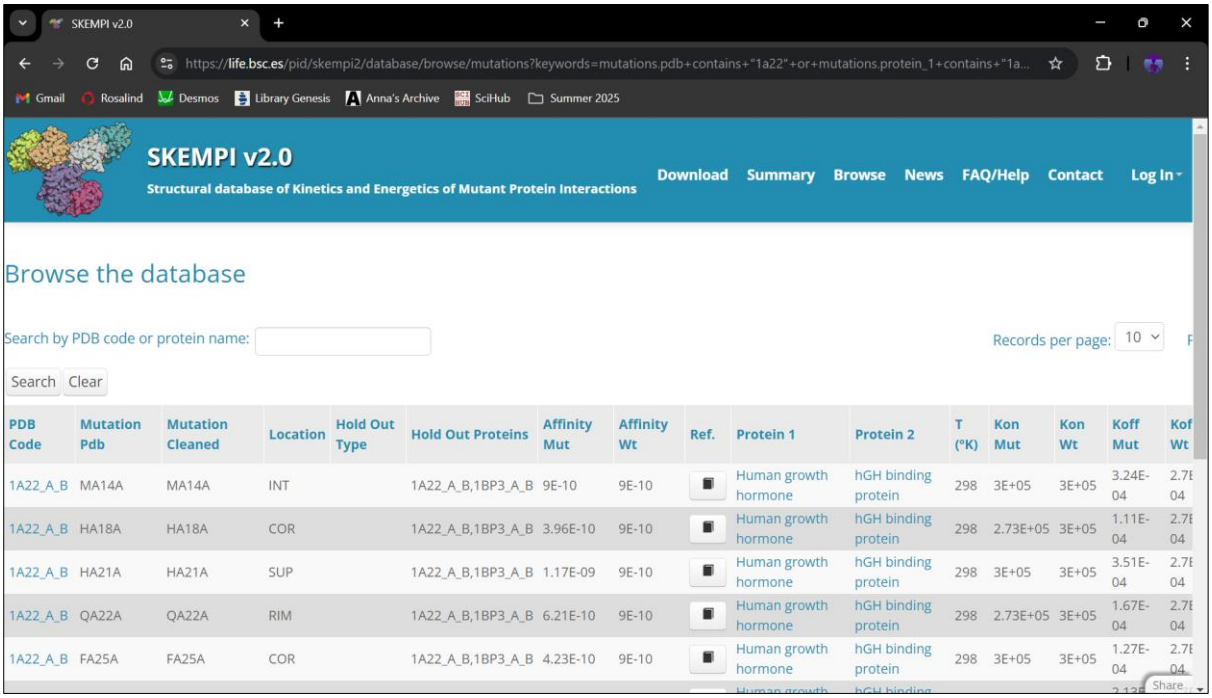
This interaction was also found in the BioGRID database.



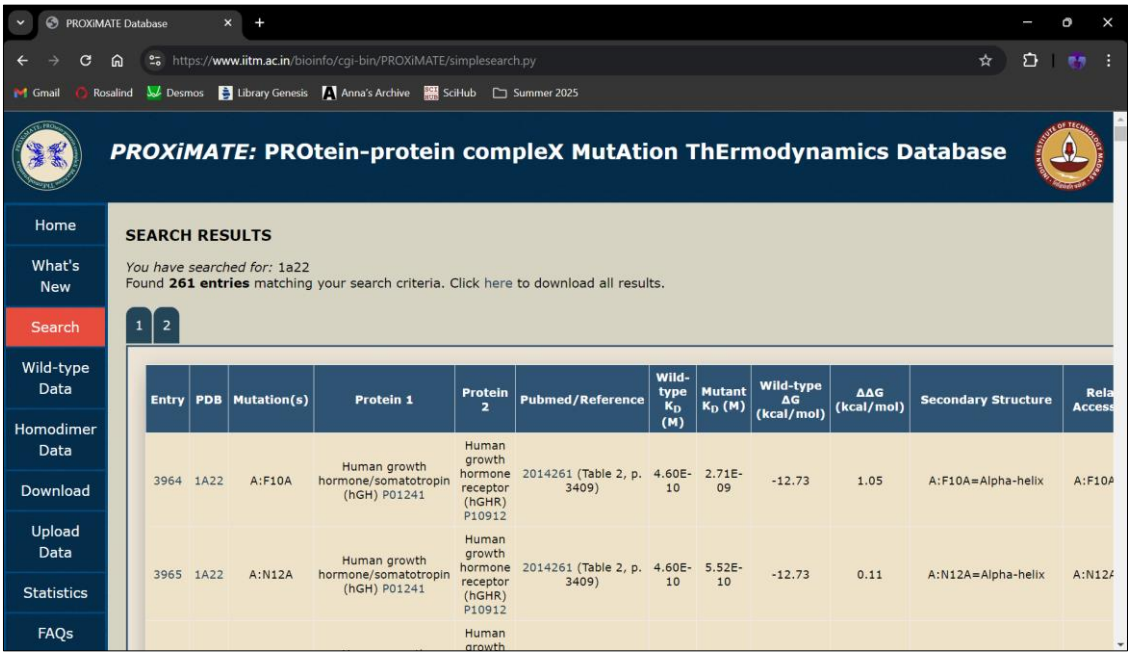
EBI's IntAct database also contained information about the interaction between the growth hormone and its receptor.



SKEMPI contains data on the changes in thermodynamic parameters and kinetic rate constants upon mutation of the protein-protein complex. There were 251 records found for this complex.



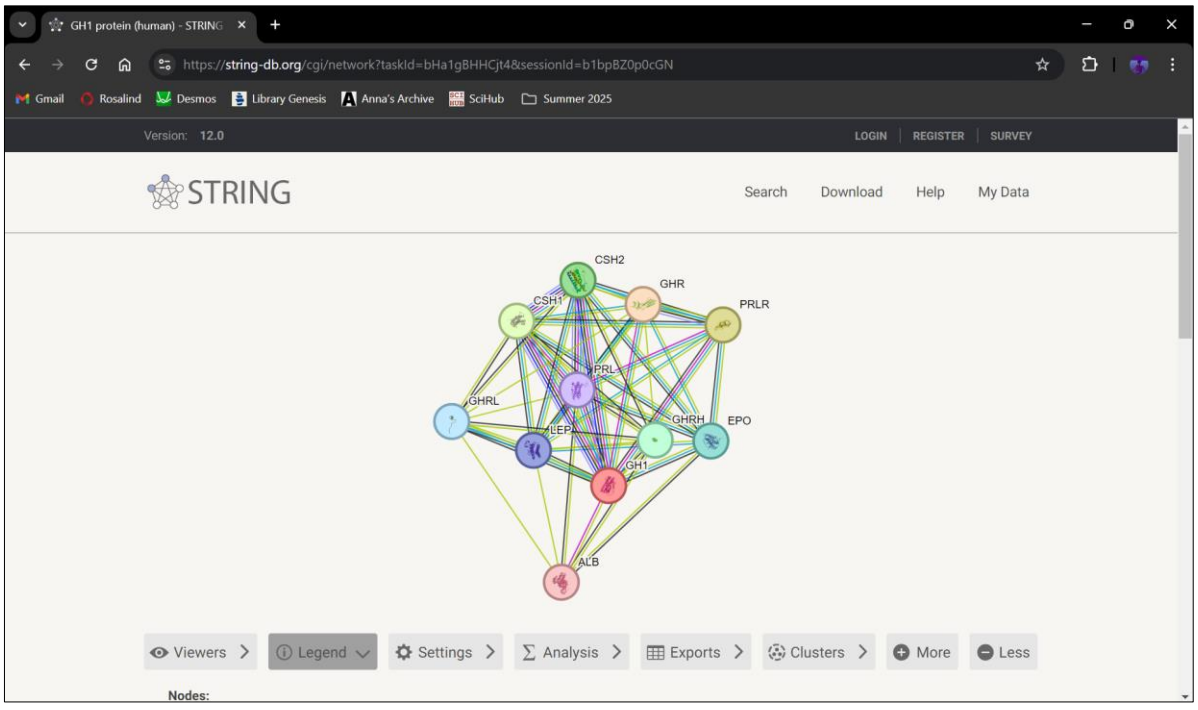
The PROXiMATE database contains 261 datapoints for the thermodynamics of mutant complexes.



The PDBbind database, Kinetic Data of Biomolecular Interaction (KDBI), and the Alanine Scanning Energetics database (ASEdb) were unresponsive and could not be queried for the complex.

6

STRING was used to construct the protein interaction network for the proteins in the complex. For the growth hormone, 10 interactors were predicted:



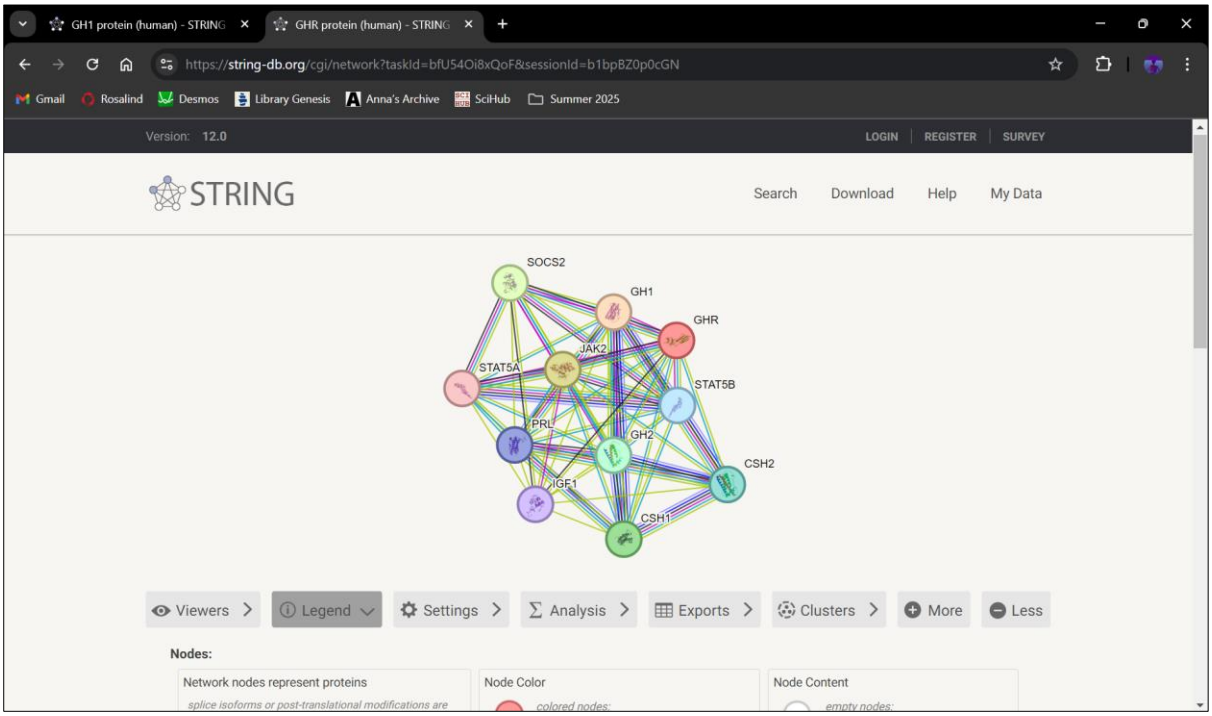
Your Input:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
GH1	*Somatotropin; Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues. (217 aa)*									

Predicted Functional Partners:

GHR	*Growth hormone-binding protein; Receptor for pituitary gland growth hormone involved in regulating postnatal body growth. On li...*					●	●	●		0.999
PRLR	*Prolactin receptor; This is a receptor for the anterior pituitary hormone prolactin (PRL). Acts as a prosurvival factor for spermatoz...*				●	●	●	●		0.993
CSH1	*Chorionic somatomammotropin hormone 1; Produced only during pregnancy and is involved in stimulating lactation, fetal growt...*				●	●	●	●	●	0.990
CSH2	*Chorionic somatomammotropin hormone 2; Produced only during pregnancy and is involved in stimulating lactation, fetal growt...*				●	●	●	●	●	0.989
GHRH	*Somatoliberin; GRF is released by the hypothalamus and acts on the adenohypophyse to stimulate the secretion of growth horm...*				●			●		0.964
EPO	*Erythropoietin; Hormone involved in the regulation of erythrocyte proliferation and differentiation and the maintenance of a physi...*						●	●		0.959
GHRL	*Appetite-regulating hormone; Ghrelin is the ligand for growth hormone secretagogue receptor type 1 (GHSR). Induces the release...*				●		●	●		0.956
LEP	*Leptin; Key player in the regulation of energy balance and body weight control. Once released into the circulation, has central and...*						●	●		0.949
PRL	*Prolactin; Prolactin acts primarily on the mammary gland by promoting lactation; Belongs to the somatotropin/prolactin family.*				●	●	●	●	●	0.944
ALB	*Serum albumin; Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acid...*					●		●		0.936

For the receptor, 10 interactors were predicted:



Your Input:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
GHR	*Growth hormone-binding protein; Receptor for pituitary gland growth hormone involved in regulating postnatal body growth. On ligand binding, couples to the JAK2/STAT5 pathway (By similarity). Isoform 2 up-regulates the production of GHBP and acts as a negative inhibitor of GH signaling. (645 aa)*									

Predicted Functional Partners:

GH1	*Somatotropin; Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and o...*					●	●	●		0.999
JAK2	*Tyrosine-protein kinase JAK2; Non-receptor tyrosine kinase involved in various processes such as cell growth, development, dif...*				●	●	●	●		0.999
SOCS2	*Suppressor of cytokine signaling 2; SOCS family proteins form part of a classical negative feedback system that regulates cyto...*				●	●	●	●		0.993
CSH1	*Chorionic somatomammotropin hormone 1; Produced only during pregnancy and is involved in stimulating lactation, fetal grow...*						●	●		0.992
GH2	*Chorionic somatomammotropin hormone 1 isoform 2; Growth hormone 2.*						●	●		0.991
CSH2	*Chorionic somatomammotropin hormone 2; Produced only during pregnancy and is involved in stimulating lactation, fetal grow...*						●	●		0.991
STAT5B	*Signal transducer and activator of transcription 5B; Carries out a dual function: signal transduction and activation of transcripti...*					●	●	●		0.977
PRL	*Prolactin; Prolactin acts primarily on the mammary gland by promoting lactation; Belongs to the somatotropin/prolactin family.*						●	●		0.970
IGF1	*Insulin-like growth factor I; The insulin-like growth factors, isolated from plasma, are structurally and functionally related to insu...*				●			●		0.956
STAT5A	*Signal transducer and activator of transcription 5A; Carries out a dual function: signal transduction and activation of transcripti...*					●	●	●		0.926

- The interaction score for the hormone and its receptors was 0.999, which is a high score.
- They both interact with CSH1 and CSH2, which are foetal growth hormones produced during pregnancy.
- The network statistics for both networks were similar.

  For the hormone,

Network Stats

number of nodes:	11	expected number of edges:	13
number of edges:	47	PPI enrichment p-value:	1.11e-13
average node degree:	8.55		
avg. local clustering coefficient:	0.877		

your network has significantly more interactions than expected (what does that mean?)

  For the receptor,

Network Stats

number of nodes:	11	expected number of edges:	12
number of edges:	52	PPI enrichment p-value:	< 1.0e-16
average node degree:	9.45		
avg. local clustering coefficient:	0.953		

your network has significantly more interactions than expected (what does that mean?)

- The functional enrichments for both networks yielded similar results, as expected.

  For the hormone,

Functional enrichments in your network

Note: some enrichments may be expected here (why?)

explain columns

Biological Process (Gene Ontology)

GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:0042320	Regulation of circadian sleep/wake cycle, REM sleep	2 of 4	2.95	0.00091
GO:0045938	Positive regulation of circadian sleep/wake cycle, sleep	2 of 5	2.86	0.0012
GO:0042976	Activation of Janus kinase activity	2 of 6	2.78	0.0015
GO:0030252	Growth hormone secretion	2 of 6	2.78	0.0015
GO:0060124	Positive regulation of growth hormone secretion	2 of 10	2.55	0.0031

(more ...)

Molecular Function (Gene Ontology)

GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:0016608	Growth hormone-releasing hormone activity	2 of 2	3.25	0.00084
GO:0005148	Prolactin receptor binding	2 of 2	3.25	0.00084
GO:0005131	Growth hormone receptor binding	3 of 11	2.69	2.89e-05
GO:0051427	Hormone receptor binding	5 of 33	2.43	2.29e-08
GO:0051428	Peptide hormone receptor binding	2 of 19	2.28	0.0224

(more ...)

  For the receptor,

Functional enrichments in your network

Note: some enrichments may be expected here (why?)

explain columns

Biological Process (Gene Ontology)

GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:0060396	Growth hormone receptor signaling pathway	9 of 21	2.89	9.97e-21
GO:0019530	Taurine metabolic process	3 of 7	2.89	6.23e-06
GO:0042976	Activation of Janus kinase activity	2 of 6	2.78	0.0014
GO:0060397	Growth hormone receptor signaling pathway via JAK-STAT	3 of 10	2.73	1.31e-05
GO:0046427	Positive regulation of receptor signaling pathway via JAK-ST...	7 of 45	2.44	4.28e-13

(more ...)

Molecular Function (Gene Ontology)

GO-term	*description*	*count in network*	*strength*	*false discovery rate*
GO:0005148	Prolactin receptor binding	2 of 2	3.25	0.00084
GO:0005131	Growth hormone receptor binding	6 of 11	2.99	3.46e-13
GO:0005159	Insulin-like growth factor receptor binding	2 of 16	2.35	0.0177
GO:0005179	Hormone activity	6 of 129	1.92	5.19e-08
GO:0008083	Growth factor activity	5 of 162	1.74	1.80e-05

(more ...)

PROXiMATE was used to obtain the wild-type thermodynamic data of the complex.

							sodium azide)			
125	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 0.1% BSA, 0.02% NaN_3	298	7.4	2014261 (Ta p. 340
126	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 0.1% BSA, 0.02% NaN_3, 50 µM $ZnCl_2$, 10 mM $MgCl_2$	298	7.4	2014261 (Ta p. 340
127	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 1 g/L BSA, 0.2 g/L NaN_3	298	7.4	2034689 (Ta p. 449
128	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 1 g/L BSA, 0.2 g/L NaN_3	298	7.4	2471267 (Ta p. 108
129	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Surface Plasmon Resonance	150 mM NaCl, 10 mM sodium phosphate, 0.02% Tween 20	298	7.4	7504735 (Ta and 2, pp. 5
130	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive Binding	50mM Tris-HCl, 10mM $CaCl_2$, 1g/L bovine albumin, 0.2 g/L NaN_3	298	7.4	7529940 (Fi p. 384
131	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Surface Plasmon Resonance	PBS with 10mM sodium phosphate, 137mM NaCl, 2.7mM KCl, 0.05% Tween 20	298	7.2	8756685 (Ta p. 1030
132	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	competitive displacement of hGH from hGHbp, using	PBS with 0.02% (v/v) Tween 20	298		9571026 (Ta p. 111

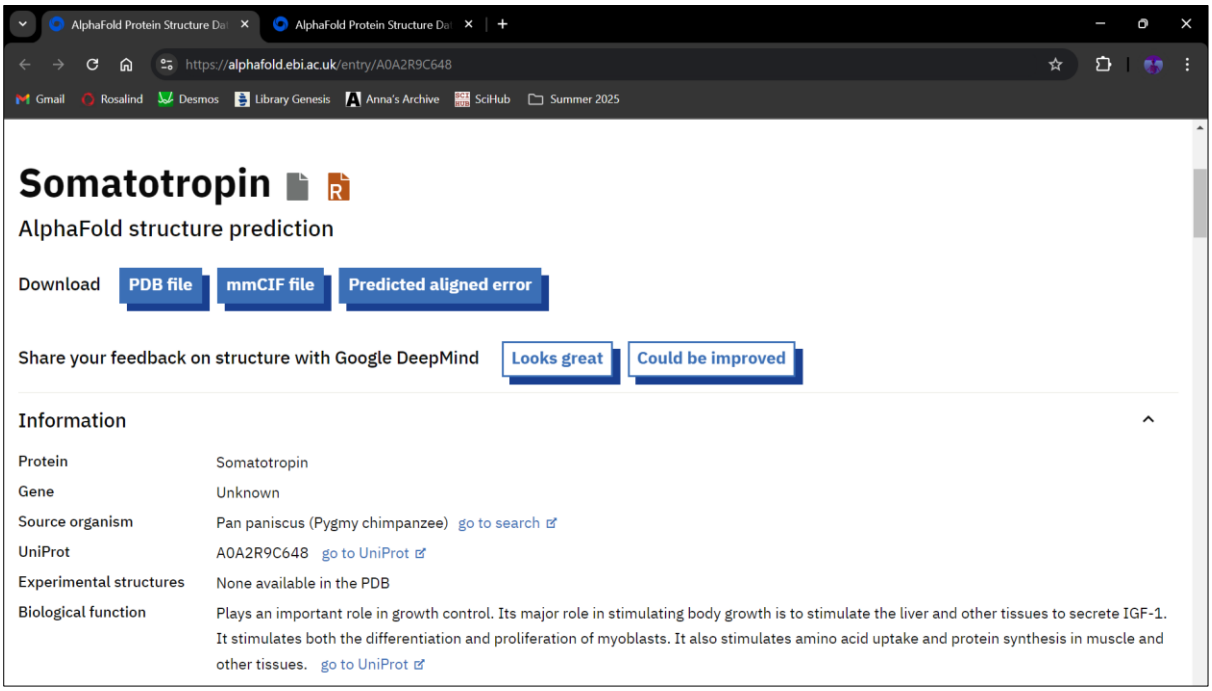
The results are summarised below:

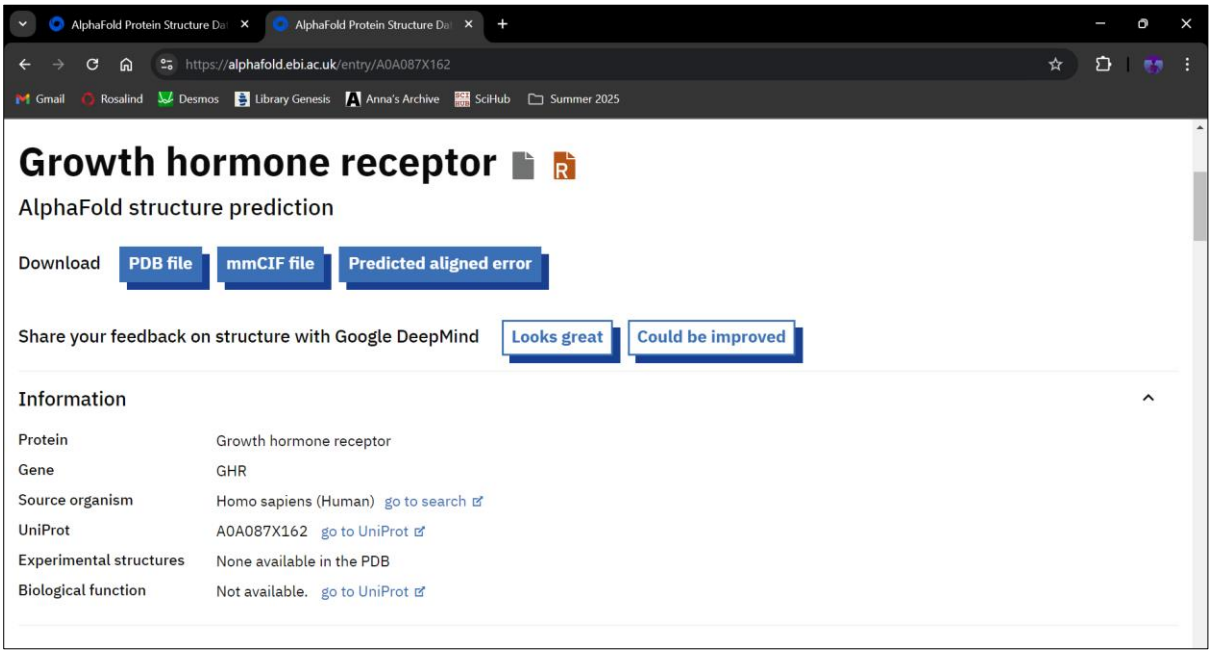
S.no.	PDB	Dissociation constant (M)	ΔG (kcal/mol)
125	1A22	4.60E-10	-12.73
126	1A22	4.40E-10	-12.76
127	1A22	4.00E-10	-12.81
128	1A22	3.40E-10	-12.91
129	1A22	9.00E-10	-12.33
130	1A22	9.63E-10	-12.29
131	1A22	2.90E-09	-11.64
132	1A22	3.40E-10	-12.91

All the datapoints were obtained at 298 K at a pH in the range 7.2-7.4.

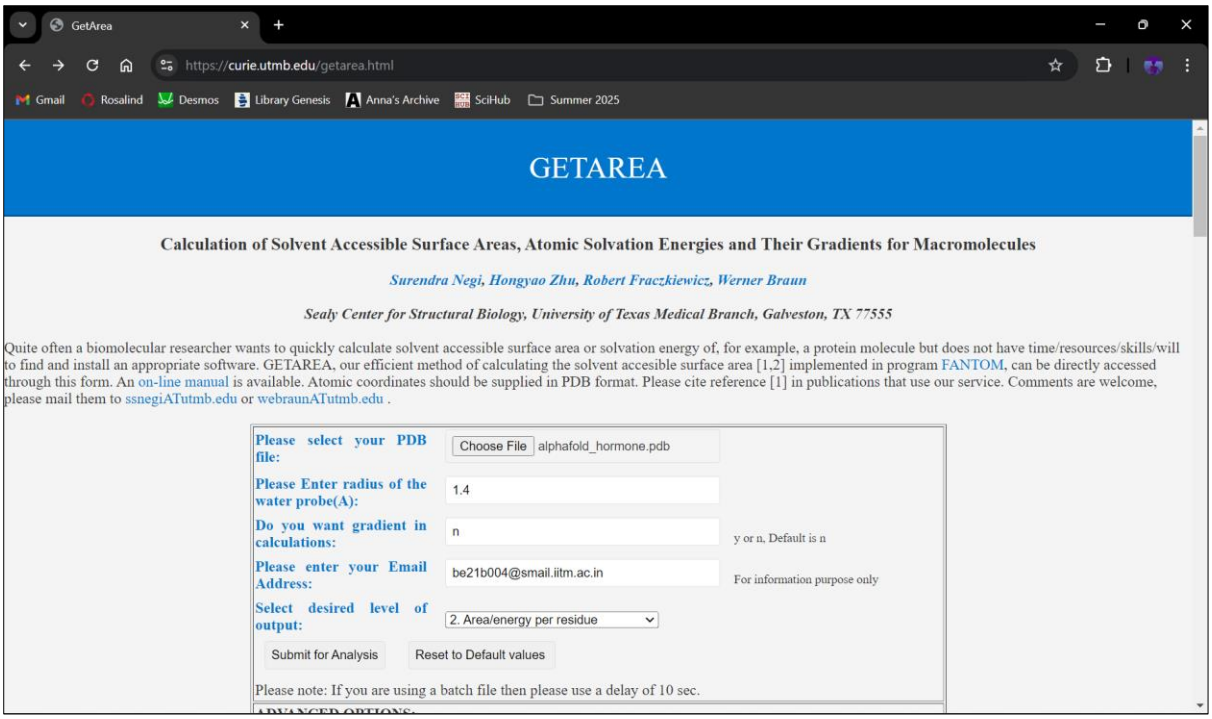
On average, the dissociation constant is 6.74×10^{-9} M and the free energy change on binding is −12.55 kcal/mol.

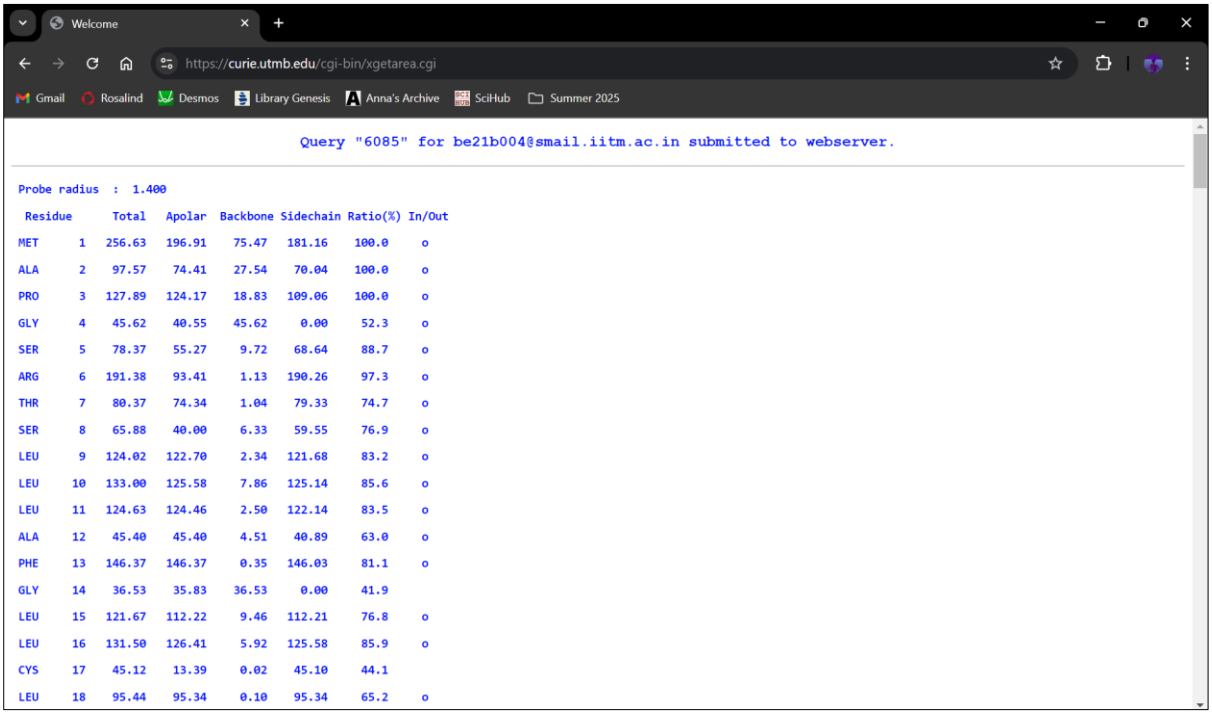
The .fasta sequences of the two proteins in the complex were obtained from PDB. These were individually fed into the AlphaFold database to obtain their separate structures in .pdb format.





Next, the residue-level solvent accessible surface areas were computed for the individual proteins and the complex using the GETAREA webserver.





If the total ASA of a residue in a protein reduces by more than 0.1 $Å^2$ in the complex, that residue can be considered an interface residue.

The ASA data can be found as a supplementary file to this report.

Based on this analysis, the following residues were found to be interface residues:

Hormone

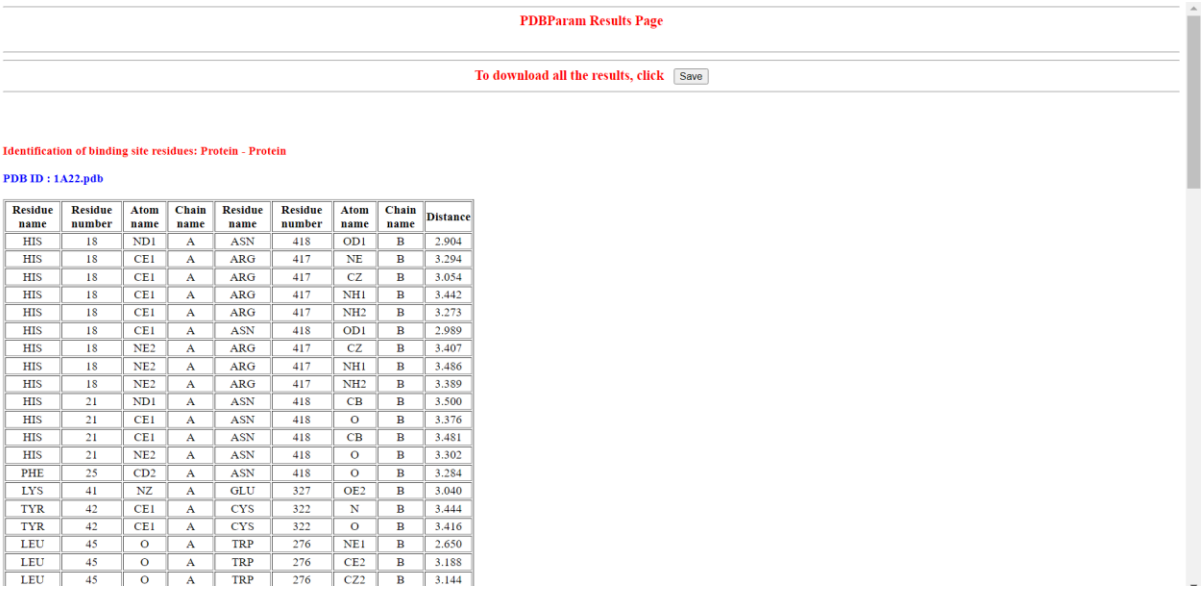
PRO2	LEU6	ASN12	ALA13	LEU15	ARG16
ALA17	HIS18	ARG19	LEU20	HIS21	GLN22
LEU23	ALA24	PHE25	ASP26	THR27	TYR28
ILE36	PRO37	LYS38	GLU39	GLN40	LYS41
TYR42	SER43	PHE44	LEU45	GLN46	PRO48
SER51	LEU52	PHE54	GLU56	SER57	ILE58
PRO59	PRO61	SER62	ASN63	ARG64	GLU65
THR67	GLN68	GLN69	LEU73	GLU74	ARG77
ILE78	LEU81	GLN84	SER85	VAL90	GLN91
LEU93	ARG94	PHE97	ALA98	ASN99	SER100
LEU101	VAL102	GLY104	ALA105	SER106	ASP107
ASN109	VAL110	TYR111	LEU113	ASP116	GLU119
GLN122	GLY126	LEU128	GLN137	ILE138	LYS140
GLN141	SER144	TYR160	LEU162	LEU163	TYR164
PHE166	ARG167	LYS168	MET170	ASP171	LYS172
THR175	PHE176	ARG178	ILE179	GLN181	CYS182
SER184	VAL185	GLU186	GLY187	SER188	CYS189
GLY190	PHE191				

Receptor

PHE235	LYS237	CYS238	PRO241	GLU242	ARG243
GLU244	THR245	SER247	TRP250	ILE264	GLN265
ARG270	ARG271	ASN272	THR273	GLN274	GLU275
TRP276	THR277	GLN278	LYS281	GLU282	ASP285
SER288	ALA289	GLY290	GLU291	ASN297	SER298
SER299	PHE300	THR301	SER302	ILE303	TRP304
ILE305	PRO306	CYS308	GLY315	THR317	ASP319
GLU320	LYS321	CYS322	PHE323	SER324	ASP326
GLU327	ILE328	VAL329	GLN330	ASP332	PRO334
ALA336	ASN338	TRP339	THR340	LEU341	LEU342
ARG356	TRP357	GLU358	ARG361	ASN362	ALA363
ASP364	ILE365	GLN366	LYS367	GLY368	TRP369
VAL371	GLU373	GLU375	LEU376	TYR378	LYS379
GLU380	VAL381	ASN382	GLU383	LYS385	TRP386
LYS387	MET388	MET389	ASP390	ILE392	THR394
THR395	SER401	LYS403	VAL404	ASP405	GLU407
TYR408	GLU409	GLN416	ARG417	ASN418	SER419
GLY420	PHE425	GLU427	VAL428	LEU429	GLN435

Here, the numbering of residues is according to the original .pdb file of the complex.

The distance-based approach to identify binding sites was carried out using PDBparam with the default distance threshold of 3.5 Å.



PDBParam Results Page

To download all the results, click Save

Identification of binding site residues: Protein - Protein

PDB ID : 1A22.pdb

Residue name	Residue number	Atom name	Chain name	Residue name	Residue number	Atom name	Chain name	Distance
HIS	18	ND1	A	ASN	418	OD1	B	2.904
HIS	18	CE1	A	ARG	417	NE	B	3.294
HIS	18	CE1	A	ARG	417	CZ	B	3.054
HIS	18	CE1	A	ARG	417	NH1	B	3.442
HIS	18	CE1	A	ARG	417	NH2	B	3.273
HIS	18	CE1	A	ASN	418	OD1	B	2.989
HIS	18	NE2	A	ARG	417	CZ	B	3.407
HIS	18	NE2	A	ARG	417	NH1	B	3.486
HIS	18	NE2	A	ARG	417	NH2	B	3.389
HIS	21	ND1	A	ASN	418	CB	B	3.500
HIS	21	CE1	A	ASN	418	O	B	3.376
HIS	21	CE1	A	ASN	418	CB	B	3.481
HIS	21	NE2	A	ASN	418	O	B	3.302
PHE	25	CD2	A	ASN	418	O	B	3.284
LYS	41	NZ	A	GLU	327	OE2	B	3.040
TYR	42	CE1	A	CYS	322	N	B	3.444
TYR	42	CE1	A	CYS	322	O	B	3.416
LEU	45	O	A	TRP	276	NE1	B	2.650
LEU	45	O	A	TRP	276	CE2	B	3.188
LEU	45	O	A	TRP	276	CZ2	B	3.144

Residue name	Residue number	Atom name	Chain name	Residue name	Residue number	Atom name	Chain name	Distance
HIS	18	ND1	A	ASN	418	OD1	B	2.904
HIS	18	CE1	A	ARG	417	NE	B	3.294
HIS	18	CE1	A	ARG	417	CZ	B	3.054
HIS	18	CE1	A	ARG	417	NH1	B	3.442
HIS	18	CE1	A	ARG	417	NH2	B	3.273
HIS	18	CE1	A	ASN	418	OD1	B	2.989
HIS	18	NE2	A	ARG	417	CZ	B	3.407
HIS	18	NE2	A	ARG	417	NH1	B	3.486
HIS	18	NE2	A	ARG	417	NH2	B	3.389
HIS	21	ND1	A	ASN	418	CB	B	3.500
HIS	21	CE1	A	ASN	418	O	B	3.376
HIS	21	CE1	A	ASN	418	CB	B	3.481
HIS	21	NE2	A	ASN	418	O	B	3.302
PHE	25	CD2	A	ASN	418	O	B	3.284
LYS	41	NZ	A	GLU	327	OE2	B	3.040
TYR	42	CE1	A	CYS	322	N	B	3.444
TYR	42	CE1	A	CYS	322	O	B	3.416
LEU	45	O	A	TRP	276	NE1	B	2.650
LEU	45	O	A	TRP	276	CE2	B	3.188

LEU	45	O	A	TRP	276	CZ2	B	3.144
GLN	46	CA	A	TRP	276	NE1	B	3.217
GLN	46	NE2	A	CYS	308	SG	B	3.424
GLN	46	NE2	A	GLU	320	OE2	B	3.047
GLN	46	NE2	A	CYS	322	SG	B	3.491
LEU	52	CD1	A	ARG	271	CZ	B	3.430
LEU	52	CD1	A	ARG	271	NH2	B	3.210
GLU	56	CD	A	ARG	271	NH2	B	3.437
GLU	56	OE2	A	ARG	271	NH2	B	2.883
SER	62	O	A	SER	302	CA	B	3.500
SER	62	O	A	ILE	303	N	B	2.713
ARG	64	N	A	TRP	369	CZ2	B	3.433
ARG	64	NE	A	ASP	364	OD2	B	2.899
ARG	64	CZ	A	ASP	364	OD2	B	3.425
ARG	64	NH2	A	ASP	364	OD2	B	3.269
ARG	64	NH2	A	LYS	367	NZ	B	3.044
GLN	68	NE2	A	ASP	364	OD2	B	2.982
GLN	68	NE2	A	LYS	367	O	B	3.398
ARG	167	NH2	A	GLU	327	OE2	B	3.041
LYS	168	NZ	A	TRP	304	O	B	3.001
ASP	171	OD1	A	ARG	243	NH1	B	3.324
ASP	171	OD2	A	ARG	243	NH1	B	3.187
ASP	171	OD2	A	TRP	304	NE1	B	3.390
GLU	174	OE1	A	ASN	418	OD1	B	3.405
THR	175	OG1	A	ARG	243	NH2	B	2.646
ARG	178	CG	A	GLY	368	O	B	3.218
CYS	182	SG	A	GLY	368	N	B	3.468
ARG	243	NH1	B	ASP	171	OD1	A	3.324
ARG	243	NH1	B	ASP	171	OD2	A	3.187
ARG	243	NH2	B	THR	175	OG1	A	2.646
ARG	271	CZ	B	LEU	52	CD1	A	3.430
ARG	271	NH2	B	LEU	52	CD1	A	3.210
ARG	271	NH2	B	GLU	56	CD	A	3.437
ARG	271	NH2	B	GLU	56	OE2	A	2.883
TRP	276	NE1	B	LEU	45	O	A	2.650
TRP	276	NE1	B	GLN	46	CA	A	3.217
TRP	276	CE2	B	LEU	45	O	A	3.188
TRP	276	CZ2	B	LEU	45	O	A	3.144

SER	302	CA	B	SER	62	O	A	3.500
ILE	303	N	B	SER	62	O	A	2.713
TRP	304	O	B	LYS	168	NZ	A	3.001
TRP	304	NE1	B	ASP	171	OD2	A	3.390
CYS	308	SG	B	GLN	46	NE2	A	3.424
GLU	320	OE2	B	GLN	46	NE2	A	3.047
CYS	322	N	B	TYR	42	CE1	A	3.444
CYS	322	O	B	TYR	42	CE1	A	3.416
CYS	322	SG	B	GLN	46	NE2	A	3.491
GLU	327	OE2	B	LYS	41	NZ	A	3.040
GLU	327	OE2	B	ARG	167	NH2	A	3.041
ASP	364	OD2	B	ARG	64	NE	A	2.899
ASP	364	OD2	B	ARG	64	CZ	A	3.425
ASP	364	OD2	B	ARG	64	NH2	A	3.269
ASP	364	OD2	B	GLN	68	NE2	A	2.982
LYS	367	O	B	GLN	68	NE2	A	3.398
LYS	367	NZ	B	ARG	64	NH2	A	3.044
GLY	368	N	B	CYS	182	SG	A	3.468
GLY	368	O	B	ARG	178	CG	A	3.218
TRP	369	CZ2	B	ARG	64	N	A	3.433
ARG	417	NE	B	HIS	18	CE1	A	3.294
ARG	417	CZ	B	HIS	18	CE1	A	3.054
ARG	417	CZ	B	HIS	18	NE2	A	3.407
ARG	417	NH1	B	HIS	18	CE1	A	3.442
ARG	417	NH1	B	HIS	18	NE2	A	3.486
ARG	417	NH2	B	HIS	18	CE1	A	3.273
ARG	417	NH2	B	HIS	18	NE2	A	3.389
ASN	418	O	B	HIS	21	CE1	A	3.376
ASN	418	O	B	HIS	21	NE2	A	3.302
ASN	418	O	B	PHE	25	CD2	A	3.284
ASN	418	CB	B	HIS	21	ND1	A	3.500
ASN	418	CB	B	HIS	21	CE1	A	3.481
ASN	418	OD1	B	HIS	18	ND1	A	2.904
ASN	418	OD1	B	HIS	18	CE1	A	2.989
ASN	418	OD1	B	GLU	174	OE1	A	3.405

.

9

Using the solvent accessibility method's results, the propensities of the residues to be binding site residues was computed. A Python script was used for this.

Hormone

HIS	100
GLN	84.62
PRO	71.43
ALA	71.43
GLY	66.67
ARG	63.64
SER	62.5
TYR	62.5
ILE	62.5
VAL	57.14
ASN	57.14
LYS	55.56
LEU	53.85
PHE	53.85
CYS	50
ASP	44.44
GLU	42.86
THR	33.33
MET	33.33

Receptor

GLU	88.24
TRP	87.5
ASP	87.5
GLN	77.78
LYS	66.67
ARG	66.67
ILE	60
ASN	60
ALA	60
THR	57.14
PHE	57.14
SER	53.33
GLY	50
CYS	50
MET	50
VAL	35.71
LEU	33.33
PRO	25
TYR	20

The residues with highest propensity to be binding site residues are charged or polar. This enables them to form interactions with the binding partner.

10

Using the list of interactions obtained from PDBsum, all residue pairs were sorted in decreasing order of number of contacts. Based on this analysis, which was carried out using a Python script, the top interacting pairs in the complex were found to be:

Hormone	Receptor	No. of Contacts
HIS18	ARG417	11
ARG64	TRP369	11
ARG64	ASP364	9
HIS18	ASN418	8
LEU45	TRP276	7
GLN46	TRP276	7

PROXiMATE was used to obtain the wild-type thermodynamic data of the complex.

								sodium azide)			
125	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 0.1% BSA, 0.02% NaN_3	298	7.4	2014261 (Ta p. 340
126	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 0.1% BSA, 0.02% NaN_3, 50 µM $ZnCl_2$, 10 mM $MgCl_2$	298	7.4	2014261 (Ta p. 340
127	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 1 g/L BSA, 0.2 g/L NaN_3	298	7.4	2034689 (Ta p. 449
128	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive displacement of ^{125}I-labeled hGH	50 mM Tris-HCl, 10 mM $CaCl_2$, 1 g/L BSA, 0.2 g/L NaN_3	298	7.4	2471267 (Ta p. 108
129	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Surface Plasmon Resonance	150 mM NaCl, 10 mM sodium phosphate, 0.02% Tween 20	298	7.4	7504735 (Ta and 2, pp. 5
130	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Competitive Binding	50mM Tris-HCl, 10mM $CaCl_2$, 1g/L bovine albumin, 0.2 g/L NaN_3	298	7.4	7529940 (Fi p. 384
131	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	Surface Plasmon Resonance	PBS with 10mM sodium phosphate, 137mM NaCl, 2.7mM KCl, 0.05% Tween 20	298	7.2	8756685 (Ta p. 1030
132	1A22	A:B	Human growth hormone/somatotropin (hGH)	Human growth hormone receptor (hGHR)	HR	competitive displacement of hGH from hGHbp, using	PBS with 0.02% (v/v) Tween 20	298		9571026 (Ta p. 111

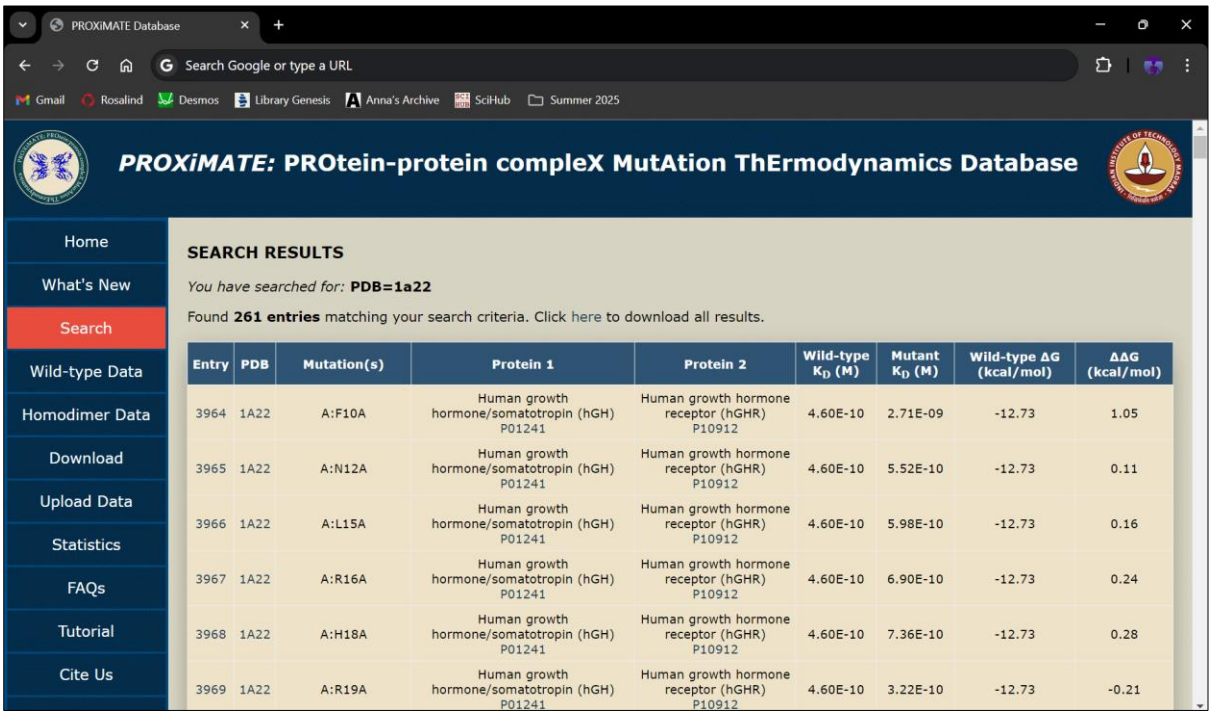
The results are summarised below:

S.no.	PDB	Dissociation constant (M)	ΔG (kcal/mol)
125	1A22	4.60E-10	-12.73
126	1A22	4.40E-10	-12.76
127	1A22	4.00E-10	-12.81
128	1A22	3.40E-10	-12.91
129	1A22	9.00E-10	-12.33
130	1A22	9.63E-10	-12.29
131	1A22	2.90E-09	-11.64
132	1A22	3.40E-10	-12.91

All the datapoints were obtained at 298 K at a pH in the range 7.2-7.4.

On average, the dissociation constant is 6.74×10^{-9} M and the free energy change on binding is −12.55 kcal/mol.

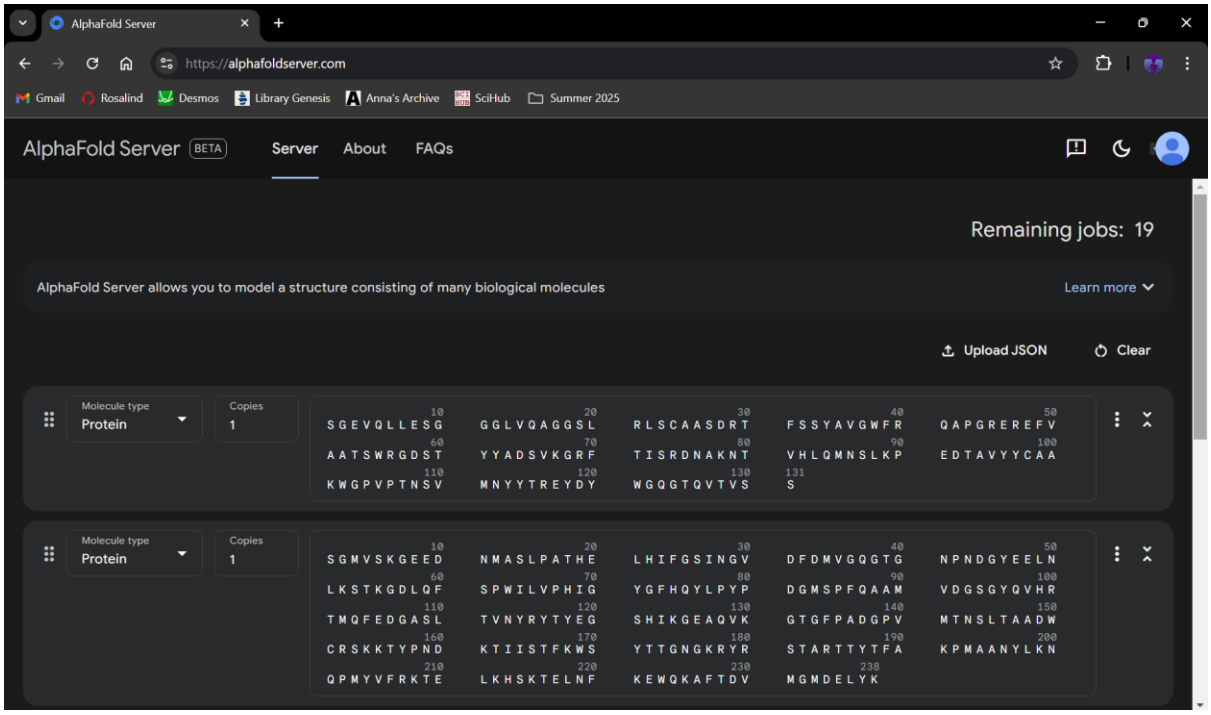
PROXiMATE was then used to obtain the thermodynamic data of the mutant forms of the complex. 261 entries were found in the database.

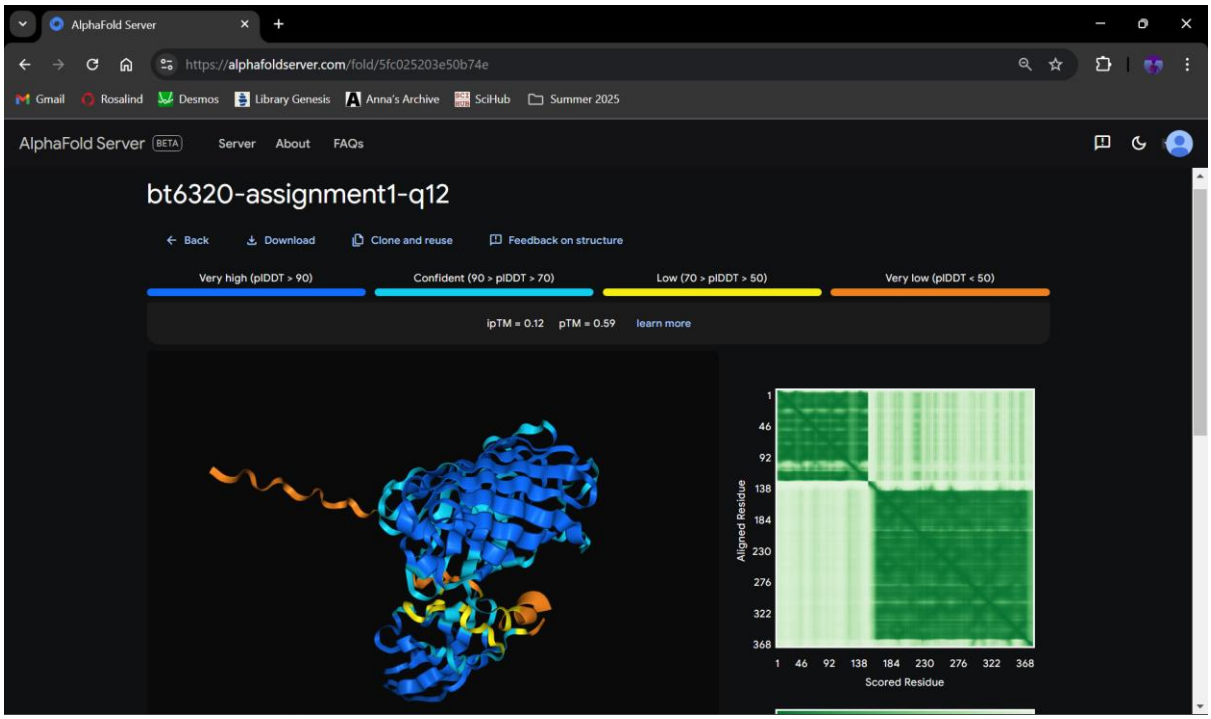


Entry	PDB	Mutation(s)	Protein 1	Protein 2	Wild-type K_D (M)	Mutant K_D (M)	Wild-type ΔG (kcal/mol)	ΔΔG (kcal/mol)
3964	1A22	A:F10A	Human growth hormone/somatotropin (hGH) P01241	Human growth hormone receptor (hGHR) P10912	4.60E-10	2.71E-09	-12.73	1.05
3965	1A22	A:N12A	Human growth hormone/somatotropin (hGH) P01241	Human growth hormone receptor (hGHR) P10912	4.60E-10	5.52E-10	-12.73	0.11
3966	1A22	A:L15A	Human growth hormone/somatotropin (hGH) P01241	Human growth hormone receptor (hGHR) P10912	4.60E-10	5.98E-10	-12.73	0.16
3967	1A22	A:R16A	Human growth hormone/somatotropin (hGH) P01241	Human growth hormone receptor (hGHR) P10912	4.60E-10	6.90E-10	-12.73	0.24
3968	1A22	A:H18A	Human growth hormone/somatotropin (hGH) P01241	Human growth hormone receptor (hGHR) P10912	4.60E-10	7.36E-10	-12.73	0.28
3969	1A22	A:R19A	Human growth hormone/somatotropin (hGH) P01241	Human growth hormone receptor (hGHR) P10912	4.60E-10	3.22E-10	-12.73	-0.21

- On average, the dissociation constant of the mutants is 4.6×10^{-9} M.
- Of the 261 mutants, 70 are stabilising mutants ($\Delta\Delta G < 0$), and 191 are destabilising mutants ($\Delta\Delta G > 0$).
- The most stabilising mutation is a E174A mutation in the hormone, which has a $\Delta\Delta G = -0.92$ kcal/mol.
- The most destabilising mutation is a double K172A and F176A mutation in the hormone, which has a $\Delta\Delta G = 3.78$ kcal/mol.

The AlphaFold3 webserver was used to predict the structure of the protein-protein complex.





- While the model predicts majority of the structure with high probability, there are some residues whose structure has very low probability.
- Both of the proteins have a characteristic beta barrel structure.
- The interface residues between the two barrels are predicted with low probability.